

Conference on Health, Inference, and Learning (CHIL) 2023

Bobak J. Mortazavi

Texas A&M University

BOBAKM@TAMU.EDU

Tasmie Sarker

Association for Health Learning and Inference

TASMIE.SARKER@AHLI.CC

Andrew Beam

Harvard University

ANDREW_BEAM@HMS.HARVARD.EDU

Joyce C. Ho

Emory University

JOYCE.C.HO@EMORY.EDU

List of organizers Matthew McDermott (Program Chair), Emily Alsentzer (Program Chair), Mike Hughes (Track Lead), Lifang He (Track Lead), Sanja Šćepanović (Track Lead), Yuyin Zhou (Track Chair), Rahul Krishnan (Track Chair), Jean Feng (Track Chair), Samantha Kleinberg (Track Chair), Elena Sizikova (Track Chair), Tom Pollard (Track Chair), Carl Yang (Track Chair), Yu Zhang (Track Chair), Stephen Pfohl (Track Chair), Dimitris Spathis (Track Chair), Bobak Mortazavi (Proceedings Chair), Huan He (Technology Chair), Jiayu Yao (Technology Chair), Brian Gow (Virtual Chair), Ioakeim Perros (Communications Chair), Anil Palepu (Communications Chair), Monica Munnangi (Logistics Chair), Tasmie Sarker (Logistics Chair), Jessica Gronsbell (Unconference Chair), Rui Duan (Unconference Chair), Edward Choi (Finance Chair), Harvinder Singh (Finance Chair), Tom Hartvigsen (Doctoral Symposium), Andrew Beam (General Chair), and Joyce Ho (General Chair)

1. Introduction

This volume contains the proceedings of the third Conference on Health, Inference, and Learning (CHIL), held in Cambridge, MA, United States on June 22–24, 2023. This is the second year of the conference being part of the Association for Health Learning and Inference (AHLI). This has continued to enable CHIL to provide access to our proceedings through open publication models.

Research in machine learning and health requires a cross-disciplinary representation of clinicians and researchers in machine learning, health policy, causality, fairness, and related areas. The goal of the con-

ference is to foster excellent research that addresses the unique challenges and opportunities that arise at the intersection of machine learning and health.

2. Conference

The CHIL 2023 program included a set of exceptional keynote speakers, invited talk, panel presentations, intentionally contentious debates, roundtables, lightning talks, doctoral symposium, in addition to the spotlight presentations and posters by authors of accepted works submitted to CHIL.

2.1. Keynotes

CHIL 2023 introduced excellent keynote speakers talks, 35 minutes each, that covered a wide variety of topics.

Dina Demner-Fushman Dr. Demner-Fushman is a tenure track investigator in the Computational Health Research Branch at the Lister Hill National Center for Biomedical Communications in the National Institutes of Health. In her talk, titled “Biomedical Question Answering Yesterday, Today, and Tomorrow,” provided an overview of the approaches to biomedical question answering as they were evolving. It covered information needs of various stakeholders and the resources created to address these information needs through Question Answering.

Ben Glocker Dr. Glocker is Professor in Machine Learning for Imaging and Kheiron Medical Technologies, Royal Academy of Engineering Research Chair

in Safe Deployment of Medical Imaging AI. He co-leads the Biomedical Image Analysis Group, leads the HeartFlow-Imperial Research Team, and is Head of ML Research at Kheiron. In his talk, titled “Safe Deployment of Medical Imaging AI,” he discussed AI safeguards from the perspective of robustness, reliability, and fairness. He explored approaches for automatic failure detection, monitoring of performance, and analysis of bias, aiming to ensure the safe and ethical use of medical imaging AI.

Jun Cheng Dr. Cheng is a Senior Research Scientist at DeepMind. His research focused on developing machine learning methods to better understand the genetic code and disease mechanisms. In his talk, titled “Biological Sequence Modeling in Research and Applications,” he provided an overview of the current techniques and status of biological sequences modeling. Additionally, specific applications of such models in genetics and immunology were discussed.

Munmun De Choudhury Dr. De Choudhury is an Associate Professor of Interactive Computing at Georgia Tech and has been recognized with the 2023 SIGCHI Societal Impact Award, the 2022 Web Science Trust Test-of-Time Award, the 2021 ACM-W Rising Star Award, the 2019 Complex Systems Society. In her talk, titled “Bridging Machine Learning and Collaborative Action Research: A Tale Engaging with Diverse Stakeholders in Digital Mental Health,” she described the experiences from working with different stakeholders in research initiatives relating to digital mental health – including with healthcare providers, grassroots advocacy and public health organizations, and people with the lived experience of mental illness. The talk presented lessons learned by way of these engagements, and to reflect on a path forward that empowers us to go beyond technical innovations to envisioning contributions that center humans’ needs, expectations, values, and voices within those technical artifacts.

Dina Katabi Dr. Katabi is the Thuan and Nicole Pham Professor of Electrical Engineering and Computer Science at MIT. She is also the director of the MIT’s Center for Wireless Networks and Mobile Computing, a member of the National Academy of Engineering, and a recipient of the MacArthur Genius Award.

Roxana Daneshjou Dr. Daneshjou (MD, PhD) is an incoming assistant professor of biomedical data science and dermatology at Stanford (Fall of 2023)

and her research interests are in developing diverse datasets and fair algorithms for applications in precision medicine. In her talk, titled “Skin in the Game: The State of AI in Dermatology,” she discussed the opportunities and challenges for AI in dermatology. Artificial intelligence tools have been touted as having performance “on par” with board certified dermatologists. However, these published claims have not translated to real world practice.

Emma Pierson Dr. Pierson is an assistant professor of computer science at the Jacobs Technion-Cornell Institute at Cornell Tech and the Technion, and a computer science field member at Cornell University. She holds a secondary joint appointment as an Assistant Professor of Population Health Sciences at Weill Cornell Medical College and has been recognized with an NSF CAREER award, a Rhodes Scholarship, Hertz Fellowship, Rising Star in EECS, MIT Technology Review 35 Innovators Under 35, and Forbes 30 Under 30 in Science. In her talk, titled “Using Machine Learning to Increase Equity in Healthcare and Public Health,” she discussed how data science and machine learning can be used to combat inequality in health care and public health by presenting several vignettes from domains like medical testing and cancer risk prediction.

2.2. Invited Talks

CHIL 2023 hosted an inaugural series of three invited talks on “The State of Machine Learning for Health: Where Are We Now, and Where Do We Go?”

Suchi Saria Dr. Saria is the John C. Malone endowed chair and is the Director of the Machine Learning, AI and Healthcare Lab at Johns Hopkins. She is also is the Founder and CEO of Bayesian Health. She highlighted recent technical advances in research in her talk titled “Research and Top Recent Papers from 2020-2022.”

Karandeep Singh Dr. Singh is an Assistant Professor of Learning Health Sciences, Internal Medicine, Urology, and Information at the University of Michigan. He discussed advances to close the translational gap in his talk titled “Recent Deployments and Real-World Impact.”

Nigam Shah Dr. Shah is a Professor of Medicine at Stanford University, and Chief Data Scientist for Stanford Health Care. He featured high-impact but under-studied research questions in his talk titled

“Under-explored Research Challenges and Opportunities.”

2.3. Panels

CHIL 2023 featured 3 panel discussions.

Sharing Health Data in an Age of Generative AI: Risks, Limitations, and Solutions. A discussion about the ways that health datasets were shared in both commercial and academic setting and how the community can navigate acquiring data in a safe, responsible, and accessible manner or ways to enable learning without data release. The panelists included Ziad Obermeyer, an Associate Professor and Blue Cross of California Distinguished Professor at UC Berkeley; John Halamka, an emergency medicine physician, medical informatics expert and president of the Mayo Clinic Platform; Elaine Nsoesie, an Associate Professor at Boston University’s School of Public Health; and Khaled El Emam, a Professor in the School of Epidemiology and Public Health at the University of Ottawa and was moderated by Marzyeh Ghassemi, an assistant professor and the Hermann L. F. von Helmholtz Professor at MIT.

Generalizability in Machine Learning for Health: Critical for Robustness, or a Distraction from Specific Validation? A discussion of the extent to which the community should focus on developing generalizable ML models including how to validate generalizability in deployments and regulatory implications of developing generalizable versus context-specific models. The panelists included Leo Celi, the clinical research director and principal research scientist at the MIT Laboratory for Computational Physiology; Jason Fries, a research scientist in the Shah Lab at Stanford University; Lauren Oakden-Rayner, a radiologist and Senior Research Fellow at the Australian Institute for Machine Learning, University of Adelaide; and Maia Hightower, the executive Vice President and Chief Digital & Technology Officer at the University of Chicago Medicine and the CEO and co-founder of Equality AI and was moderated by Isaac Kohane, the inaugural chair of Harvard Medical School’s Department of Biomedical Informatics.

Machine Learning for Healthcare in the Era of ChatGPT. A discussion on the emergence of large language models and the impact they will have on AI for medicine. The panelists included Karandeep Singh and Nigam Shah and was moderated by

Byron Wallace, the Sy and Laurie Sternberg Interdisciplinary Associate Professor and Director of the BS in Data Science program at Northeastern University in the Khoury College of Computer Sciences.

2.4. Debates

CHIL 2023 hosted 3 intentionally contentious debates as a new part of the program.

Network studies: As many databases as possible or enough to answer the question quickly? The debate involved Dr. Christopher Chute and Dr. Robert Platt. Dr. Chute, the Bloomberg Distinguished Professor of Health Informatics, Professor of Medicine, Public Health, and Nursing at Johns Hopkins University, and Chief Research Information Officer for Johns Hopkins Medicine. Dr. Platt is Professor in the Departments of Epidemiology, Biostatistics, and Occupational Health, and of Pediatrics, at McGill University. He holds the Albert Boehringer I endowed chair in Pharmacoepidemiology, and is Principal Investigator of the Canadian Network for Observational Drug Effect Studies (CNODES).

Data Heterogeneity: More Heterogeneous Data or Less Homogeneous Data? The debate involved Dr. Tianxi Cai and Dr. Yong Chen. Dr. Cai is John Rock Professor of Translational Data Science at Harvard, with joint appointments in the Biostatistics Department and the Department of Biomedical Informatics. She directs the Translational Data Science Center for a Learning Health System at Harvard Medical School and co-directs the Applied Bioinformatics Core at VA MAVERIC. Dr. Chen is Professor of Biostatistics at the Department of Biostatistics, Epidemiology, and Informatics at the University of Pennsylvania (Penn). He directs a Computing, Inference and Learning Lab at University of Pennsylvania, which focuses on integrating fundamental principles and wisdoms of statistics into quantitative methods for tackling key challenges in modern biomedical data.

Differential Privacy vs. Synthetic Data The debate involved Dr. Khaled El Emam and Dr. Li Xiong. Dr. El Emam is the Canada Research Chair (Tier 1) in Medical AI at the University of Ottawa, where he is a Professor in the School of Epidemiology and Public Health. He is also a Senior Scientist at the Children’s Hospital of Eastern Ontario Research Institute and Director of the multidisciplinary Electronic Health Information Labora-

tory, conducting research on privacy enhancing technologies to enable the sharing of health data for secondary purposes, including synthetic data generation and de-identification methods. Dr. Xiong is a Samuel Candler Dobbs Professor of Computer Science and Professor of Biomedical Informatics at Emory University. She held a Winship Distinguished Research Professorship from 2015-2018. She has served and serves as associate editor for IEEE TKDE, IEEE TDSC, and VLDBJ, general co-chair for ACM CIKM 2022, program co-chair for IEEE BigData 2020 and ACM SIGSPATIAL 2018, 2020, program vice-chair for ACM SIGMOD 2024, 2022, and IEEE ICDE 2023, 2020, and VLDB Sponsorship Ambassador and is an IEEE Fellow.

2.5. Roundtables

CHIL 2023 hosted 5 research roundtables, including 3 focused on the continuation of the intentionally contentious debates topics of Network Studies, Data Heterogeneity, and Data Privacy.

Bridging the gap between the business of value-based care and the research of health AI by Yubin Park. This roundtable focused on Value-Based Care (VBC) and the challenges of satisfying 1) better care for individuals, 2) better health for populations, and 3) lower cost.

Auditing Algorithm Performance and Equity by Ziad Obermeyer. This roundtable focused on the opportunities and challenges of auditing algorithm performance and equity.

2.6. Doctoral Symposium

CHIL 2023 provided an opportunity for PhD students to broadcast their research and get feedback on their directions from CHIL attendees and leaders in the field. Participants presented a poster and had opportunities to connect with and meet established researchers throughout the conference.

Participants. PhD students presenting a poster on their dissertation include Yasmeena Akhter, Vishwali Mhasawade, Yuan Zhao, Weston Hughes, Zhe Huang, Shira Zilberstein, Changye Li, Aziliz Cottin, Vincent Jeanselme, Wenjing Ma, Jiaying Lu, Jiacheng Zhu, Sanjana Kulkarni, Peniel Argaw, Md Saiful Islam, Shengpu Tang, Dongyeop Jang, Md Mozaharul Motalib, Fahmida Liza Piya, Bo Liu, Kyle Heuton, Divya Shanmugam, Vinyas Harish, Mirza Farhan Bin

Tarek, Lief Pagalan, Matthew Peroni, Jiawei He, Mahapara Khurshid, Bradley Segal, Helen Coupland, Xin Huang, Thomas Beaney, Elizabeth Healey, Akshay Mohan, Shekufeh Shafeie, Preetish Rath, Van Truong, Xing Han

Mentors. We are grateful to all the mentors who volunteered their time: Megan Coffee, Roy Adams, Mariia Sidulova, Bill Lotter, Niharika DSouza, Sumit Mukherjee, Valentyn Stadnytskyi, Dominik Dahlem, Omid Bazgir, Joyce Ho, Sanja Šćepanović, Michael Hughes, Edward Choi, Stephen Pfohl, Swami Sankaranarayanan, Matthew McDermott, Shalmali Joshi, Fraizer Huo, and Irene Chen.

2.7. Lightning Talks

CHIL 2023 hosted 5 5-minute lightning talks on the topic “I Can’t Believe It’s Not Better” that highlighted ML in health failures that still surprise. The talks were by David Bellamy, Bhawesh Kumar, Cindy Wang, and Andrew Beam on whether pre-trained transformers beat simple baselines on lab data; Hiba Ahsan, Silvio Amir, Byron Wallace on the difficulty of disentangling race in representations of clinical notes; Olga Demler on intransitivity of win ratio and area under the receiver operating characteristics curve; Wouter van Amsterdam, Rajesh Ranganath on whether an accurate prediction model will be useful for treatment decision making; and Yuan Zhao, David Benkeser, Russell Kempker on how doubly robust approaches for estimating treatment effect in observational studies are not better than g-computation.

3. Papers: Guidelines and Selection

We now discuss the guidelines for paper submissions and then present submission and reviewer statistics.

3.1. Paper submission guidelines

This year we continued with the mandatory requirement of the two sections: (1) Data and Code Availability, and (2) Institutional Review Board (IRB). Authors had to comment on these even if they were not planning on sharing code or data, or if their paper did not require IRB approval. Our goal was to better highlight what data are being used (with citations as appropriate), to promote sharing of code (authors who declined to share code had to explicitly say that they are not sharing code), and to make sure that

authors did provide IRB information or clearly state that their research does not require an IRB approval.

3.2. Submission statistics

We received 92 total submissions across three submission tracks. The total number of submissions increased from 71 in 2022, down from 110 in 2021 and equal 92 in 2020.

Out of the 92 papers, a strong set of 33 papers were selected for inclusion in the proceedings (35.9% acceptance rate). Each paper received at least two reviews with an average number of 4.4 reviews. Each submission also received a meta-review that reflected collaborative decision making both within each track and across all three tracks. This led to more consistent acceptance decisions.

3.3. Reviewer statistics

We had a total of 134 reviewers. To improve the matching process, we invited reviewers to bid on the papers during a 4-day window prior to paper assignment. Each reviewer was assigned an average of 2.93 papers. During the meta-review process we noted quality of review to continue to refine our reviewer pool.

4. Acknowledgements

We thank AHLI for their support and suggestions for this year’s CHIL.

4.1. Sponsors

CHIL 2023 was sponsored by ApolloMed, Dandelion Health, Genentech, Apple, Sage Bionetworks, and [AI]TRICS.

4.2. List of Reviewers

The proceedings would not have been possible without the diligent and generous work of our reviewers. They provided meaningful feedback and evaluation of all proceedings papers. We are grateful to all the reviewers: Lucia Chen, Adrienne Pichon, Agata Foryciarz, Ajay Subramanian, Ajinkya K Mulay, Alex Fedorov, Alexej Gossmann, Aniruddh Raghu, Anna Zink, Aparna Balagopalan, Arinbjörn Kolbeinsson, Arjun Sondhi, Bobak J Mortazavi, Bret Nestor, Chaoqi Yang, Chen Lin, Chen Yanover, Chuizheng Meng, Danny Eytan, David Madras, Edward Choi,

Elizabeth Healey, Elliot Creager, Emma Charlotte Rocheteau, Emmanuel Klu, Erika Bondareva, Fahad Kamran, Faysal Hossain Shezan, Frank Rudzicz, Hammaad Adam, Haoran Zhang, Hejie Cui, Helen Zhou, Houliang Zhou, Ioakeim Perros, Ismael Villanueva-Miranda, Ivana Malenica, Jason Alan Fries, Jessica Dafflon, Jiacheng Zhu, Jielin Qiu, Jieshi Chen, Joel Persson, Jun Yu, Katharina V Hoebel, Kejing Yin, Kruti Pandya, Kyle Heuton, Lorenzo A. Rossi, Lucia Chen, Lucía Prieto Santamaría, Luna Zhang, Marta Avalos, Mathias Unberath, Matthew M. Engelhard, Rakib Islam, Megan Coffee, Mehak Gupta, Mehul Motani, Melanie F. Pradier, Melissa Danielle McCradden, Michael W Dusenberry, Mollie Marian Mckillop, Monica Agrawal, Pablo Moreno-Muñoz, Pavitra Krishnaswamy, Peniel N Argaw, Preetish Rath, Rahmatollah Beheshti, Ran Xu, Richard J. Chen, Roozbeh Yousefzadeh, Roxana Daneshjou, Rudraksh Tuwani, Ruizhi Liao, Rumi Chunara, Sandeep Angara, Seongsu Bae, Shalini Saini, Shayan Fazeli, Shubhranshu Shekhar, Stefan Feuerriegel, Stefan Heggemann, Stephanie Hyland, Sumit Mukherjee, Sungjin Park, Tal El Hay, Taylor W. Killian, Thomas Hartvigsen, Tong Xia, Vincent Jeanselme, Vineeth S Bhaskara, Wangzhi Dai, Wenbin Zhang, William Boag, Xiaolei Huang, Xuan Kan, Yanchao Tan, Yao Su, Yates Coley, Yonatan Dov Mintz, Zhe Huang, Zhenbang Wu, Zhixuan Chu.