Zhou (Joe) Lan

 lanzhouBWH | lanzhoubwh.github.io/ |
 zhou.joe.lan@gmail.com; zlan@bwh.harvard.edu | 404-834-2768

HIGHLIGHTS

I am computational and statistical scientist at Brigham and Women's Hospital / Harvard Medical School with expertise in real-world evidence, machine learning, digital health, causal inference, and translational biomedical data science with **6+ years of working and research experience**. My work spans the **full spectrum** of biomedical/clinical investigation. I focus on leveraging large-scale clinical/healthcare data and machine learning to accelerate clinical development, improve patient outcomes, and inform therapeutic strategies.

- **Large Biobank Digital Health Analytics:** Large-scale imaging/digital-device/omics research using biobanks and population cohorts (e.g., Human Connectome Project, ABCD Study, Boston Scientific's ALTITUDE database and NCDR ICD Registry) for biomarker discovery and disease stratification using through machine learning, data science and digital health approaches;.
- **Epidemiological Real-World Evidence Studies & Outcomes Research:** Analysis of EHRs, insurance claims, registry data (e.g., National Inpatient Sample) for treatment effectiveness, safety evaluation, and patient trajectory modeling using advanced causal inference and survival methods.
- **Clinical Trial Design and Analytics:** Experience supporting clinical studies including Phase IV Design, trial endpoint evaluation using rigorous statistical methodology.

I have participated in a number of studies with my colleagues from **a variety of therapeutic areas** (Radiology, Neurology, Cardiovascular Medicine, Oncology, Endocrinology). I have been **project leaders** on 15+ projects with first-author publications (*Imaging Neuroscience, NeuroImage: Clinical, Bioinformatics, Science Data, JAMIA Open*); I also have been **collaborators** as co-author on 50+ high-impact clinical papers including in *NEJM, Circulation, JAHA, Human Brain Mapping, Neuro-Oncology, Radiology*. I am an developer of statistical and health analytics software; code available at `https://github.com/lanzhouBWH`.

WORK EXPERIENCE

Brigham and Women's Hospital, affiliated with Harvard Medical School   Oct 2022 – present
Computational and Biostatistical Investigator; Faculty Member in Medicine

Center for Outcomes Research & Evaluation, affiliated with Yale School of Medicine   Jul 2020 – Sep 2022
Statistician Consultant

Pennsylvania State University   Aug 2019 – Jul 2020
Bruce Lindsay Fellow

Eli Lilly & Company   May 2018 – Aug 2018
Ph.D. Statistics Intern

Ventana Medical Systems, a part of Roche   May 2016 – Jul 2016
Imaging Scientist Intern

EDUCATION

North Carolina State University, Raleigh, North Carolina

  Ph.D. in Statistics, July 2019

Georgia Institute of Technology, Atlanta, Georgia

  M.S., Statistics (Applied Mathematics), May 2015

Zhejiang University, Hangzhou, China

  B.S., Biology, May 2012

EXPERTISE

Programming	Python, R, C/C++, Bash, Slurm Workload Manager, SAS, SQL
Statistics & Machine Learning	Survival Analysis, Bayesian Modeling, Probabilistic Graphical Models, Quantile Regression, Spatial Statistics, High-Dimensional Data Analysis, Causal Inference
Computation	High-Performance Computing, Parallel Computing, Algorithm Development
Therapeutic Areas	Radiology, Neurology, Cardiovascular Medicine, Oncology, Endocrinology

WORK AUTHORIZATION

United States Permanent Resident (Green Card)

MAIN PROJECTS

Large Biobank Digital Health Analytics:

- **Project 1:** Machine Learning and AI methods for Analyzing Brain White Matter Bundles using Large Biobank Data: Human Connectome Project-Young Adults Data (`https://www.humanconnectome.org/study/hcp-young-adult`) (Lauren J. O'Donnell at Brigham and Women's Hospital, Harvard Medical School)

    - Used scalable **multimodal neuroimaging analytics pipeline** using *UKFTractography* and *WhiteMatterAnalysis* for 1,000+ participants from the Human Connectome Project.
    - Developed **novel statistical framework (Fiber Microstructure Quantile Regression)** for white-matter biomarker discovery, addressing high-dimensional tractography, spatial correlation, and heterogeneity across fiber bundles.
    - Applied and evaluated **deep learning techniques** (neural networks) for cerebellar tract segmentation and neuroanatomical mapping.
    - Validated digital endpoints (FA, MD, NoS) through **numerical studies** for reliability and biomarker robustness.
    - Built reproducible, Docker-based **open-source software pipeline** for tractography analytics and visualization via computer-vision techniques (Github: `https://github.com/lanzhouBWH`).
    - Publications/Reports

        * <u>Zhou Lan</u>, Yuqian Chen, Jarrett Rushmore, Leo Zekelman, Nikos Makris, Yogesh Rathi, Alexandra J. Golby, Fan Zhang, and Lauren J. O'Donnell, "Fiber Microstructure Quantile (FMQ) Regression: A Novel Statistical Approach for Analyzing White Matter Bundles from Periphery to Core", *Imaging Neuroscience* **3**, 11 (2025).(doi: 10.1162/imag_a_00569)
        * Leo R. Zekelman, Suheyla Cetin-Karayumak, Yuqian Chen, Melyssa Almeida, Jon Haitz Legarreta, Jarrett Rushmore, Steve Pieper, <u>Zhou Lan</u>, John E. Desmond, Lissa C. Baird, Nikos Makris, Yogesh Rathi, Fan Zhang, Alexandra J. Golby, Lauren J. O'Donnell, "Consistent cerebellar pathway-cognition associations across pre-adolescents and young adults: a diffusion MRI study of 9000+ participants", *bioRxiv* , (2025). (doi.org/10.1101/2025.02
        * Legarreta, Jon Haitz, <u>Zhou Lan</u>, Yuqian Chen, Fan Zhang, Edward Yeterian, Nikos Makris, Jarrett Rushmore, Yogesh Rathi, and Lauren J. O'Donnell, "Towards an Informed Choice of Diffusion MRI Image Contrasts for Cerebellar Segmentation", *Human Brain Mapping* **46**, 11 (2025).(doi: doi.org/10.1002/hbm.70317)

- **Project 2:** Magnetic Resonance Spectroscopy (MRS)-based Imaging Biomarkers Development functional neurological disorder in children and adolescents using machine learning methods

    - Developed **NMetNet**, a novel Bayesian network–based statistical framework (Bayesian graphical lasso) for identifying **neurometabolic imaging biomarkers** in pediatric FND.
    - Led analysis of high-dimensional MRS data to quantify metabolite interactions and distinguish FND patients from controls with **interpretable network-based machine learning**.
    - Advanced noninvasive biomarker research for pediatric neurological disease, supporting precision diagnostics and early disease characterization.
    - Publications/Reports: <u>Zhou Lan</u>, Sheryl Foster, Molly Charney, Max van Grinsven, Katherine Breedlove, Kasia Kozlowska, Alexander Lin, "Neurometabolic Network (NMetNet) for Functional Neurological Disorder in Children and Adolescents", *NeuroImage: Clinical* **46**, (2025).(doi.org/10.1016/j.nicl.2025.103767)

- **Project 3:** Novel automated pipeline development to assess MR spectroscopy data quality using signal processing/computer vision approaches

    - Mentored an intern to develop an **automatic MRS data processing pipeline** using signal processing approaches, avoiding manual data quality control.
    - Publications/Reports: Beroukhim B, McComas S, Joyce JM, Schuhmacher LS, Koerte I, <u>Lan Z</u>, Lin A, "A novel automated pipeline to assess MR spectroscopy quality control: Comparing current standards and manual assessment.", *J Neuroimaging* **35**, 1 (2025). (DOI: doi.org/10.1080/26941899.2024.2412017)

- **Project 4:** Digital Cardiac Monitoring and Predictive Modeling for ICD Pacing Risk

    - Built study cohort from **multi-institutional cardiac device registry data**.
    - Led statistical modeling to identify **predictors of right ventricular pacing** in patients receiving implantable cardioverter-defibrillators (ICDs) without a pacing indication, using variable selection and predictive modeling approaches.
    - Contributed expertise in **clinical device signal analytics, clinical interpretation, and validation** in collaboration with cardiologists, electrophysiologists, and device-engineering specialists.

- Publication/Reports: Hummel JP, <u>Lan Z</u>, Jones PW, Khera R, Stein K, Curtis JP, Akar J, "Predictors of Right Ventricular Pacing in Patients Undergoing Implantable Defibrillator Placement", *Journal of Cardiovascular Electrophysiology* **36**, 4 (2025).

- **Project 5:** Non-invasive MRI/PET biomarkers for tumor molecular profiling
  Radiogenomics, multimodal imaging, predictive oncology

  - Curated and maintained a large **clinical and imaging database** to evaluate MRI/PET-based radiological biomarkers against biopsy-confirmed tumor molecular characteristics.
  - Applied **predictive modeling and machine learning methods** (logistic regression, ROC analysis, LASSO feature selection) to develop and validate non-invasive biomarkers for tumor aggressiveness and molecular subtype prediction.
  - Supported radiology-oncology collaboration to advance **precision oncology imaging endpoints** and imaging-driven patient triage strategies.
  - Publications/Reports:

    * Allyson L Chesebro, Susan C Lester, <u>Zhou Lan</u>, Catherine S Giess, "Imaging, Patient, and Diagnostic Radiologist Factors Associated With Malignancy for Mammographic Asymmetries Undergoing Biopsy", *Journal of Breast Imaging* **7**, 4 (2025). (DOI: doi.org/10.1093/jbi/wbaf008)
    * Jirarat Jirarayapong, ..., <u>Zhou Lan</u>, ..., Eva C Gombos, "Intraoperative Supine Breast MRI for Residual Tumor Assessment after Breast-Conserving Therapy in Early-Stage Breast Cancer", *Radiology: Imaging Cancer* **7**, 2 (2025). (DOI: doi.org/10.1148/rycan.240158)
    * Jessie L Chai, ..., <u>Zhou Lan</u>, ..., Atul B Shinagare, "Performance of VI-RADS in predicting muscle-invasive bladder cancer after transurethral resection: a single center retrospective analysis", *Abdominal Radiology* **49**, 5 (2025). (DOI: doi.org/10.1007/s00261-024-04245-4)

Epidemiological Real-World Evidence & Outcomes Research

- **Project 1:** Real-world evaluation of statin therapy effectiveness in patients with diabetes
  Electronic health records–based causal inference, cardiometabolic outcomes, survival modeling

  - Built and validated patient cohorts from large **EHR repositories** to study statin initiation and treatment patterns in diabetes care.
  - Applied **advanced causal inference** methods including propensity score weighting, mediation modeling, and competing-risks Cox survival models to evaluate cardiovascular outcomes.
  - Led statistical methodology and results interpretation; contributed methods and results sections for peer-reviewed publications.

    * Shah, Nisarg, Zhou Lan, C. Justin Brown, Seth S. Martin, and Alexander Turchin, "Impact of Statin Nonacceptance on Cardiovascular Outcomes in Patients With Diabetes", *Journal of the American Heart Association* **14**, 11 (2025).(doi: https://doi.org/10.1161/JAHA.124.040464)
    * C. Justin Brown, PharmD; Lee-Shing Chang, MD; Naoshi Hosomura, DDS, DMSc, MBA; Shervin Malmasi, PhD; Fritha Morrison, PhD; Maria Shubina, ScD; <u>Zhou Lan, PhD</u>; Alexander Turchin, MD, MS , "Assessment of Sex Disparities in Nonacceptance of Statin Therapy and Low-Density Lipoprotein Cholesterol Levels Among Patients at High Cardiovascular Risk", *JAMA Network Open* **6**, 2 (2023). (doi:10.1001/jamanetworkopen.2023.1047)

- **Project 2:** Evaluating reliability of NLP-derived clinical variables in real-world epidemiologic analyses
  Methodological validation, Monte Carlo simulation, clinical NLP quality assessment

  - Proposed and led a methodological study on how **NLP extraction errors from clinical notes** affect causal and epidemiologic inference in real-world data research.
  - Designed and executed **Monte Carlo simulation experiments** using high-performance computing to quantify downstream bias under different NLP error scenarios. *Code:* `https://github.com/lanzhouBWH/NLP-Downstream-Impact`
  - Demonstrated the importance of **data validation and error modeling** in clinical AI pipelines to support robust RWE and regulatory-grade analytics.
  - Publications/Reports:

    * <u>Zhou Lan</u>, Alexander Turchin, "Impact of Possible Errors in NLP-Derived Data on Downstream Epidemiologic Analysis", *JAMIA Open* **6**, 4 (2023). (DOI: 10.1093/jamiaopen/ooad111)

- **Project 3:** Prognostic modeling for patients with primary brain tumor
  Multimodal real-world oncology data, survival modeling, radiogenomics

  - Constructed real-world patient cohorts from clinical records to evaluate **clinical, molecular, and radiological predictors** of outcomes in astrocytoma and IDH-mutant glioma.

- Applied **predictive modeling and survival analysis** (including weighted Cox models for non-proportional hazards) to assess associations with overall survival (OS) and progression-free survival (PFS).
- Integrated **radiology, genomics, and clinical features** to support precision oncology and treatment response profiling.
- Performed **multiple imputation** to address missing data and improve robustness of real-world evidence analyses.
- Publications/Reports:
  * Aleksandra B Lasica, <u>Zhou Lan</u>, ..., Raymond Y Huang, Patrick Y Wen, Gilbert Youssef, "Clinical, molecular and radiological predictors of prognosis in newly diagnosed astrocytoma, IDH-mutant, WHO grade 4", *Neuro-Oncology* **7**, 1 (2025).(doi: https://doi.org/10.1093/neuonc/noaf133)
  * Tyler A Lanman, ..., <u>Zhou Lan</u>, ..., Patrick Y Wen, L Nicolas Gonzalez Castro, "Ivosidenib for the Treatment of IDH1-mutant Glioma, Grades 2 to 4: Tolerability, Predictors of Response, and Outcomes", *Neuro-Oncology Advances* **7**, 1 (2025).(doi: https://doi.org/10.1093/noajnl/vdae227)

- **Project 4:** National Inpatient Sample (NIS) — Registry-scale real-world data analytics
  Large-scale clinical registry data, ICD coding, orthopedic outcomes, reproducible pipelines

  - Designed and implemented a **Python toolbox** for NIS data processing and analysis, including **ICD-9/ICD-10-based cohort and outcome definitions**. *Code: $https://github.com/lanzhouBWH/Toolbox-for-National-Inpatien$*
  - Conducted **epidemiologic and outcomes research** to evaluate venous thromboembolism (VTE) prophylaxis strategies in hip fracture, total knee arthroplasty, and total hip arthroplasty populations.
  - Delivered end-to-end real-world analysis workflow including cohort creation, risk adjustment, and interpretation of comparative effectiveness results.
  - Publications/Reports:
    * Leyva, H., Fanikos, J., <u>Lan, Z.</u>, Scimeca, G., Bikdeli, B., Piazza, G. and Goldhaber, S.Z, "Aspirin for Venous Thromboembolism Prophylaxis in Hip Fracture, Total Knee, and Hip Arthroplasty", *The American journal of medicine* **138**, 9 (2025).(doi: https://doi.org/10.1016/j.amjmed.2025.03.020)

- **Project 5:** CORONA-VTE Multicenter Registry — Cardiovascular and thrombotic outcomes in COVID-19
  Multi-institution real-world evidence, longitudinal cardiovascular outcomes, registry harmonization

  - Built a harmonized **multi-center clinical database** integrating EHR data from five major systems (Ann Arundel, Beth Israel, Colorado, Mass General Brigham, Thomas Jefferson, University of Virginia) for COVID-19 cardiovascular outcomes research.
  - Performed **cluster Cox regression** to address site-level heterogeneity and evaluate risk factors for venous thromboembolism and cardiovascular events.
  - Conducted **longitudinal analyses** to assess vaccination effects on cardiovascular outcomes among inpatient and outpatient COVID-19 populations.
  - Publications/Reports:
    * Giovanni Scimeca, ...,<u>Zhou Lan</u>,..., Behnood Bikdeli, Gregory Piazza, "Predictors of venous thromboembolic events in hospitalized patients with COVID-19", *Journal of Thrombosis and Thrombolysis* **58**, 4 (2025).(doi: https://doi.org/10.1007/s11239-025-03078-2)
    * B Bikdeli, D Krishnathasan, CD Khairani, ..., <u>Z Lan</u>, S Goldhaber, G Piazza, "Adjudicated cardiovascular events in a cotemporary multicenter cohort of patients with COVID-19: findings from the CORONA-VTE Registry", *European Heart Journal* **138**, 9 (2025).(doi: https://doi.org/10.1093/eurheartj/ehad655.1223)
    * Behnood Bikdeli, ..., <u>Zhou Lan</u>, Gregory Piazza, "Low absolute risk of thrombotic and cardiovascular events in outpatient pregnant women with COVID-19", *European Heart Journal* **138**, 9 (2025).(doi: https://doi.org/10.1016/j.thr

- **Project 6:** CMS Excess Days in Acute Care (EDAC) Quality Measure
  Medicare claims analytics, healthcare quality measurement, regulatory-grade modeling

  - Served as the **Lead Measure Analyst** for the U.S. CMS EDAC program, quantifying post-discharge acute care utilization (readmissions, observation stays, ED visits) to evaluate national hospital performance.
  - Maintained and enhanced **production SAS codebases** for data preprocessing, risk adjustment, model validation, and ICD-10 clinical code updates, ensuring accuracy and regulatory compliance.
  - Prepared technical documentation and contributed to public reporting materials published on CMS QualityNet, supporting transparency and federal reporting requirements. *Link: $https://qualitynet.cms.gov/files/64501bdf9920e90$ $filename=2022\_EDAC\_MeasMethRpt.zip$*

Clinical Trials Design

- **Project 1:** Apixaban for Extended Treatment of Provoked Venous Thromboembolism (HI-PRO Trial) (Brigham and Women's Hospital Thrombosis Research Group `https://thrombosisresearch.bwh.harvard.edu`)
  - Performed **study design**, **primary and secondary endpoints analyses**, generating **regulatory-grade evidence** consistent with clinical trial standards.
  - Contributed to **clinical trial manuscripts and regulatory documentation**, including statistical interpretation and reporting in alignment with ICH E9 guidelines.
  - Publications/Reports:
    * Gregory Piazza, Behnood Bikdeli, Arvind K. Pandey, Darsiya Krishnathasan, Candrika D. Khairani, Antoine Bejjani, Ruth H. Morrison, Heather Hogan, Sina Rashedi, Mariana Pfeferman, Junyang Lou, John Fanikos, Nicole Porio, Lisa Rosenbaum, Piotr Sobieszczyk, Zhou Lan, Marie Gerhard-Herman, Umberto Campia, Samuel Z. Goldhaber, for the HI-PRO Trial Investigators, "Apixaban for Extended Treatment of Provoked Venous Thromboembolism (HI-PRO Trial)", *New England Journal of Medicine* **393**, (2025). (DOI: 10.1056/NEJMoa2509426)

PHARMACEUTICAL, BIOTECH, AND HEALTHCARE ANALYTICS PARTNERSHIP

- Eli Lilly And Company (Complications of Obesity),
- Pfizer Inc (Flutemetamol-ATTR),
- Janssen Research and Development (Covid-19 PREVENT-HD)
- Boston Scientific Corporation (OPTALYSE-PE Trial)

PUBLICATIONS

Complete List of Published Work and Preprints in google scholar https://scholar.google.com/citations?user=fB87fIgAAAAJ

PATENTS

Chukka Srinivas, Zhou Lan, "System and Method for Color Deconvolution of a Slide Image to Assist in the Analysis of Tissue", , US20200167965A1 (Priority Date: 2017/08/04; Publication Date: 5/28/2020).