Hyejung Maybell Lee

Salt Lake City, UT hyejung.lee@utah.edu
Website

EDUCATION

Doctor of Philosophy, Population Health Sciences emphasis in Biostatistics **GPA: 3.99** **December, 2025**
University of Utah, Salt Lake City, UT, USA

- Courses: Bayesian Inference, Machine Learning, Causal Inference(theory and applied), Advanced Statistical Inference, Epidemiologic Methods, Clinical Trials, Multilevel Modeling, Survival Analysis, Research Ethics
- Improving patient care and disease management in chronic diseases (chair: Jincheng She, Ph.D., co-chair: Tom Greene, Ph.D.)

Master of Science, Statistics **GPA: 3.82** **August, 2020**
Brigham Young University, Provo, UT, USA

- Courses: Generalized Linear Models, Applications in biostatistics, Bayesian Methods, Multivariate Statistical Methods, Statistical Computation, Probability Theory and Mathematical Statistics 1, Linear Models, Statistical Learning & Data Mining, Mixed Model Methods, Probability Theory and Mathematical Statistics 2

Bachelor of Science, Pure Mathematics, Minor: Applied Mathematics **GPA: 3.69** **November, 2014**
University of Calgary, Calgary, AB, Canada

- Graduate with Distinction on a competitive GPA basis
- Courses: Survival Analysis, Introduction to Biostatistics, Applied Regression Analysis, Introduction to Stochastic Process, Introduction to Biochemistry, Calculus, Linear Algebra, Analysis, Principles of Genetics, Organic Chemistry

SKILLS

- Advanced R Programming
- Parallel computing
- High performance computing
- R packaging – example: LMER bootstrap
- Multivariate statistical modeling
- Data interpretation to clinical relevance
- Causal inference
- Bayesian Statistics
- Monte Carlo and Bootstrap simulation in R
- Git/GitHub

WORK EXPERIENCE

Research Associate **2025-Current**
Department of Internal Medicine, Division of Epidemiology, University of Utah

- Develope and validate novel statistical methods of surrogate endpoints to replace expensive clinical endpoint as a part of the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI)
- Develope custom function in R using Bayesian modeling, along with STAN script
- Create web-based dissemination tools

Research Assistant **2021-2025**
University of Utah, Huntsman Cancer Institute – Biostatistics Shared Resources in SLC, UT. Lead statistician for 3 collaboration projects.

- Identify patient characteristics and clinical histology that can determine whether the advanced non-small cell lung cancer patient should be waiting for biomarker test results prior to starting treatment
    - Data: a nation-wide longitudinal electronic health record of advanced non-small cell lung cancer from Flatiron Health.

Hyejung Maybell Lee

Salt Lake City, UT hyejung.lee@utah.edu
Website

  - Method: Causal inference using inverse probability weighting on accelerated failure time model
- Identify association between longitudinal body composition and survival in metastatic non-small cell lung cancer patients
  - Data: Observational study of advanced non-small-cell lung cancer from Huntsman Cancer Institue
  - Method: Joint modeling of linear mixed effects model and Cox proportional hazard model
- Identify circulatory proteomes associated with prostate cancer progression
  - Data: Protein assays of protstate cancer patients at Huntsman Cancer Institue
  - Method: linear model, false discovery rate control, Cox proportional hazards model
- Investigate associations between overall survival, hospitalization-free survival, and patient reported outcomes of late-stage lung cancer patients and fat and muscles

Biostatistics Intern **Summer 2024**

Regeneron Pharmaceuticals Inc.

- Model natural disease progression in Parkinson's Disease using pseudotime
- Novel misclassification error evaluation
- Unsupervised clustering

Research Assistant **2021-2022**

University of Utah, Study Design & Biostatistics Center in SLC, UT. Lead statistician for 2 collaboration projects and assisted 1 collaboration project.

- Assess the effect of MRI-guided ablation on recurrence of atrial arrythmia in patients with cystic fibrosis
  - Data: Clinical trials data analysis from a international multi-center randomized clinical trial
  - Method: Apply multivariable Cox proportional hazards regression models
- Predict adverse events and venous thromboembolism among the patients with T cell acute leukemia using logistic regression
- Investigate inter and intra-rater reliability of fetal heart measures to establish standardized Z-score for scales

Research Assistant **2020 –2021**

Rocky Mountain Center for Occupational and Environmental Health in SLC, UT. Lead statistician for a state-funded project.

- Analyze association between major depressive disorder and work-industry and individual behavioral factors

Research Assistant **2018-2020**

Dr.Tolley, Department of Statistics, Brigham Young University, Provo, UT. Master's thesis.

- Develop statistical methods of deconvolving target compounds that have overlapping elution patterns on mass spectrometry by using the known ratios of absorbances at the two light way lengths
- Use of skewed normal distribution for functional data analysis – modeling chemical compounds from mass spectrometry using skewed normal distribution
- Develop user-interactive function in R

Research Assistant **2018**

Department of Statistics, Brigham Young University, Provo, UT.

- Grade homework and provide feedback for improvement
- Explain concepts in multiple different ways using basic vocabulary

PUBLICATIONS

1. **Lee, H. […],** Greene, T., (2025) Optimization of treatment effect on Clinical Endpoint in Chronic Kidney Disease Phase II/III trials by Incorporating Surrogate eGFR slopes and Clinical Endpoints using 2-stage Bayesian Mixed Effect Meta-Regression Model . (In-progress)

2. **Lee, H. […],** Greene, T., (2025) Optimization of treatment effect on Clinical Endpoint in Chronic Kidney Disease Phase II/III trials by Incorporating Surrogate eGFR slopes, Albuminaria, and Clinical Endpoints using 2-stage Bayesian Mixed Effect Meta-Regression Model. (In-progress)

Hyejung Maybell Lee

Salt Lake City, UT hyejung.lee@utah.edu
Website

3. **Lee, H.,** Akerley, W., Alhamad, K., Haaland, B., Shen, J., (2025) Immpact of Biomarker Testing Wait Time on Treatment Decision Optimization in Advanced Non-Small Cell Lung Cancer. (In-progress)

4. **Lee, H.**, Shen, J., Fadlullah, M.Z.H., Neibling, A., Hanson , C., Lin , T., Larsen , M., Lloyd , J., Maughan , B.L., Swami , U., Agarwal , N., Gupta , S., Tward , J.D., Johnson , S.B., ONeil , B., Dechet , C.B., Haaland , B., Wang , L., Tan , A.-C. and Kohli , M. (2025). Circulatory Prostate Cancer Proteome Landscapes and Prognostic Biomarkers in Metastatic Castrate Resistant Prostate Cancer. *Clinical Proteomics,* 22(1) 13. https://doi.org/10.1186/s12014-025-09536-6 (Published)

5. Wang, X., **Lee, H.,** Haaland, B., Kerrigan, K., Puri, S., Akerley, W. and Shen, J., 2024. A matching-based machine learning approach to estimating optimal dynamic treatment regimes with time-to-event outcomes. *Statistical Methods in Medical Research,* 33(5), pp.794-806 (Published)

6. Coletta, A.M., **Lee, H.**, Puri, S., Culleton, S., Covington, M.F., Yap, J.T., Maslana, K.E., Haaland, B. and Akerley, W. (2025), The Association Between Body Composition, Overall Survival, Treatment Decisions, and Patient-Reported Outcomes in Metastatic Non-Small-Cell Lung Cancer. *Cancer Med*, 14: e70534. https://doi.org/10.1002/cam4.70534 (Published)

7. Akoum, N., Mekhael, M., Bisbal, F., Wazni, O., McGann, C., **Lee, H**., Bardsley, T., Greene, T., Dean, J.M., Dagher, L. and Kholmovski, E., 2024. Lesion Delivery and Scar Formation in Catheter Ablation for Atrial Fibrillation The DECAAF II Trial. *Heart Rhythm,* 2024-09 (Published)

8. Arnold, L.M., Hoshina, Y., **Lee, H.**, Colman, H. and Mendez, J., 2024. Effect of Pneumocystis jirovecii pneumonia prophylaxis on hematologic toxicity in patients receiving chemoradiation for primary brain tumors. *Journal of Neuro-Oncology*, pp.1-7. (Published)

9. Peterson, L., **Lee, H.**, Huybrechts, I., Biessy, C., Neuhouser, M.L., Haaland, B., Krick, B., Gunter, M., Schulze, M.B., Jannasch, F. and Coletta, A.M., 2023. Reliability estimates for assessing meal timing derived from longitudinal repeated 24-hour dietary recalls. *The American Journal of Clinical Nutrition*, *117*(5), pp.964-975. (Published)

10. Moon-Grady, A.J., **Lee, H.**, Lopez, L., Fatusin, O., Freud, L.R., Hogan, W., Krishnan, A., McFarland, C., Minich, L.L., Morris, S.A. and Pinto, N., 2023. Fetal Echocardiographic Z-Score Pilot Project: Study Design and Impact of Gestational Age and Variable Type on Reproducibility of Measurements Within and Across Investigators. *Journal of the American Society of Echocardiography*. (Published)

PRESENTATIONS

Joint Statistical Meeting (JSM) Presentation **August 4, 2025**

- Modeling disease progression in Parkinson's Disease with cross-section data using pseudo-time

Regeneron High Performance Computation Talk **August 2, 2024**

- Introduction on how to use high performance computing within Regeneron's internal system, focusing on basic language that are useful to do parallel with controlled random seed

Olink Technology Seminar **May 1, 2024**

- PowerPoint presentation: Identifying plasma-based Proteins related to Prostate Cancer Progression and Death

ASCO Quality Care Symposium **Sept. 30 – Oct 1, 2022**

- Title: The association between body composition, quality of life (QoL), overall survival (OS) and decision to treat (DTT) in patients with metastatic non–small cell lung cancer (mNSCLC)
- Poster presentation

Hyejung Maybell Lee

Salt Lake City, UT hyejung.lee@utah.edu
Website

Journal of the American Society of Echocardiography **July 2022**

- 2022 Arthur E Weyman Young Investigator's Award Competition Finalists
- Abstract Title: Fetal Echocardiographic Z-Score Pilot Project: Study Design and Impact of Gestational Age and Variable Type on Reproducibility of Measurements within and Across Investigators
- Inter and intra-rater reliability of fetal heart measures

19th Annual Regional National Occupational Research Agenda (NORA) Young/New Investigators Virtual Symposium Summer Research **April 2021**

- Association Between Major Depressive Disorder and Personal Factors Among Utah Lawyers – peer reviewed poster presentation

SCHOLARSHIPS/AWARDS

- 25th Summer Institute in Statistical Genetics (SISG) by University of Washington, School of Public Health, Department of Biostatistics (2020)
- Dean's list (2011,2013,2014)
- Jason Lang Scholarship (2011, 2013)
- Education Matters Scholarship (2010)
- Alexander Rutherford Scholarship (2010)

TEACHING EXPERIENCE

Teaching Assistant **2018**

Methods and Mechanics of Secondary Data Analysis, Department of Population Health Sciences, University of Utah, SLC, UT.

- Lead applied exercises in class
- Grade homework and provide feedback for improvement

Teaching Assistant **2018**

Probability and Inference I, Department of Statistics, Brigham Young University, Provo, UT.

- Grade homework and provide feedback for improvement
- Explain concepts in multiple different ways using lay language

Private Mathematics Tutor **2010-2015**

Self-employed, AB, Canada

- Teach private lessons in beginning math to 10-15 years old
- Explain mathematical concepts using objects and writing

LEADERSHIP EXPERIENCE

Judge at Brigham Young University Student Research Conference **March 8, 2025**
High School Mentoring **2024**
Student Advisory Committee for Faculty Review and Advancement **2023**
Korean Young Single Adult(KYSA) Conference Co-Chair **2019**

VOLUNTEER EXPERIENCE

Undergraduate Research Summer Symposium at the University of Utah **2024**
Community Math Tutoring **2023**
Gradudate School Diversity Preview Day **2022-2023**
Student Advisory Committee for Faculty Review and Advancement **2023**

Hyejung Maybell Lee

Salt Lake City, UT hyejung.lee@utah.edu
Website

Utah Food Bank	**2022-2023**
Graduate School Diversity Office Representative at the Utah Conference on Undergraudate Research	**2023**
Population Health Sciences Virtual Open House	**2022**