Conference on Health, Inference, and Learning (CHIL) 2024

Tom Pollard[*]   TPOLLARD@MIT.EDU
Massachusetts Institute of Technology, Cambridge, MA, USA

Edward Choi[*]   EDWARDCHOI@KAIST.AC.KR
Korea Advanced Institute of Science & Technology (KAIST), Daejeon, South Korea

Pankhuri Singhal[*]   SINGHALP@PENNMEDICINE.UPENN.EDU
University of Pennsylvania, Philadelphia, PA, USA

Michael C. Hughes[*]   MICHAEL.HUGHES@TUFTS.EDU
Tufts University, Medford, MA, USA

Elena Sizikova[*]   ELENA.SIZIKOVA@FDA.HHS.GOV
U.S. Food and Drug Administration Silver Spring, MD, USA

Bobak J. Mortazavi[*]   BOBAKM@TAMU.EDU
Texas A&M University, College Station, TX, USA

Irene Chen[*]   IYCHEN@BERKELEY.EDU
University of California, Berkeley, CA, USA

Fei Wang[*]   FEW2001@MED.CORNELL.EDU
Weill Cornell Medicine, New York, NY, USA

Tasmie Sarker[*]   TASMIE.SARKER@AHLI.CC
Association for Health Learning and Inference, New York, NY, USA

Matthew McDermott[*]   MATTHEW_MCDERMOTT@HMS.HARVARD.EDU
Harvard Medical School, Boston, MA, USA

Marzyeh Ghassemi[*]   MGHASSEM@MIT.EDU
Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

The Conference on Health, Inference, and Learning (CHIL), brings together a cross-disciplinary group of clinicians and researchers, from industry and academia, with the goal of advancing machine learning for health. CHIL has been an official conference of the Association for Health Learning and Inference (AHLI) since 2022. This volume contains proceedings of the fifth annual CHIL conference, held at Cornell Tech in New York City, US.

1. Introduction

The Conference on Health, Inference, and Learning (CHIL), brings together a cross-disciplinary group of clinicians and researchers, from industry and academia, with the goal of advancing machine learn-

ing for health (Conference on Health, Inference, and Learning (CHIL), 2024). CHIL is a platform for the machine learning for the community to gather, network, and learn, and to showcase the future directions of our field.

In the past eighteen months, machine learning and "AI" have surged into mainstream interest. It is now more important than ever to ensure that we stay grounded as researchers and practitioners, striving for innovation while keeping focus on advances that are safe, fair, and ultimately beneficial for patients and providers of care. Doing so requires collaboration between a diverse group, including clinical experts, computer scientists, and social scientists.

Through CHIL, our hope is to foster these critical collaborations and drive forward meaningful progress in machine learning for health. This volume contains the proceedings of the fifth annual CHIL conference, held at Cornell Tech in New York City, US.

---

[*] These authors contributed equally

1.1. Organizing Committee

Marzyeh Ghassemi (General Chair), Matthew Mc-Dermott (General Chair), Fei Wang (Program Chair), Irene Chen (Program Chair), George Chen (Program Sub-Chair, Research Roundtables), Xiaoxiao Li (Program Sub-Chair, Research Roundtables), Monica Agrawal (Program Sub-Chair, Doctoral Symposium), Emma Rocheteau (Program Sub-Chair, Doctoral Symposium), Tom Pollard (Proceedings Chair), Edward Choi (Proceedings Chair), Pankhuri Singhal (Proceedings Sub-Chair, Workflow), Michael C. Hughes (Senior Area Chair, Track 1), Elena Sizikova (Senior Area Chair, Track 2), Bobak Mortazavi (Senior Area Chair, Track 3), Zehra Abedi (Logistics Chair), Chengxi Zang (Logistics Sub-Chair, Local), Kaveri Thakoor (Logistics Sub-Chair, Local), Roshan Kenia (Logistics Sub-Chair, Local), Elizabeth Healey (Logistics Sub-Chair, Volunteer Coordinator), Ahmed Alaa (Finance Chair), Kai Wang (Finance Chair), Monica Munnangi (Comms Chair), Jiacheng Zhu (Comms Sub-Chair), Brian Gow (Comms Sub-Chair), Koyena Pal (Comms Sub-Chair).

2. Program

CHIL's program was designed to showcase exceptional work through invited talks, to open opportunities for debate on topical issues, and to promote collaboration and learning through roundtables, lightening talks, posters, tutorials, and a doctoral symposium.

2.1. Keynote Speakers

Our keynote speakers were selected to encompass a wide range of critical topics in machine learning for health:

Hamsa Bastani: Hamsa Bastani is an Associate Professor of Operations, Information, and Decisions at the Wharton School, University of Pennsylvania. Her research focuses on developing novel machine learning algorithms for data-driven decision-making, with applications to healthcare operations, social good, and revenue management.

Noémie Elhadad: Noémie Elhadad is Associate Professor and Chair of the Department of Biomedical Informatics at Columbia University Vagelos College of Physicians and Surgeons. Her research is at the intersection of artificial intelligence, human-centered computing, and medicine, with a focus on developing novel machine-learning methods

Nils Gehlenborg: Nils Gehlenborg is an Associate Professor of Biomedical Informatics At Harvard Medical School. His efforts are focused on improving human health by developing computational techniques and interfaces that enable scientists and clinicians to efficiently interact with biomedical data.

Samantha Kleinberg: Samantha Kleinberg is an Associate Professor in the Computer Science department at Stevens Institute of Technology, where she works on development of AI methods, with a focus on causal inference and time series data.

Sanmi (Oluwasanmi) Koyejo: Sanmi Koyejo is an Assistant Professor in the Department of Computer Science at Stanford University. His interests are in developing the principles and practice of trustworthy machine learning, focusing on applications to neuroscience and healthcare

Deborah Raji: Deborah Raji is a Mozilla fellow and Computer Science PhD student at University of California, Berkeley, where she works on questions of algorithmic auditing and evaluation.

2.2. State of Machine Learning

Last year we introduced a series of talks on the State of Machine Learning, which proved to be highly popular (Mortazavi et al.). This year, we have again invited three experts to share their perspective on the state of machine learning:

Deployments & real-world impact: Leo Celi, MD, Massachusetts Institute of Technology.

Large language models in health: Kyunghyun Cho, PhD, New York University.

Underloved innovations (Rescue From obscurity): Zak Kohane, MD, PhD, Harvard University.

2.3. Panels

Our panel sessions explore topical and controversial issues in machine learning for health, bringing together experts to share their insights:

Health economics and behavior: David Meltzer, MD, PhD, University of Chicago, Walter Dempsey, PhD, University of Michigan, F. Perry Wilson, MD, Yale School of Medicine. *Moderated by Kyra Gan, PhD, Cornell University.*

Real deployments, and how to find them: Girish Nadkarni, MD, MPH, Mount Sinai, Roy Perlis, MD, Harvard University, Ashley Beecy, MD, NewYork-Presbyterian. *Moderated by Leo Celi, PhD, Massachusetts Institute of Technology.*

3. Papers: Guidelines and Selection

Papers for the proceedings were solicited across three tracks at the intersection of health and machine learning: *Models and Methods*; *Applications and Practice*; and *Impact and Society*. The submission and review process was entirely managed through the OpenReview platform (OpenReview, 2024).

3.1. Submissions

We received a total of 116 submissions across the three tracks, up 25% on the previous year and the greatest number of submissions since CHIL began in 2020. All submissions were reviewed by a minimum of three reviewers, except where the submission was subsequently desk rejected (n=3) or withdrawn (n=13).

3.2. Peer review

CHIL follows a mutually anonymous ("double blind") review process with five broad phases:

1. Reviewer bidding
2. Assignment of papers to reviewers
3. Completion of reviews
4. Discussion, and author rebuttals
5. Decision and notification

Reviewers were asked to consider aspects of relevance, quality, originality, clarity, and significance when preparing their reviews. After reviews were completed, a discussion phase allowed open communication between authors, reviewers, and chairs. Area Chairs then created meta-reviews and recommendations based on the reviews. Final decisions were made by Senior Area Chairs in discussion with Area Chairs, Proceedings Chairs, and reviewers, where appropriate.

4. Proceedings

Thirty-nine submissions (34%) were accepted for publication in the archived conference proceedings. Some prominent themes amongst the accepted papers include:

- Clinical documentation and language models: Han et al. (2024); Kim et al. (2024a); Lee et al. (2024); Hegselmann et al. (2024); Ahsan et al. (2024); Kim et al. (2024b);

- Data integration and multi-modal analysis: Zhang et al. (2024); Ravva et al. (2024); En and Guo (2024); Behrouz and Hashemi (2024);

- Safety and fairness: Nakamura Sakai et al. (2024); Fong and Motani (2024); Hoche et al. (2024); Cusick et al. (2024); Telukunta et al. (2024); Wang et al. (2024); Wu et al. (2024);

- Waveform processing and analysis: Yan et al. (2024); Han et al. (2024); Amirshahi et al. (2024); Blanks et al. (2024);

- Wearables and health monitoring: Zhang et al. (2024); Yan et al. (2024); Wei et al. (2024); Wang et al. (2024); Kim et al. (2024b); Nzeyimana et al. (2024); Visy et al. (2024); Kasl et al. (2024);

4.1. Awards

Three papers were selected for a Best Paper Award across the conference tracks.

4.1.1. MODELS AND METHODS

Brain-Mamba: Encoding Brain Activity via Selective State Space Models by Ali Behrouz and Farnoosh Hashemi (Behrouz and Hashemi, 2024).

4.1.2. APPLICATIONS AND PRACTICE

Addressing Wearable Sleep Tracking Inequity: A new dataset and novel methods for a population with sleep disorders by Will Ke Wang, Jiamu Yang, Leeor Hershkovich, Hayoung Jeong, Bill Chen, Karnika Singh, Ali R Roghanizad, Md Mobashir Hasan Shandhi, Andrew R Spector, and Jessilyn Dunn (Wang et al., 2024).

4.1.3. IMPACT AND SOCIETY

A Cross-study Analysis of Wearable Datasets and the Generalizability of Acute Illness Monitoring Models by Patrick Kasl, Severine Soltani, Lauryn Keeler Bruce, Varun Kumar Viswanath, Wendy Hartogensis, Amarnath Gupta, Ilkay Altintas, Stephan Dilchert, Frederick M Hecht, Ashley Mason, and Benjamin L Smarr (Kasl et al., 2024).

5. CHIL 2025 in California, USA

CHIL 2025 will be held in California, USA in June. Save the date! The conference will be led by general chairs Matthew McDermott and Irene Chen.

If you would like to volunteer at the conference or serve on our committee, please contact us at info@chilconference.org.

Acknowledgments

5.1. Sponsors

5.2. Reviewers

The proceedings would not have been possible without the diligent and generous work of our reviewers.

5.2.1. List of Senior Area Chairs

Michael C Hughes, Bobak J Mortazavi, and Elena Sizikova.

5.2.2. List of Area Chairs

Prince Ebenezer Adjei, Hossein Azizpour, Jessica Dafflon, Martine De Cock, Neel Dey, Jean Feng, Zepeng Frazier Huo, Daniel Moyer, Stephen R Pfohl, Vivek Kumar Singh, Sarah Tan, Shengpu Tang, Sana Tonekaboni, Yuyin Zhou, Weimin Zhou, and Ayah Zirikly.

5.2.3. List of Reviewers

Vasundhara Agrawal, Sandeep Angara, Peniel N Argaw, Marta Avalos, Seongsu Bae, Aparna Balagopalan, Rahmatollah Beheshti, Karan Bhanot, Erika Bondareva, Vinod Kumar Chauhan, Jieshi Chen, Jichan Chung, Megan Coffee, Elliot Creager, Wangzhi Dai, Charles B Delahunt, Jessilyn Dunn, Tal El Hay, Matthew M Engelhard, Xiao Fan, Shayan Fazeli, Alex Fedorov, Stefan Feuerriegel, Walter Gerych, Mehak Gupta, Thomas Hartvigsen, Ethan Harvey, Elizabeth Healey, Stefan Hegselmann, Kyle Heuton, Max Homilius, Chang Hu, Zhe Huang, Sicong Huang, Kyunghoon Hur, Stephanie Hyland, Wisdom Oluchi Ikezogwo, Vincent Jeanselme, Emmanuel Klu, Arinbjörn Kolbeinsson, Daeun Kyung, Julien Le Kernec, Hyungyung Lee, Gyubok Lee, Heike Leutheuser, Lei Lu, Abdullah Mamun, Melissa Danielle McCradden, Chuizheng Meng, Kenny Moise, Intae Moon, Jong Hak Moon, Pablo Moreno-Muñoz, Mehul Motani, Purity Mugambi, Ajinkya K Mulay, Bret Nestor, Shahriar Noroozizadeh, Declan O'Loughlin, Jungwoo Oh, Karla Paniagua, Ioakeim Perros, Adrienne Pichon, Srivamshi Pittala, Lucía Prieto Santamaría, Jielin Qiu, Aniruddh Raghu, Preetish Rath, Emma Charlotte Rocheteau, Lorenzo A Rossi, Frank Rudzicz, Shalini Saini, Antonio-José Sánchez-Salmerón, Carlo Sansone, MohammadAli Shaeri, Afrah Shafquat, Shubhranshu Shekhar, Xiaobin Shen, Bowen Song, Yao Su, Yanchao Tan, Salvatore Tedesco, Rudraksh Tuwani, Ismael Villanueva-Miranda, Jing Wang, Cédric Wemmert, Alexander Woyczyk, Zhenbang Wu, Chenwei Wu, Tong Xia, Ran Xu, Xuhai Xu, Jiajun Xu, Chaoqi Yang, Esma Yildirim, Roozbeh Yousefzadeh, Jun Yu, Wenbin Zhang, Luna Zhang, Yuan Zhao, Helen Zhou, Houliang Zhou, and Jiacheng Zhu.

References

Hiba Ahsan, Denis Jered McInerney, Jisoo Kim, Christopher A Potter, Geoffrey Young, Silvio Amir, and Byron C Wallace. Retrieving evidence from ehrs with llms: Possibilities and challenges. volume 248 of *Proceedings of Machine Learning Research*, pages 489–505. PMLR, 2024.

Alireza Amirshahi, Jonathan Dan, Jose Angel Miranda, Amir Aminifar, and David Atienza. Fetch: A fast and efficient technique for channel selection in eeg wearable systems. volume 248 of *Proceedings of Machine Learning Research*, pages 397–409. PMLR, 2024.

Ali Behrouz and Farnoosh Hashemi. Brain-mamba: Encoding brain activity via selective state space models. volume 248 of *Proceedings of Machine Learning Research*, pages 233–250. PMLR, 2024.

Zachary Blanks, Donald E Brown, Marc A Adams, and Siddhartha S Angadi. An improved bayesian

permutation entropy estimator with wasserstein-optimized hierarchical priors. volume 248 of *Proceedings of Machine Learning Research*, pages 120–136. PMLR, 2024.

Conference on Health, Inference, and Learning (CHIL). Chil conference, 2024. URL https://chilconference.org/. Accessed: 2024-06-24.

Marika M Cusick, Glenn M Chertow, Douglas K Owens, Michelle Y Williams, and Sherri Rose. Algorithmic changes are not enough: Evaluating the removal of race adjustment from the egfr equation. volume 248 of *Proceedings of Machine Learning Research*, pages 619–643. PMLR, 2024.

Qing En and Yuhong Guo. Unsupervised domain adaptation for medical image segmentation with dynamic prototype-based contrastive learning. volume 248 of *Proceedings of Machine Learning Research*, pages 312–325. PMLR, 2024.

Kei Sen Fong and Mehul Motani. Explainable and privacy-preserving machine learning via domain-aware symbolic regression. volume 248 of *Proceedings of Machine Learning Research*, pages 198–216. PMLR, 2024.

William Han, Diana Guadalupe Gomez, Avi Alok, Chaojing Duan, Michael A Rosenberg, Douglas J Weber, Emerson Liu, and Ding Zhao. Interpretation of intracardiac electrograms through textual representations. volume 248 of *Proceedings of Machine Learning Research*, pages 7–23. PMLR, 2024.

Stefan Hegselmann, Zejiang Shen, Florian Gierse, Monica Agrawal, David Sontag, and Xiaoyi Jiang. A data-centric approach to generate faithful and high quality patient summaries with large language models. volume 248 of *Proceedings of Machine Learning Research*, pages 339–379. PMLR, 2024.

Marine Hoche, Olga Mineeva, Manuel Burger, Alessandro Blasimme, and Gunnar Ratsch. Famews: a fairness auditing tool for medical early-warning systems. volume 248 of *Proceedings of Machine Learning Research*, pages 297–311. PMLR, 2024.

Patrick Kasl, Severine Soltani, Lauryn Keeler Bruce, Varun Kumar Viswanath, Wendy Hartogensis, Amarnath Gupta, Ilkay Altintas, Stephan Dilchert, Frederick M Hecht, Ashley Mason, and Benjamin L Smarr. A cross-study analysis of wearable datasets and the generalizability of acute illness monitoring models. volume 248 of *Proceedings of Machine Learning Research*, pages 644–682. PMLR, 2024.

Kyungsu Kim, Junhyun Park, Saul Langarica, Adham Mahmoud Alkhadrawi, and Synho Do. Integrating chatgpt into secure hospital networks: A case study on improving radiology report analysis. volume 248 of *Proceedings of Machine Learning Research*, pages 72–87. PMLR, 2024a.

Yubin Kim, Xuhai Xu, Daniel McDuff, Cynthia Breazeal, and Hae Won Park. Health-llm: Large language models for health prediction via wearable sensor data. volume 248 of *Proceedings of Machine Learning Research*, pages 522–539. PMLR, 2024b.

Hyungyung Lee, Da Young Lee, Wonjae Kim, Jin-Hwa Kim, Tackeun Kim, Jihang Kim, Leonard Sunwoo, and Edward Choi. Vision-language generative model for view-specific chest x-ray generation. volume 248 of *Proceedings of Machine Learning Research*, pages 280–296. PMLR, 2024.

Bobak J. Mortazavi, Tasmie Sarker, Andrew Beam, and Joyce C. Ho, editors. *Proceedings of the Conference on Health, Inference, and Learning*, volume 209 of *Proceedings of Machine Learning Research*. PMLR.

Shinpei Nakamura Sakai, Dennis Shung, and Jasjeet S Sekhon. Enhancing collaborative medical outcomes through private synthetic hypercube augmentation: Prisha. volume 248 of *Proceedings of Machine Learning Research*, pages 55–71. PMLR, 2024.

Antoine Nzeyimana, Anthony Campbell, James M Scanlan, Joanne D Stekler, Jenna Marquard, Barry G Saver, and Jeremy Gummeson. Dosemate: A real-world evaluation of machine learning classification of pill taking using wrist-worn motion sensors. volume 248 of *Proceedings of Machine Learning Research*, pages 566–581. PMLR, 2024.

OpenReview. Openreview, 2024. URL https://openreview.net/. Accessed: 2024-06-24.

Pavan Uttej Ravva, Pinar Kullu, Mohammad Fahim Abrar, and Roghayeh Leila Barmaki. A machine learning approach for predicting upper limb motion intentions with multimodal data. volume 248 of *Proceedings of Machine Learning Research*, pages 169–181. PMLR, 2024.

Mukund Telukunta, Sukruth Rao, Gabriella Stickney, Venkata Sriram Siddhardh Nadendla, and Casey Canfield. Learning social fairness preferences from non-expert stakeholder opinions in kidney placement. volume 248 of *Proceedings of Machine Learning Research*, pages 683–695. PMLR, 2024.

Tamás Visy, Rita Kuznetsova, Christian Holz, and Shkurta Gashi. Systematic evaluation of self-supervised learning approaches for wearable-based fatigue recognition. volume 248 of *Proceedings of Machine Learning Research*, pages 582–596. PMLR, 2024.

Will Ke Wang, Jiamu Yang, Leeor Hershkovich, Hayoung Jeong, Bill Chen, Karnika Singh, Ali R Roghanizad, Md Mobashir Hasan Shandhi, Andrew R Spector, and Jessilyn Dunn. Addressing wearable sleep tracking inequity: A new dataset and novel methods for a population with sleep disorders. volume 248 of *Proceedings of Machine Learning Research*, pages 380–396. PMLR, 2024.

Hui Wei, Maxwell A Xu, Colin Samplawski, James Matthew Rehg, Santosh Kumar, and Benjamin Marlin. Temporally multi-scale sparse self-attention for physical activity data imputation. volume 248 of *Proceedings of Machine Learning Research*, pages 137–154. PMLR, 2024.

Kevin Wu, Eric Wu, Kit Rodolfa, Daniel E Ho, and James Zou. Regulating ai adaptation: An analysis of ai medical device updates. volume 248 of *Proceedings of Machine Learning Research*, pages 477–488. PMLR, 2024.

Runze Yan, Cheng Ding, Ran Xiao, Alex Fedorov, Randall J Lee, Fadi Nahab, and Xiao Hu. Squwa: Signal quality aware dnn architecture for enhanced accuracy in atrial fibrillation detection from noisy ppg signals. volume 248 of *Proceedings of Machine Learning Research*, pages 105–119. PMLR, 2024.

Haoting Zhang, Donglin Zhan, Yunduan Lin, Jinghai He, Qing Zhu, Zuo-Jun Shen, and Zeyu Zheng. Daily physical activity monitoring: Adaptive learning from multi-source motion sensor data. volume 248 of *Proceedings of Machine Learning Research*, pages 39–54. PMLR, 2024.