

# Stanley Z. Hua

Berkeley, California

(+1) 4379863444 | [stanley.z.hua@gmail.com](mailto:stanley.z.hua@gmail.com) | [linkedin](#) | [github](#) | [website](#)

## EDUCATION

---

<b>University of California, Berkeley</b> <i>PhD Computational Precision Health</i> (Primary Advisor: Irene Chen)	Berkeley, USA <i>Aug. 2025 – Present</i>
<b>University of California, San Francisco</b> <i>PhD Computational Precision Health</i> (Secondary Advisor: Jean Feng)	San Francisco, USA <i>Aug. 2025 – Present</i>
<b>University of Toronto</b> <i>Honours BS Computer Science Specialist</i> (GPA: 3.84/4)	Toronto, Canada <i>Sept. 2022 – June 2024</i>
<b>University of Toronto</b> <i>Honours BS Bioinformatics &amp; Computational Biology Specialist</i>	Toronto, Canada <i>Sept. 2019 – Sept. 2022</i> (Switched Majors)

## TECHNICAL SKILLS

---

**Languages:** Python, SQL, Shell Script, Git, Javascript, HTML/CSS, C/C++, Java, R, Assembly  
**Software Tools:** Airflow, AWS, Docker, GitHub, Redis, Figma  
**Data Skills:** Data Engineering, Data Visualization, Machine Learning, Computer Vision, NLP  
**Python Libraries:** Pandas, NumPy, Matplotlib, Dask, Ray, psycogp2  
**ML Libraries:** PyTorch (PyTorch Lightning), Tensorflow, Scikit-Learn, Comet  
**CV/NLP Libraries:** Transformers, vLLM, Langchain, Guardrails, nltk, OpenCV

## PUBLICATIONS

---

- Hua SBZ**, Lotfi S, Chen I. Uncertainty Drives Social Bias in Quantized Large Language Models. ICLR Workshop on Principled Design for Trustworthy AI. April 2026.
- Hua SBZ**, He P, Towbin A, Heller N, Chen I, Lu A, Erdman L. Underrepresentation of children in public medical imaging datasets. Nature Health. April 2026.
- Hua SBZ\***, Khondker A\*, . . . , Rickard M, Erdman L. Longitudinal image-based prediction of surgical intervention in infants with hydronephrosis using deep learning: Is a single ultrasound enough? PLOS Digital Health. August 2025.
- Erdman L, Rickard M, Drysdale E, Skreta M, **Hua SBZ**, . . . , Lorenzo AJ, Goldenberg A. The Hydronephrosis Severity Index guides paediatric antenatal hydronephrosis management based on artificial intelligence applied to ultrasound images alone. Scientific Reports. October 2024
- Hua SBZ**, Rickard M, Weaver J, Xiang A, Alvarez D, Velear KN, Sheth K, Tasian GE, Lorenzo AJ, Goldenberg A, Erdman L. From Single-Visit to Multi-Visit Image-Based Models: Single-Visit Models are Enough to Predict Obstructive Hydronephrosis. 18th Symposium on Medical Information Processing and Analysis (SIPAIM). November 2022.
- Hua SBZ**, Lu AX, Moses AM. CytoImageNet: A large-scale pretraining dataset for bioimage transfer learning. NeurIPS Workshop on Learning Meaningful Representations for Life. December 2021.

## POSTER PRESENTATIONS

---

- Uncertainty Drives Social Bias in Quantized Large Language Models**  
ICLR 2026 Workshop on Principled Design for Trustworthy AI (Rio de Janeiro, Brazil)
- Underrepresentation of Children in Public Medical Imaging Datasets**  
Frontiers in Computational Precision Health 2026 (San Francisco, USA)
- Machine Learning-Enabled Renal Ultrasound View Labeling to Expand Use of Point-Of-Care Imaging in Community Settings**  
Nature Conference on Precision Child Health 2024 (Toronto, Canada)
- Supervised Contrastive Learning for Improved View Labeling in Pediatric Renal Ultrasound Videos**  
ISBI 2023 (Cartagena, Colombia)
- CytoImageNet: A large-scale pretraining dataset for bioimage transfer learning**  
NeurIPS 2021 Workshop on Learning Meaningful Representations for Life (Remote)

## ORAL PRESENTATIONS

---

### **Uncertainty Drives Social Bias in Quantized Large Language Models**

ICLR 2026 Workshop on Principled Design for Trustworthy AI (Rio de Janeiro, Brazil)

### **From Single-Visit to Multi-Visit Image-Based Models: Single-Visit Models are Enough to Predict Obstructive Hydronephrosis**

SIPAIM 2022 (Valparaiso, Chile)

## INVITED TALKS

---

### **Towards Meaningful Pretraining Data** (with Alex Lu and Alexander Lin)

Models, Inference & Algorithms Seminar, Broad Institute  
Boston, USA, October 25th, 2023

## HONORS & AWARDS

---

- 2024 **St. Michael's College In-Course Scholarship**, \$3000
- 2024 **Hosinec Family Scholarship**, \$3000
- 2023 **Samuel Beatty Fund Travel Grant**, \$600
- 2021 **University of Toronto CSB Undergraduate Research Award**, \$4000
- 2020-24 **Dean's List Award**
- 2021 **The F. M. Hill Scholarship in Biology**, \$1100

## RESEARCH EXPERIENCE

---

### **Data Scientist**

June 2024 – Aug. 2025

*The Hospital for Sick Children* (Supervisor: Prof. Lauren Erdman)

*Toronto, Canada*

- (With Lauren Erdman & Mandy Rickard) Develop RenalView – a real-time model to guide the acquisition of pediatric renal ultrasound views and improve community point-of-care ultrasound. Validated retrospectively on data from SickKids, Stanford, CHOP and UIowa.
- (With Lauren Erdman) Aggregate and clean public ultrasound data to develop an OOD (or selective classification) benchmark for real-time ultrasound.
- (With Lauren Erdman & Alex Lu) Performed a large-scale review of public medical imaging datasets, demonstrating that less than 1% of public data is pediatric. We further show that young children are at risk if using adult X-ray models.
- (With Lauren Erdman & Michael Chua) Fine-tuned a SAM2-UNet model to segment penises with hypospadias to assist in surgical planning.
- (With Lauren Erdman, Tim Van Mieghem & Catherine Windrim) Developed a segment-reconstruct-segment algorithm for detecting abnormalities – anencephaly – in first trimester fetal ultrasounds.
- Spearheaded the design and approval of 2 REB/IRBs for a prospective silent trial for RenalView, and for a retrospective study on LLMs for referral note prioritization

### **ML Researcher**

May 2024 – Aug. 2025

*University of California, Berkeley*

*Remote*

- (With Irene Chen & Sanae Lotfi) Investigated the impact of quantization on social bias in LLMs, finding that model-bias specific uncertainty is a key contributor of behavioral changes.

### **ML4H Researcher**

Jan. 2024 – June 2024

*Vector Institute* (Supervisor: Prof. Rahul Krishnan)

*Toronto, Canada*

- Benchmarked multi-agent prompting methods to improve zero-shot performance of open-source large multi-modal models (Mixtral 8x7B, LLaVA) for general and healthcare-specific MCQA datasets.
- (With Ian Shi and Philip Fradkin) Assisted in evaluation of mRNA foundation model IsoCLR against RNA foundation models (mRNA-FM and CodonBERT) to perform RNA-binding protein prediction

## Junior ML Specialist

May 2022 – Aug. 2023

*The Hospital for Sick Children* (Supervisors: Prof. Lauren Erdman, Alex Lu, Prof. Irene Chen) *Toronto, Canada*

- Demonstrated that (MoCo) **supervised contrastive image pre-training** can improve **in-domain** and **out-of-distribution generalization** for automatic view labeling on renal **ultrasound videos**.
- Trained and deployed CNN to forecast ER patient volumes, and explored Bayesian methods (GP, Bayesian NNs) for confidence interval estimation.

## ML4H Researcher

Sept. 2021 – May 2022

*The Hospital for Sick Children* (Supervisors: Prof. Lauren Erdman, Prof. Anna Goldenberg) *Toronto, Canada*

- Adapted video-based deep learning methods for **feature aggregation** (Conv.Pooling, CNN-LSTM, TSM) to predict if a child needs kidney surgery from ultrasound images taken over multiple hospital visits.
- Demonstrated that ultrasounds from the first hospital visit alone is enough to predict the need for surgery.

## BioML Researcher

May 2021 – Sept. 2021

*University of Toronto* (Supervisors: Prof. Alan Moses, Alex Lu) *Toronto, Canada*

- Created the **CytoImageNet dataset** (890K images, 894 classes) from 20 TB of open-source microscopy images.
- Showed that CytoImageNet-pretrained models are competitive with ImageNet-pretrained models on downstream microscopy datasets, despite weak labels and minimal hyperparameter tuning (compared to ImageNet).
- The CytoImageNet dataset attracted attention on Kaggle (10696 views, 526 downloads).

## Stats/ML Researcher

Jul. 2020 – Jul. 2021

*University of Toronto* (Supervisor: Prof. Pascal Tyrrell) *Toronto, Canada*

- Showed that choice of dimensionality following dim. reduction (PCA, autoencoder) is important for clustering (K-Means, DBSCAN, Agglomerative) of medical images under small sample sizes using Tensorflow.
- Performed registration and skull-stripping on brain MRIs of children with vasculitis using FSL

## ENGINEERING EXPERIENCE

---

### Data Science Intern

June 2023 – Sept. 2023

*Wealthsimple* *Toronto, Canada*

- Collaborated with client-facing teams to create a **LLM framework** to perform zero-shot **topic/subtopic classification** of **customer service tickets** for ticket routing, resulting in 15% less errors versus humans.
- Designed, implemented and deployed model to an API endpoint that detects if a user entered invalid account details when initiating an account transfer, preventing a projected \$12M in failed transfers annually.

### Software Engineer Intern

May 2022 – May 2023

*Intel Corporation* *Toronto, Canada*

- Refactored a mission-critical **data extraction tool** to extract FPGA benchmarking results from a SQL database or file system, resulting in improved efficiency, code maintainability and unit-test coverage (99%).
- Optimized **SQL** and Pandas code for existing dashboards, speeding up loading time by 400%.
- Developed dashboards and automated checks, to ensure CRON jobs are spaced apart and benchmarking jobs are not abusing high-priority job queues, increasing stability of jobs and saving cloud compute by +1K hours per month

## TEACHING EXPERIENCE

---

### Data Scientist

June 2024 – Aug. 2025

*The Hospital for Sick Children – Centre for Computational Medicine* *Toronto, Canada*

- Organized workshops for the broader SickKids Research Institute on: a) deep learning for medical image analysis, b) accelerating scientific computation with GPUs on the SickKids HPC SLURM cluster and c) a two-part SQL tutorial on the Expedition cloud platform.

### Teaching Assistant

June 2021 – Aug. 2021

*University of Toronto – Independent Summer Statistics Community* *Toronto, Canada*

- Organized data science - related workshops for undergraduate statistics students.
- Presented hands-on workshop introducing the Pandas library.

## ACADEMIC SERVICE

---

### **Program Committee**

Jul. 2026 – May 2026

*Conference on Neural Information Processing Systems (NeurIPS)*

*Remote*

- Served as Assistant Chair of the Positions Paper track.
- Determined desk rejections and assisted in the handling of paper submissions and matching.

### **Round-table Volunteer**

Sept. 2025 – Sept. 2025

*Frontiers in Computational Precision Health Conference*

*Berkeley, USA*

- Led round-table discussions for one group of attendees and coordinated with other round-table facilitators.

### **Volunteer**

Sept. 2025 – Sept. 2025

*Health Futures Unconference*

*Berkeley, USA*

- Assisted in the setup of the conference.

## COMMUNITY SERVICE

---

### **Club President & Founder**

Sept. 2021 – Sept. 2022

*University of Toronto – Offline*

*Toronto, Canada*

- Led a team to build an online community of 200+ undergraduate students on Discord during the pandemic.
- Hosted multiple game nights and movie nights.

### **Events Coordinator**

May 2020 – May 2021

*University of Toronto – Table Tennis Club*

*Toronto, Canada*

- Led new initiatives to increase social media presence and spread information on table tennis players and competitions.

### **General Council Member**

Sept. 2020 – May 2021

*University of Toronto – Bioinformatics & Computational Biology Student Union*

*Toronto, Canada*

- Assisted in planning of online social activities for bioinformatics undergraduate students.

### **Volunteer Chess Teacher**

Aug. 2020 – Aug. 2020

*Kensington Health – Second Mile Club*

*Toronto, Canada*

- Organized, promoted and hosted beginner chess workshops for seniors.

### **Volunteer Photographer**

Feb. 2020 – Feb. 2020

*West Neighborhood House*

*Toronto, Canada*

- Took photographs and videos to help promote the not-for-profit's community services.