

Jonghyun Lee

PostDoc · Department Biostatistics, Epidemiology, and Informatics

University of Pennsylvania, Philadelphia, PA 19104

✉ Jonghyun.Lee@penmedicine.upenn.edu | <https://www.linkedin.com/in/jonghyunlee1993/> |

<https://scholar.google.com/citations?user=B892PiYAAAAJ&hl=en>

Research Interests

My research develops **multimodal AI systems that integrate biological data across scales**, from single-cell cytometry to tissue histology to population-level imaging. By combining representation learning, foundation model architectures, and large-scale clinical data, I build frameworks that are scalable, interpretable, and clinically translatable. My work spans computational pathology, cytometric data modeling, and data-driven epidemiology, unified by the goal of connecting diverse biomedical modalities to enable precision medicine.

Education

Hanyang University

Seoul, Korea

PhD Medical and Digital Engineering

2021 - 2024

- Advisor: Dae Won Jun, MD-PhD
- Focus: Statistical medical data analysis

Yonsei University

Seoul, Korea

MS Cognitive Science

2017 - 2019

- Advisor: Sanghoon Han PhD
- Focus: Episodic memory & brain stimulation

Yonsei University

Seoul, Korea

BS Korean Language and Literature & Psychology

2012 - 2017

- Graduated 1st in class (1/66)

Professional Experience

Apr 2024 - **Post Doctoral Researcher**, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
Present Advisor: Dokyoon Kim PhD - Multi-modal medical AI

Mar 2023 - **Research Instructor**, Department of Pathology, Korea University Anam Medical Center, Seoul, Korea
Mar 2024 Advisor: Sangjeong Ahn MD-PhD - Computational pathology

Jan 2020 - **Researcher**, Data Analysis Team, Fittogether, Seoul, Korea
May 2021 Sports AI modeling, especially GPS data

Publications

Selected Papers

- Jonghyun Lee**, Lina Takemaru, Dewan M Bappy, Ye Sul Jeong, Won-Ki Jeong, Derek Oldridge, Dokyoon Kim, Sangjeong Ahn, Seong Hak Lee. 2025. Adaptive Compression Framework for Giga-pixel Whole Slide Images. *Nature Communications*.
- Jonghyun Lee**, Jakob Leiby, ..., Dokyoon Kim. 2026. Clinically Aligned Multi-Modal Image-Text Model for Pan-Cancer Prognosis Prediction. *BioData Mining*.
- Jonghyun Lee**, Sangjeong Ahn, Hyun-Soo Kim, Jungsuk An, Jongmin Sim. 2024. Robust model training strategy via hard negative mining in the weakly labeled dataset for lymphatic invasion in gastric cancer. *Journal of Pathology: Clinical Research*.

Published Papers

- Eileen Laurel Yoon, Hyo Young Lee, **Jonghyun Lee**, Joo Hyun Oh, Inseok Hwang, Chan Hee Park, Sejung Kim, Huiyul Park, Ae Jeong Jo, Dae Won Jun. 2026. Cryptogenic Steatotic Liver Disease: A Lean Phenotype Associated with Increased Liver-Related Mortality. *Gut*.
- Seri Jeong, Keun-Young Lee, Ga-Hyun Son, Kyong-No Lee, Won Kyong Cho, Hoseong Choi, **Jonghyun Lee**, Joanne Kwak-Kim. 2026. Dynamics of Immune-Checkpoint Regulators and Inflammatory Cytokines and Chemokines Associated with Preterm Birth Before and After Cervical Cerclage. *Frontiers in Immunology*.
- Joo Hyun Oh*, **Jonghyun Lee***, Miyoung Choi, Kento Imajo, Masato Yoneda, Hideki Fujii, Hyo Young Lee, Eileen L Yoon, Mimi Kim, Dae Won Jun. 2026. Rule-out Cut-off of 2D-Shear Wave Elastography for Significant Fibrosis in MASLD. *Clinical Gastroenterology and Hepatology*.
- Guk Jin Lee, Sang-Hyuk Jung, **Jonghyun Lee**, ..., Dokyoon Kim. 2026. Polygenic Risk Score Predicts Prostate Cancer Risk Independent of Type 2 Diabetes. *Diabetes, Obesity and Metabolism*.
- Hyeseong Lee, Yoo Jin Lee, Eunsu Kim, **Jonghyun Lee**, Sangjeong Ahn, Sung Hak Lee. 2026. Subtype classification of gastric spindle cell tumors in whole slide images. *Computers in Biology and Medicine*.
- Eunsu Kim, Yoo Jin Lee, **Jonghyun Lee**, Hyeseong Lee, DM Bappy, Sangjeong Ahn, Sung Hak Lee. 2025. Fusing Domains for Tumor Detection in Frozen Section Histopathology Using Joint Feature Alignment and Self-Knowledge Distillation. *IEEE Access*.
- Young Eun Chon, **Jonghyun Lee**, Eileen L Yoon, Soon Sun Kim, Sang Bong Ahn, Soung Won Jeong, Dae Won Jun. 2025. Changes in antiviral treatment rate for hepatitis B virus before hepatocellular carcinoma diagnosis: a nationwide Korean study. *European Journal of Gastroenterology & Hepatology*.
- Soon Sun Kim, **Jonghyun Lee**, Sang Bong Ahn, Young Eun Chon, Eileen Yoon, Soung Won Jeong, Dae Won Jun. 2025. Clinical Course and Prognosis of Long Term Survivors of Hepatocellular Carcinoma. *Alimentary Pharmacology & Therapeutics*.
- Jonghyun Lee**, Dokyoon Kim, Dae Won Jun, Yun Kim. 2024. Multi-Modal Fusion-based Lightweight Model for Enhanced Generalization in Drug-Target Interaction Prediction. *Journal of Chemical Information and Modeling*.
- Jonghyun Lee**, Dae Won Jun, Ildae Song, Yun Kim. 2024. DLM-DTI: A Dual Language Model for the Prediction of Drug-Target Interaction with Hint-Based Learning. *Journal of Cheminformatics*.
- Jonghyun Lee**, Seunghyun Cha, ..., Myung-Giun Noh. 2024. Ensemble Deep Learning Model to Predict Lymphovascular Invasion in Gastric Cancer. *Cancers*.
- Jonghyun Lee**, In-Soo Myeong, Yun Kim. 2023. The drug-like molecule pre-training strategy for drug discovery. *IEEE Access*.
- Ju Won Kim*, **Jonghyun Lee***, Sung Hak Lee, Sangjeong Ahn, Kyong Hwa Park. 2024. Machine Learning-Based Prognostic Gene Signature for Early Triple Negative Breast Cancer. *Cancer Research and Treatment*.
- Wooyoung Jang, **Jonghyun Lee**, Kyong Hwa Park, Aeree Kim, Sung Hak Lee, Sangjeong Ahn. 2024. Molecular Classification of Breast Cancer Using Weakly Supervised Learning. *Cancer Research and Treatment*.

- Jeong Hoon Lee, Ga Young Song, **Jonghyun Lee**, ..., Deok Hwan Yang. 2024. Prediction of immunochemo therapy response for diffuse large B cell lymphoma using artificial intelligence digital pathology. *The Journal of Pathology: Clinical Research*.
- Joo Hyun Oh, **Jonghyun Lee**, Eileen L Yoon, Soung Won Jeong, Soon Sun Kim, Young Eun Chon, Sang Bong Ahn, Dae Won Jun. 2023. Regular Alpha-Fetoprotein Tests Boost Curative Treatment and Survival for Hepatocellular Carcinoma Patients in an Endemic Area. *Cancers*.
- Huiyul Park, Eileen L Yoon, Mimi Kim, **Jonghyun Lee**, Hye-Lin Kim, Seon Cho, Eun-Hee Nah, Dae Won Jun. 2023. Diagnostic performance of the fibrosis-4 index and the NAFLD fibrosis score for screening at-risk individuals in a health check-up setting. *Hepatology Communications*.
- Huiyul Park, Eileen L Yoon, Takanori Ito, Ae Jung Jo, Mimi Kim, **Jonghyun Lee**, ..., Mindie H Nguyen. 2023. Diagnostic Performance of the Fibrosis-4 Index and Nonalcoholic Fatty Liver Disease Fibrosis Score in Lean Adults With Nonalcoholic Fatty Liver Disease. *JAMA Network Open*.
- Mimi Kim, Eileen Laurel Yoon, **Jonghyun Lee**, Seon Cho, Chul-Min Lee, Bo Kyeong Kang, Huiyul Park, Dae Won Jun, Eun-Hee Nah. 2023. Diagnostic Performance of Noninvasive Tests for Advanced Hepatic Fibrosis in Young Age Population. *Clinical Gastroenterology and Hepatology*.
- Young Eun Chon, Aejeong Jo, Eileen L Yoon, **Jonghyun Lee**, Ho Gyun Shin, Min Jung Ko, Dae Won Jun. 2023. The Incidence and Care Cascade of the Hepatitis C Virus in Korea. *Gut and Liver*.
- Young Eun Chon, Seong Yong Park, Han Pyo Hong, Donghee Son, **Jonghyun Lee**, Eileen Yoon, Soon Sun Kim, Sang Bong Ahn, Soung Won Jeong, Dae Won Jun. 2023. Hepatocellular carcinoma incidence is decreasing in Korea but increasing in the very elderly. *Clinical and Molecular Hepatology*.
- Yejin Jeong, Cristina Eunbee Cho, Ji-Eon Kim, **Jonghyun Lee**, Namkug Kim, ... , Kyoung Min Moon, Sangjeong Ahn. 2022. Deep learning model to predict Epstein-Barr virus-associated gastric cancer in histology. *Scientific Reports*.
- Jeongbin Kim, Hyunsung Kim, **Jonghyun Lee**, Jaechan Lee, Jinsung Yoon, Sang-Ki Ko. 2022. A Deep Learning Approach for Fatigue Prediction in Sports Using GPS Data and Rate of Perceived Exertion. *IEEE Access*.
- Chul-min Lee, Eileen L Yoon, Atsushi Nakajima, Masato Yoneda, Hidenori Toyoda, Satoshi Yasuda, **Jonghyun Lee**, ..., Yoshio Sumida. 2022. A Reappraisal of the Diagnostic Performance of B-Mode Ultrasonography for Mild Liver Steatosis. *American Journal of Gastroenterology*.
- Kyung A Kang, **Jonghyun Lee**, Hyun-Soo Kim, Sung Yoon Park. 2022. Ovarian masses in patients with breast cancer. *Journal of Computer Assisted Tomography*.
- Huiyul Park, Eileen L Yoon, Mimi Kim, **Jonghyun Lee**, Seon Cho, Dae Won Jun, Eun-Hee Nah. 2022. Reappraisal of Fibrosis-4 Index and Non-alcoholic Fatty Liver Disease Fibrosis Score for Advanced Fibrosis in Average-Risk Population. *Frontiers in Medicine*.
- Huiyul Park, Eileen L. Yoon, Mimi Kim, **Jonghyun Lee**, Jung-Hwan Kim, Seon Cho, Dae Won Jun, Eun-Hee Nah. 2021. Comparison of diagnostic performance between FIB-4 and NFS in metabolic-associated fatty liver disease era. *Hepatology Research*.
- Su Hyun Park*, **Jonghyun Lee***, Dae Won Jun, Kyung A Kang, Ji Na Kim, Hee Jin Park, Han Pyo Hong. 2021. Determining the Target Population That Would Most Benefit from Screening for Hepatic Fibrosis in a Primary Care Setting. *Diagnostics*.
- Yoonjin Nah, **Jonghyun Lee**, Sanghoon Han. 2018. Interactivity within large-scale brain network recruited for retrieval of temporally organized events. *Korean Journal of Cognitive Science*.
- Jeesung Ahn, **Jonghyun Lee**, Jin Hyuk Han, Min Seong Kang, Sanghoon Han. 2018. Group analysis data representing the effects of frontopolar transcranial direct current stimulation on the default mode network. *Data in Brief*.
- Jonghyun Lee**, Sanghoon Han. 2019. The effect of transcranial direct current stimulation on hippocampus for episodic memory modulation: simultaneous tDCS-fMRI approach. *Master's Thesis*.

Conferences

Akira A Nair, Jaehyun Joo, **Jonghyun Lee**, Lina Takemaru, Yidi Huang, Manu Shivakumar, Matthew E Lee, Jaesik Kim, Sokratis Apostolidis, Dokyoon Kim. 2026. Interpreting Genomic Language Models using Sparse Autoencoders. *Mechanistic Interpretability Workshop, ICML*.

Jonghyun Lee*, Sang-Hyuk Jung*, ..., and Dokyoon Kim. 2026. Large-Scale Image-Derived Environmental Signatures for Personalized Risk Prediction. *DBEI Research Day, UPenn*.

Jonghyun Lee, Jakob Leiby, ..., Dokyoon Kim. 2025. Clinically Aligned Multi-Modal Image-Text Model for Pan-Cancer Prognosis Prediction. *MidAtlantic Bioinformatics Conference*.

Matthew E. Lee, Jaesik Kim, Matei Ionita, **Jonghyun Lee**, ..., Allison R Greenplate. 2025. MAESTRO: Masked Encoding Set Transformer with Self-Distillation. *ICLR*.

Hyesung Lee, **Jonghyun Lee**, Eunsoo Kim, Sangjeong Ahn, Sunghak Lee. 2023. Revisiting Active Learning in Histopathology Segmentation: Representativeness and Cost-Effective Active Learning. *MICCAI Annual Meeting, MedAGI Workshop*.

Cristina Eunbee Cho, Yeojin Jeong, Ji-Eon Kim, **Jonghyun Lee**, ..., Sangjeong Ahn. 2022. Deep Learning Model for the Prediction of EBV-associated Gastric Cancer. *USCAP Annual Meeting*.

Jonghyun Lee, Dae Won Jun, Yun Kim. 2021. Predicting drug molecule structure using a pre-trained BERT-based language model. *Korean Society of Artificial Intelligence in Medicine*.

Hyunsung Kim, Jihun Kim, Dongwook Chung, **Jonghyun Lee**, Jinsung Yoon, Sang-Ki Ko. 2021. 6MapNet: Representing Soccer Players from Tracking Data by a Triplet Network. *8th Workshop on Machine Learning and Data Mining for Sports Analytics*.

Seokyoung Min, Je Hoon Chae, **Jonghyun Lee**, Sang Yup Lee. 2021. Identifying individuals using topic patterns of Instagram photos. *7th International Conference on Computational Social Science*.

Jonghyun Lee, Hojoon Lee, Jeesung Ahn, Seung-Koo Lee, Sanghoon Han. 2018. Exploring the high-resolution EPI fMRI protocol to reduce susceptibility-related BOLD signal dropout. *Organization for Human Brain Mapping*.

Seokyoung Min, Suhnyoung Jun, Jeesung Ahn, **Jonghyun Lee**, Soohyun Park, Seung Koo Lee, Sanghoon Han. 2018. Intrinsic functional connectivity in emotion regulation network is altered in emotional laborers. *Society for Neuroscience*.

Jeesung Ahn, Suhnyoung Jun, **Jonghyun Lee**, Seokyoung Min, Seung-Koo Lee, Soo Hyun Park, Sanghoon Han. 2018. Altered Emotional Attention and Brain Functional Connectivity Networks of Emotional Laborers. *Society for Neuroscience*.

In Review

Jonghyun Lee, Chang-Uk Jeong, Tianhua Zhai, Xia Ning, Li Shen, Zuyi Huang, Dokyoon Kim, Yun Kim. 2026. OptLLaVA: Lead Optimization Using Lightweight Language Model and Versatile Assistants. *Under review at Journal of Cheminformatics*.

Jonghyun Lee, Sang-Hyuk Jung, ..., Dokyoon Kim. 2026. Large-Scale Image-Derived Environmental Signatures for Personalized Risk Prediction. *Under review at Nature Human Behaviour*.

In Preparation

Jonghyun Lee, Dokyoon Kim. 2026. MoCA: Multimodal Cytometry Translation using Cyclic Alignment.

Jonghyun Lee, Dokyoon Kim. 2026. Cell to Sample Cytometry Foundation Model using Mixture of Experts Network.

Jonghyun Lee, Dokyoon Kim. 2026. Personalized Inpatient Glucose Optimization via Offline Reinforcement Learning.

Jonghyun Lee, Myung-Giun Noh. 2026. Phenotypic Deconstruction of Lymphovascular Invasion Using Active Learning and Virtual Multiplex Immunoprofiling.

Patents

International Patent (PCT)

2024. Image Processing Method and Device Using Reinforcement Learning. *PCT International Application*, Filed; PCT/KR2025/007948.

Korean Patent

2024. Image Processing Method and Device Using Reinforcement Learning. *Korea Patent*, Granted; 1020240117625.

2024. A method for detecting sports events and system performing the same. *Korea Patent*, Granted; 1026246500000.

2023. Method for determining tissue invasion from pathological image based on deep learning and electronic device thereof. *Korea Patent*, Filed; 1020230050276.

2023. Method for tracking sports participants, device for tracking sports participants, and system for tracking sports participants. *Korea Patent*, Granted; 1025449720000.

2023. A method for providing exercise load information. *Korea Patent*, Granted; 1025466130000.

2023. A method for detecting sports events and system performing the same. *Korea Patent*, Granted; 1024504320000.

2022. A method for providing exercise load information. *Korea Patent*, Filed; 1020220102046.

2022. A method for providing exercise load information. *Korea Patent*, Filed; 1020220102047.

2022. A method for providing exercise load information. *Korea Patent*, Filed; 1020220102050.

2022. Method for tracking sports participants, device for tracking sports participants, and system for tracking sports participants. *Korea Patent*, Granted; 1024589380000.

Teaching Experience

Fall 2023	Computational Pathology Tutorial (6 hours) , Medical School, Korea University	Lecturer
Fall 2022	Creative Computing , Department of Engineering, Hanyang University	TA
Spring 2022	Creative Computing , Department of Engineering, Hanyang University	TA
Spring 2021	Deep Learning Tutorial (12 hours) , Aerospace Research Institute, Inha University	Lecturer

Mentoring

2024 - Present	Lina Takemaru , PhD student, University of Pennsylvania - Computational pathology
2025 - Present	Youngseok Lee , MS student, Korea University - Computational pathology
2026 - Present	Irina Zhu , PhD student, University of Pennsylvania - Computational pathology
Summer 2025	Miranda Song , PhD student, University of Pennsylvania - Prognosis prediction using multiple instance learning of CyTOF data

Academic Service

Reviewer

Journals: Nature Communications; npj Digital Medicine; npj Precision Oncology; Scientific Reports; IEEE Access; BMC Cancer; BMC Medical Informatics and Decision Making; Computer Methods in Biomechanics and Biomedical Engineering; Journal of Korean Medical Science; Yonsei Medical Journal.

Conferences: Pacific Symposium on Biocomputing (PSB); Annual International Conference on Research in Computational Molecular Biology (RECOMB); Machine Learning for Health Symposium (ML4H).