U.S. FOOD & DRUG ADMINISTRATION

MAQC Massive Analysis and Quality Control Society

FDA/NCTR-MAQC 2022 CONFERENCE

(5th Annual Meeting of the MAQC Society)

VENUE: U.S. FDA WHITE OAK CAMPUS (SILVER SPRING, MD)
GREAT CONFERENCE ROOM

SEPTEMBER 26-27, 2022

PROGRAM COMMITTEE

Wendell Jones, PhD
Joshua Xu, PhD
Rebecca Kusko, PhD
Joaquin Dopazo, PhD
Samir Lababidi, PhD
Shraddha Thakkar, PhD
Dongying Li, PhD

BOARD OF DIRECTORS

Executive Chair:

Wendell Jones, PhD, Q2 Solutions, USA

Deputy Chair:

Rebecca Kusko, PhD, Immuneering Corporation, USA

Treasurer:

Wenjun Bao, PhD, JMP Statistical Discovery-SAS Institute Inc., USA

Executive Secretary:

Joshua Xu, PhD, US FDA, USA

President:

Joaquin Dopazo, PhD, Fundacion Progreso y Salud, Spain

Vice-President:

Xiaohui Fan, PhD, Zhejiang University, China

President-Elect:

Shraddha Thakkar, PhD, US FDA, USA

Chair Emeritus:

Weida Tong, PhD, US FDA, USA

Chair of Awards Committee:

Cesare Furlanello, PhD, HK3 Lab S.R.L & Orobix Life, Italy

Board Members:

Ben Haibe-Kains, PhD, Princess Margaret Cancer Centre, Canada
Leming Shi, PhD, Chief Science Officer, Fudan University, China
Russ Wolfinger, PhD, JMP Statistical Discovery-SAS Institute Inc., USA
Chris Mason, PhD, Weill Cornell Medicine, USA
Matthias Fisher, PhD, Cologne University, Germany
Susanna-Assunta Sansone, PhD, University of Oxford, UK
Andreas Scherer, PhD, Business Liaison at Institute for Molecular Medicine, Finland
Aedín C. Culhane, PhD, Professor of Cancer Genomics, University of Limerick, Ireland
Tim Mercer, PhD, Garvan Institute of Medical Research, Australia

FDA/NCTR-MAQC 2022 CONFERENCE

(5th Annual Meeting of the MAQC Society)

Venue: U.S. FDA White Oak Campus Great Conference Room (Silver Spring, MD)

AGENDA

Note: All times are shown in Eastern Time Zone.

DAY 1 MONDAY SEPTEMBER 26, 2022

8:00am-8:45am	Registration	
8:45am-10:25am	**Opening Session – Room B**	
8:45am-9:00am	**Welcome and Introduction**	Dr. Wendell Jones *Q^2 Solutions* Dr. Tucker Patterson *FDA/NCTR*
9:00am-9:10am	**Opening Remarks**	Dr. Namandjé Bumpus FDA Chief Scientist
9:10am-9:30am	"From Genomics to AI – Reflections on the MAQC's Journey to Improve Reproducibility in Life and Health Science"	MAQC Panel Dr. Leming Shi *Fudan Univ.* Dr. Benjamin Haibe-Kains *Princess Margaret Cancer Center*
9:30am-10:15am	Special Invited Speaker – "Leakage and the Reproducibility Crisis in ML-based Science"	Sayash Kapoor *Princeton University*
10:15am-10:25am	Q&A	
10:25am-10:40am	**Refreshment Break**	
10:40am-12:01pm	**Parallel Sessions I-A and I-B**	
	Session I-A: XAI Programs - Room A	Session Chair: Dr. Cesare Furlanello *HK3 Lab*
10:40am-11:20am	Special Invited Speaker – "DARPA's explainable AI (XAI) program: A retrospective"	Dr. Matt Turek *DARPA*
11:20am-11:50am	"Drug repurposing for COVID-19 using explainable machine learning and mechanistic models of signal transduction circuits related to SARS-CoV-2 infection with real world data validation"	Dr. Carlos Loucera *Andalusian Public Foundation Progress & Health, Spain*
11:50am-12:00pm	Q&A	

	Session I-B: Real World NGS Challenges - Room B	**Session Chair:** Dr. Joshua Xu *FDA/NCTR*	
10:40am-11:10am	"Beyond SEQC2 – Next Phase with Real World Data"	Dr. Don Johann *UAMS Health*	
11:10am-11:40am	"Some nonobvious factors that enhance or diminish observed and theoretical Illumina library complexity"	Dr. Wendell Jones *Q^2 Solutions*	
11:40am-11:50am	Q&A		
12:00pm-1:15pm	**Lunch Break**		
1:15pm-3:00pm	**Parallel Sessions II-A and II-B (Monday, September 26)**		
	Session II-A: Reproducible AI/ML in Toxicology - Room A	**Session Chair:** Dr. Shraddha Thakkar *FDA/CDER*	
1:15pm-1:40pm	"AI or Animal – A reproducibility perspective"	Dr. Weida Tong *DBB Director, FDA/NCTR*	
1:40pm-2:05pm	"Reproducible AI for supporting regulatory applications – a case study"	Dr. Ting Li *FDA/NCTR/DBB*	
2:05pm-2:30pm	"Enhancing reproducibility of language models in BERTox"	Dr. Leihong Wu *FDA/NCTR/DBB*	
2:30pm-2:55pm	"Improving preclinical pathology analysis with PathologAI"	Dr. Cesare Furlanello *HK3 Lab*	
2:55pm-3:00pm	Q&A		
	Session II-B: Tools and Methods for Reproducibility + "Hot Topics" (Posters) - Room B	**Session Chair:** Dr. Rebecca Kusko *Immuneering*	
1:15pm-1:40pm	"Automated bioinformatic software testing and benchmarking to standardize the development of accurate, reproducible, and scalable omics analysis tools"	Gwenn Berry *Magna Labs, Inc.*	
1:40pm-2:05pm	"Reproducible toxicogenomics analysis in the three-sample scenario"	Dr. Dongying Li *FDA/NCTR/DBB*	
2:05pm-2:10pm	Q&A		
2:10pm-2:25pm	"LeaSH: a High-Throughput COVID-19 Test to Profile Viral and Host Transcriptomes by Next Generation Sequencing" (Poster)	Dr. Oswaldo Lozoya *NIH/NIEHS*	
2:25pm-2:40pm	"Evaluation of mutagenic susceptibility of different stages in germ cell development of Caenorhabditis elegans using next generation sequencing" (Poster)	Dr. Tao Chen *FDA/NCTR/DGMT*	
2:40pm-2:55pm	"Explainable approach for species identification using LIME" (Poster) - cancelled	Dr. Aarti Milind Karande *S.P.I.T.*	
2:55pm-3:00pm	Q&A		

3:00pm-3:20pm	**Refreshment Break and Posters - Room C**	
3:20pm-5:20pm	**Session III: precisionFDA NCTR Indel Challenge - Room B**	**Session Chair:** Dr. Samir Lababidi *FDA/ODT*
3:20pm-3:40pm	"precisionFDA overview"	Dr. Samir Lababidi *FDA/ODT*
3:40pm-3:58pm	"NCTR indel calling challenge from oncopanel sequencing data"	Dr. Binsheng Gong *FDA/NCTR/DBB*
3:58pm-4:00pm	Announcement of Top Performers	Sarah Prezek *FDA/CDER*
4:00pm-4:15pm	Top performer's presentation #1	Mahdi Golkaram *DRAGEN team at Illumina*
4:15pm-4:30pm	Top performer's presentation #2	Konstantinos Kyriakidis
4:30pm-4:45pm	Top performer's presentation #3	Limin Peng *Genetalks*
4:45pm-5:00pm	Top performer's presentation #4	Colin Hercus *Novocraft*
5:00pm-5:10pm	Q&A	
5:15pm-5:20pm	**Closing Announcements and Adjourn - Room B**	

DAY2 TUESDAY SEPTEMBER 27, 2022

8:00am-8:45am	Registration	
8:45am-9:35am	**Opening Keynote - Room B**	
8:45am-8:50am	Announcements and Introduction	Dr. Wendell Jones *Q^2 Solutions*
8:50am-9:30am	Keynote: "Reproducibility and Transparency in Medical AI"	Dr. Hugo Aerts *Harvard Univ., Dana-Farber*
9:30am-9:35am	Q&A	
9:35am-10:40am	**Session IV MAQC Society Awards - Room B**	
9:35am - 9:40am	Award announcements – MAQC Society Award and Outstanding Reproducibility in Science Award MAQC Society Award – Dr. Ira Deveson, Garvan Institute of Medical Research Outstanding Reproducibility in Science Award – Dr. Marzyeh Ghassemi, MIT	Dr. Rebecca Kusko *Immuneering* Dr. Cesare Furlanello *HK3 Lab*
9:40am-10:10am	Outstanding Reproducibility in Science Award – "Reproducibility Challenges in AI/ML"	Dr. Marzyeh Ghassemi *MIT*
10:10am-10:40am	Outstanding Reproducibility in Science Award Runner-up – "Reproducibility Issues with Pre-Clinical Data"	Dr. Brian Nosek *University of Virginia*
10:40am-10:50am	**Refreshment Break**	

10:50am-12:00pm	**Parallel Sessions V-A and V-B**	
	Session V-A: Single Cell Annotation and Inference - Room A	**Session Chair:** Dr. Wendell Jones, *Q^2 Solutions*
10:50am-11:20am	"De novo analysis of bulk RNA-seq data at spatially resolved single-cell resolution"	Dr. Jie Liao, *Zhejiang University, China*
11:20am-11:50am	"A signaling-informed neural network for scRNA-seq annotation of known and unknown cell types"	Pelin Gundogdu, *University of Seville/ Andalusian Public Foundation Progress and Health-FPS, Spain*
11:50am-12:00pm	Q&A	
	Session V-B: Enhancing Reproducibility of and Confidence in Studies - Room B	**Session Chair:** Dr. Joshua Xu, *FDA/NCTR*
10:50am-11:20am	"Systematizing Confidence in Open Research and Evidence (SCORE)"	Dr. Brian Nosek, *University of Virginia*
11:20am-11:50pm	"The FAIR Cookbook: pharmas and academics join forces to make data FAIR"	Dr. Susanna-Assunta Sansone, *University of Oxford*
11:50am-12:00pm	Q&A	
12:00pm-1:00pm	**Lunch Break**	
1:00pm-2:00pm	**Afternoon Keynote - Room B**	
1:00pm-1:45pm	Keynote: "The Mythos of Model Interpretability"	Dr. Zachary Lipton, *Carnegie Mellon Univ.*
1:45pm-2:00pm	Q&A and Refreshment Break	
2:00pm-3:40pm	**Parallel Sessions VI-A and VI-B**	
	Session VI-A: SEQC2 Targeted RNA-seq - Room A	**Session Chair:** Dr. Wendell Jones, *Q^2 Solutions*
2:00pm-2:30pm	"Targeted RNA-seq for small variant detection to enhance precision medicine"	Dr. Dan Li, *FDA/NCTR/DBB*
2:30pm-3:00 pm	"Integrated Calling and Evaluation of Gene Fusion Detection by Targeted RNA Sequencing of Reference Samples with Long and Short Reads"	Dr. Joshua Xu, *FDA/NCTR*
3:00pm-3:30pm	"Resolving complex gene expression and splicing with targeted long-read RNA sequencing"	Dr. David P Kreil, *Boku University, Austria* Dr. Pawel Labaj, *Jagiellonian University, Poland*
3:30pm-3:40pm	Q&A	

	Session VI-B: New Single-Cell Assays and Tools + "Hot Topics" (Posters) - Room B	**Session Chair:** Dr. Rebecca Kusko *Immuneering*
2:00pm-2:30pm	"Knowledge-graph-based cell-cell communication inference for spatially resolved transcriptomic data with SpaTalk"	Dr. Xin Shao *Zhejiang University, China*
2:30pm-3:00pm	"Mutual genome and transcriptome enrichment and sequencing in single cells with ResolveOME amplification chemistry to illuminate complex biological mechanisms"	Dr. Isai Salas-Gonzalez *BioSkryb Genomics*
3:00pm-3:05pm	Q&A	
3:05pm-3:20pm	"Using Inherently Interpretable Models to Understand Sources of Variant Calling Errors" (Poster)	Dr. Nathan Dwarshuis *NIST*
3:20pm-3:35pm	"Radiological heterogeneity of intrapatient lesion-specific responses to chemotherapy is associated with survival in advanced leiomyosarcoma (Poster)"	Caryn Geady *Princess Margaret Cancer Center*
3:35pm-3:50pm	**Refreshment Break and Posters - Room C**	
3:50pm-5:05pm	**Closing Session - Room B**	
3:50pm-4:00pm	Plans for 6th Annual Meeting of the MAQC Society in 2023	
4:00pm-5:00pm	Discussion for new MAQC Society Working Groups	
5:00pm-5:05pm	**Closing Remarks**	

AWARD WINNERS

2022

MAQC Society Award

Winner: Ira Deveson, PhD

Outstanding Reproducibility in Science Award

Winner: Marzyeh Ghassemi, PhD
Runner-up: Brian Nosek, PhD

2021

MAQC Society Award

Winner: Charles Wang, MD, PhD, MPH
Runner-up: Benjamin Haibe-Kains, PhD

Outstanding Reproducibility in Science Award

Winner: Davide Chicco, PhD
Runner-up: Tsuyoshi Miyakawa, PhD

MAQC HISTORY

Since 2005, the FDA has led the MicroArray Quality Control (MAQC) consortium, which is a community wide effort to address reproducibility issues relating to the application of constantly evolving high-throughput genomics technologies to either assess safety and efficacy of FDA regulated products and their safe and effective use in clinical applications. The MAQC consortium completed three projects between 2005 -2014 (namely MAQC-I, II and III), resulting in ~30 publications. The fourth consortium, termed the Sequencing Quality Control Phase 2 (SEQC2) project is still on-going resulting in over 20 publications in several journals over the last two years. The MAQC Society grew out of the MAQC consortium.

MAQC MISSION

The objective of the MAQC Society is to communicate, promote, and advance reproducible science principles and quality control for analysis of the massive data generated from the existing and emerging technologies in solving biological, health, and medical problems. Thus, the goals of the Society are to

(1) advocate and facilitate the development and application of quality control practice and standard analysis protocols of bioinformatics and biostatistics for enhanced reproducibility across multiple experiments, laboratories, and data analysis methods, and

(2) advance our understanding and best practices in the analysis of massive data from emerging technologies applied in drug development, clinical application, and safety/risk assessment.

KEYNOTE, INVITED, AND MAQC SOCIETY AWARD SPEAKER BIOGRAPHIES

Dr. Namandjé N. Bumpus (FDA Chief Scientist) – Opening Remarks
Monday, September 26, 2022, 9:00am-9:10am

Dr. Namandjé N. Bumpus was named as the FDA's Chief Scientist on June 30, 2022. The Office of the Chief Scientist supports the research foundation, science, and innovation that underpins the FDA's regulatory mission. It does this through a broad framework that encompasses scientific collaborations, laboratory safety, the transfer of FDA inventions to the private sector, scientific integrity in FDA policy- and decision-making, the professional development of regulatory scientists, and its core research component — the FDA's National Center for Toxicological Research—which generates the vital data that the FDA requires for its regulatory decision-making and development of sound regulatory policy.

Before joining the FDA, Dr. Bumpus was the E.K. Marshall and Thomas H. Maren Professor and chair of the Department of Pharmacology and Molecular Sciences at the Johns Hopkins University School of Medicine. She served previously as associate dean for basic research in the Johns Hopkins University School of Medicine. Dr. Bumpus' research has focused on drug metabolism, pharmacogenetics, bioanalytical chemistry, and infectious disease pharmacology. Dr. Bumpus joined the faculty at Johns Hopkins in 2010 as an assistant professor. She earned a bachelor's degree in biology at Occidental College in 2003, a doctorate in pharmacology at the University of Michigan in 2007 and completed a postdoctoral fellowship in molecular and experimental medicine at The Scripps Research Institute in La Jolla, CA in 2010.

Dr. Bumpus currently serves as president-elect of the American Society for Pharmacology and Experimental Therapeutics. She previously served as chair of the NIH Xenobiotic and Nutrient Disposition and Action study section.

Her many honors include the Leon I. Goldberg Award from the American Society for Clinical Pharmacology and Therapeutics, the James Gillette Award from the International Society for the Study of Xenobiotics, the John J. Abel Award in Pharmacology from the American Society for Pharmacology and Experimental Therapeutics and the Presidential Early Career Award for Scientists and Engineers, which is the highest honor bestowed by the United States government on early career scientists and engineers. Dr. Bumpus is an elected fellow of the American Association for the Advancement of Science.

Sayash Kapoor (Princeton University) – Special Invited Speaker
Monday, September 26, 2022, 9:30am-10:15am

Sayash Kapoor is a Ph.D. candidate at Princeton University's Center for Information Technology Policy. His research critically investigates Machine Learning methods and their use in science and has been featured in WIRED and Nature among other media outlets. His work has been published in conferences and journals such as ACM FAccT, CSCW, AIES, IJCAI, Machine Learning, and AI Communications. He has received a best paper award by ACM FAccT. At Princeton University, he organized a workshop on "The Reproducibility Crisis in ML-based Science", which saw more than 1,700 registrations. He has worked on Machine Learning in several institutions in the industry and academia, including Facebook, Columbia University, and EPFL Switzerland.

Dr. Matt Turek (DARPA) – Special Invited Speaker

Monday, September 26, 2022, 10:40am-11:20am

Dr. Matt Turek assumed the role of deputy office director for DARPA's Information Innovation Office (I2O) in May 2022. In this position, he provides technical leadership and works with program managers to envision, create, and transition capabilities that ensure enduring information advantage for the United States and its allies.

Turek joined DARPA in July 2018 as an I2O program manager and served as acting deputy director of I2O from June 2021 to October 2021. He previously managed the Media Forensics (MediFor), Semantic Forensics (SemaFor), Machine Common Sense (MCS), and Explainable AI (XAI) programs as well as the Reverse Engineering of Deception (RED) AI Exploration program (AIE). His research interests include computer vision, machine learning, artificial intelligence, and their application to problems with significant societal impact.

Prior to his position at DARPA, Turek was at Kitware, Inc., where he led a team developing computer vision technologies. His research focused on multiple areas, including large scale behavior recognition and modeling; object detection and tracking; activity recognition; normalcy modeling and anomaly detection; and image indexing and retrieval. Turek has made significant contributions to multiple DARPA and Air Force Research Lab (AFRL) efforts and has transitioned large scale systems for operational use. Before joining Kitware, Turek worked for GE Global Research, conducting research in medical imaging and industrial inspection.

Turek holds a Doctor of Philosophy degree in computer science from Rensselaer Polytechnic Institute, a Master of Science in electrical engineering from Marquette University, and a Bachelor of Science in electrical engineering from Clarkson University. His doctoral work focused on combinatorial optimization techniques for computer vision problems. Turek is a co-inventor on several patents and co-author of multiple publications, primarily in computer vision.

Dr. Hugo Aerts (Harvard University, Dana-Farber) – Keynote Speaker

Tuesday, September 27, 2022, 8:50am-9:30am

Hugo Aerts PhD is Director of the Artificial Intelligence in Medicine (AIM) Program at Harvard-MGB. AIM's mission is to accelerate the application of AI algorithms in medical sciences and clinical practice. This academic program centralizes AI expertise stimulating cross-pollination among clinical and technical expertise areas, and provides a common platform to address a wide range of clinical challenges.

Dr. Aerts is a leader in medical AI and Principal Investigator on major NIH-supported efforts, including the Quantitative Imaging Network (U01) and Informatics Technology for Cancer Research (U24) initiatives of the NCI. In 2020 he was awarded a prestigious ERC Consolidator grant of the Horizon program from the European Union. His research has resulted in numerous peer-reviewed publications in top-tier journals. In 2022 he was awarded by Web of Science as he was among the top 1% highest cited scientists worldwide.

Dr. Aerts is an Associate Professor at Harvard University and a Full Professor at Maastricht University. Dr. Aerts earned his Master in Engineering from Eindhoven Institute of Technology, his PhD from Maastricht University, and his postdoctoral fellowship from Harvard School of Public Health.

Dr. Marzyeh Ghassemi (MIT) – Outstanding Reproducibility in Science Award Winner

Tuesday, September 27, 2022, 9:40am-10:10am

Dr. Marzyeh Ghassemi is an Assistant Professor at MIT in Electrical Engineering and Computer Science (EECS) and Institute for Medical Engineering & Science (IMES), and a Vector Institute faculty member holding a Canadian CIFAR AI Chair and Canada Research Chair. She holds MIT affiliations with the Jameel Clinic and CSAIL.

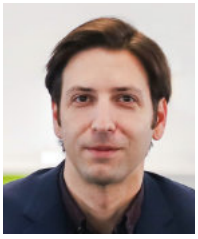
Professor Ghassemi holds a Herman L. F. von Helmholtz Career Development Professorship, and was named a CIFAR Azrieli Global Scholar and one of MIT Tech Review's 35 Innovators Under 35. Previously, she was a Visiting Researcher with Alphabet's Verily. She is currently on leave from the University of Toronto Departments of Computer Science and Medicine. Prior to her PhD in Computer Science at MIT, she received an MSc. degree in biomedical engineering from Oxford University as a Marshall Scholar, and B.S. degrees in computer science and electrical engineering as a Goldwater Scholar at New Mexico State University.

Professor Ghassemi has previously served as a NeurIPS Workshop Co-Chair and General Chair for the ACM Conference on Health, Inference and Learning (CHIL). She also founded the non-profit Association for Health Learning and Inference. Professor Ghassemi has published across computer science and clinical venues, including NeurIPS, KDD, AAAI, MLHC, JAMIA, JMIR, JMLR, AMIA-CRI, Nature Medicine, Nature Translational Psychiatry, and Critical Care. Her work has been featured in popular press such as MIT News, NVIDIA, and The Huffington Post.

Dr. Brian Nosek (University of Virginia) – Outstanding Reproducibility in Science Award Runner-up

Tuesday, September 27, 2022, 10:10am-10:40am

Brian Nosek co-developed the Implicit Association Test, a method that advanced research and public interest in implicit bias. Nosek co-founded three non-profit organizations: Project Implicit to advance research and education about implicit bias (http://projectimplicit.net/), the Society for the Improvement of Psychological Science to improve the research culture in his home discipline (http://improvingpsych.org/), and the Center for Open Science (COS; http://cos.io/) to improve rigor, transparency, integrity, and reproducibility across research disciplines. Nosek is Executive Director of COS and a professor at the University of Virginia. Nosek's research and applied interests are to understand why people and systems produce behaviors that are contrary to intentions and values; to develop, implement, and evaluate solutions to align practices with values; and, to improve research credibility and cultures to accelerate progress.

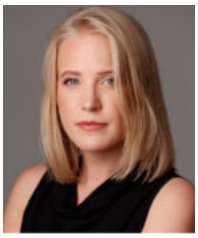


Dr. Zachary Lipton (Carnegie Mellon University) – Keynote Speaker

Tuesday, September 27, 2022, 1:00pm-1:45pm

Zachary Chase Lipton is an Assistant Professor of Machine Learning and Operations Research at Carnegie Mellon University and a Visiting Scientist at Amazon AI. He directs the Approximately Correct Machine Intelligence (ACMI) lab, whose research focuses including the theoretical and engineering foundations of robust and adaptive machine learning algorithms, applications to both prediction and decision-making problems in clinical medicine, natural language processing, and the impact of machine learning systems on society. A key theme in his current work is to take advantage of causal structure underlying the observed data while dealing with the messy high-dimensional data that typifies deep learning settings. He is the founder of the Approximately Correct blog (approximatelycorrect.com) and a co-author of Dive Into Deep Learning, an interactive open-source book drafted entirely through Jupyter notebooks. He can be found on Twitter (@zacharylipton), GitHub (@zackchase), or his lab's website (acmilab.org).

CONTRIBUTING SPEAKER BIOGRAPHIES (IN ALPHABETICAL ORDER BY SURNAME)



Gwenn Berry

Gwenn Berry is the founding engineer and CEO of Magna Labs Inc., whose mission is to accelerate the innovation and delivery of precision science & medicine by empowering scientists to build and test high-quality bioinformatics tools. Gwenn has a decade of experience in applying and adapting software best practices to bioinformatics algorithm development and transitioning research software into both commercial research-use-only (RUO) and in-vitro diagnostic (IVD) on-market products in the genomics industry. Her work specializes in production-grade bioinformatics algorithm and pipeline development, including: methods evaluation and optimization, novel methods development, verification and validation, and remediation of prototype or research-grade software for commercial and clinical readiness.

Dr. Nathan Dwarshuis

Dr. Dwarshuis currently works at the National Institute of Standards and Technology (NIST) where he is part of the team leading the Genome in a Bottle Consortium (GIAB). GIAB develops infrastructure to enable translation of human sequencing research to clinical practice and innovation; his involvement with this has focused on developing machine learning models to understand "difficult" parts of the human genome. Prior to joining NIST Dr. Dwarshuis received his PhD from Georgia Tech/Emory in biomedical engineering, during which his research focus was in CAR T cell manufacturing.

Dr. Tao Chen, Ph.D., D.A.B.T.

Dr. Tao Chen received his Ph.D. degree in Toxicology from the University of Arkansas for Medical Sciences in 1997 and received his diplomat of the American Board of Toxicology in 1999. He was a postdoctoral scientist in Duke University during 1998-2000. He joined the Division of Genetic and Molecular Toxicology, National Center for Toxicological Research, U.S. Food and Drug Administration in 2000 as a research toxicologist. He is also an adjunct professor in several universities. Dr. Chen has served as an editor or an editor board member for more than six scientific journals. He has been a consultant in the World Health Organization (WHO) and the Organization for Economic Co-operation and Development (OECD) for development of regulatory documents and a grant reviewer for the U.S. National Science Foundation and European Research Council. Dr. Chen has served as an organization committee member or a chair for a meeting or a meeting session many times. He has also been invited to present several keynote speeches and planetary lectures in national and international scientific meeting. He has published more than 150 articles in peer-reviewed scientific journals and books. Dr. Chen's current approaches addresses on evaluation of mutagenicity and carcinogenicity of FDA regulated agents using next generation sequencing.

Dr. Cesare Furlanello

Cesare Furlanello is an expert in Machine Learning, predictive models, and reproducibility of AI; he is the CEO of HK3 Lab (http://hk3lab.ai), an AI startup for Predictive Health. He is also Founder and Chief Scientific Officer of Orobix Life. Former Director of Data Science and Head of the Predictive Models for Biomedical & Environmental Data research Unit at Fondazione Bruno Kessler (FBK, Italy). National habilitation: full professor in Biomedical Engineering. Adjunct faculty at Wistar Institute (Philadelphia, USA). Editorial board member of Nature Scientific Data and British Journal of Clinical Pharmacology. AI and Bioinformatics collaborator within the MAQC/SEQC initiative, and of the FANTOM5 consortium. Formerly, President of the international Massive Analysis and Quality Control (MAQC) Society.

Caryn Geady

Caryn is a PhD student in the Department of Medical Biophysics at the University of Toronto, Canada. She has a Bachelor's degree with a focus on Medical Physics and Imaging and a Master's degree with a focus on Medical Image Analysis. Her MSc thesis focused on texture analysis of pathology images towards biological interpretation of radiomics features in pancreatic cancer. For her PhD, she is investigating machine learning techniques for treatment response monitoring and prognostication in sarcoma.

Dr. Mahdi Golkaram

Mahdi Golkaram, PhD. is a Bioinformatics Scientist and a member of DRAGEN team at Illumina Inc. After receiving his PhD from University of California Santa Barbara and interning at Genentech, Mahdi joined Illumina Inc. at 2018. Currently, Mahdi leads the translational oncology bioinformatics research team at Illumina.

The DRAGEN team started at Edico Genome in 2013 and has been blessed with additional talent through the acquisition of Illumina and thereafter. The team is responsible for the secondary analysis tools at Illumina and the on-market DRAGEN product. The team is formed from very talented computational biologists, information theory experts, bioinformatician, amazing software quality engineers, and excellent software engineers. The team has won several awards including from the PrecFDA, and is involved in many worldwide genomics projects.

Other DRAGEN team members: Konrad Scheffler, Theo Heyns, Varun Jain, Rami Mehio, Severine Catreux, James Han.

Dr. Binsheng Gong

Dr. Gong has 20 years of distinguished experience in Bioinformatics research, education, and support, including next-generation sequencing (NGS) and microarray data analysis, development of tools, pipelines, strategies for processing, modeling complex omics data from various sources. He has participated as one of the major investigators in SEQC and SEQC2 projects and published several leading author papers in Nature Biotechnology, Genome Biology, and other prestigious journals. Dr. Gong also brought his expertise and knowledge gained from the SEQC/SEQC2 projects to support multiple projects at NCTR and provided training to NCTR scientists. Dr. Gong won several FDA and NCTR Scientific awards.

Pelin Gundogdu

Pelin Gundogdu obtained her degree in Computer Engineering at TOBB University of Economics and Technology in 2015, and two Master's in Big Data Solutions and Logic, Computation and Artificial Intelligence in 2019 at the University of Barcelona and 2020 at University of Seville, respectively. She is currently a PhD student in the Computer Engineering program at the University of Seville.

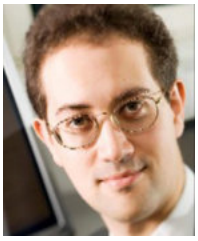
Nowadays, she works as a Machine Learning Researcher in Precision Medicine at the Computational Medicine Platform (Fundación Progreso y Salud) where she researches, develops, and evaluates algorithm in the context of interpretable machine learning model in cancer patient via using signalling pathways, single-cell RNA sequencing data. Her projects have been funded from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813533.

Dr. Donald J. Johann, Jr., MD, MSc, FACP

Dr. Johann is a physician/scientist (medical oncologist), Professor at UAMS and the inaugural Director of the UAMS Genomics Sequencing Facility. His scientific focus concerns advanced molecular profiling and high-throughput technologies for the characterization of molecular alterations in cancer cells and bio-liquids. Previously, he was an assistant investigator at the National Cancer Institute (NCI), Center for Cancer Research (CCR), in Bethesda, MD. Prior to attending medical school he worked as an engineer for the Unisys Corporation for six years, where he directed a team of five engineers on projects involving advanced avionic software design and instrumentation.

Dr. Wendell Jones

Dr. Wendell Jones is Principal Bioinformaticist and Scientific Advisor at Q2 Solutions | Genomics. He conducts collaborative scientific research in multiple molecular biology areas, most recently in immuno-oncology. His background includes bioinformatic systems that process complex genomic assays, including NGS assays, evaluating emerging genomic technologies, and developing bioinformatic strategies. He consults with clients and provides thought leadership in industry and public consortiums. Dr. Jones has over 20 years of experience in advanced genomic technologies and is the Executive Chair and past President of the MAQC Society. He has authored/co-authored of over 70 peer-reviewed publications in genomics and engineering.



Dr. David P Kreil

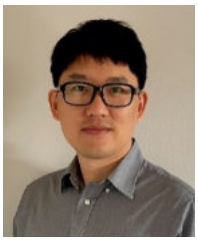
David has followed a call 2005 to Boku University Vienna (www.bioinf.boku.ac.at). In 2019 he has also founded the Institute of Advanced Research in Artificial Intelligence (www.iarai.ac.at). Over the years he has gained extensive expertise in the quantitative analysis of complex data sets from a diverse set of domains. He was educated at the Technical University of Munich (Dept of Physics) and the University of Cambridge where he obtained his PhD as EMBL fellow at the European Bioinformatics Institute. He then received interdisciplinary training as a Research Fellow of Darwin College and the Medical Research Council UK at the University of Cambridge (Dept of Genetics, Cavendish). His research interests include the analysis of life-science data using probabilistic and statistical methods, deep sequence analysis, and the integration of heterogeneous databases and bioinformatics tools.

Dr. Samir Lababidi

Dr. Samir Lababidi is a Mathematical Statistician and a Bioinformatician in the Office of Data, Analytics & Research, Office of the FDA Commissioner (OC), working on further development of the health informatics platforms: precisionFDA, openFDA, and the Global Substance Registration System (GSRS). He participated in several PrecisionFDA Challenges. He also developed the analytical and statistical dashboard for GSRS. Prior to OC, Dr. Lababidi worked in the Office of Biostatistics and Epidemiology, at the FDA's Center for Biologics Evaluation and Research (CBER), where he was a Principal Investigator for the NGS clinical study of genetic risk factors of febrile seizure following measles-containing vaccines and for database development of genetics of autoimmune diseases for vaccine safety signal detection. Prior to CBER, he served as a Statistical Reviewer in the Division of Surveillance and Biometrics, at FDA's Center of Devices and Radiological Health (CDRH), working on diagnostic clinical studies and genomic research studies. Before joining the FDA, he was a research fellow in the Genomics and Bioinformatics Group at the National Cancer Institute/NIH. Dr. Lababidi served on several NIH panels, has a patent, and authored/co-authored over 60 publications in peer-reviewed journals including Nature Biotechnology, Cell, PNAS, and Cancer Research.

Dr. Paweł P Łabaj

Dr Łabaj has graduated at the Silesian University of Technology in 2007 in field of Computer Science in Medicine. He then has joined the lab of Prof. Kreil at Boku University Vienna where he worked on the topic of Measurement and data analysis in the face of noise and complex backgrounds — Advances from improved bioinformatics algorithms. After obtaining his PhD in 2012 he has focused on the Power and limitations of high throughput expression profiling technologies within the MAQC-III/SEQC consortium and later as a Fellow of the Austrian Academy of Sciences on Assessing the impact of alternative transcript discrimination for RNA-Seq gene expression analysis. He has obtained his habilitation (DSc) in Austria in the field of Bioinformatics in 2017 and in Poland in the field of Biomedical Engineering and Cybernetics in 2018. Since 2017 he is head of the Bioinformatics research group in the Małopolska Centre of Biotechnology at the Jagiellonian University in Kraków, Poland.

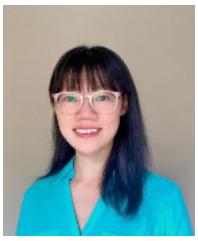


Dr. Dan Li

Dr. Dan Li has a broad background in computer science, software engineering and bioinformatics. He joined the Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR) as a Visiting Scientist in 2018. Since then, Dr. Li has been involved in diverse research projects, including new software and system development, data analysis and data mining, and conducting bioinformatics research.

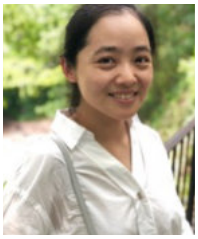
Dr. Li's current research interests include:

- Development of interactive systems to store and manage pharmacology and toxicology review data, captured from documents or entered by reviewers, in database applications
- Designing and developing web-based applications to support and improve regulatory informatics and toxicogenomics
- Working in a sub-group of the FDA-led Sequencing Quality Control Phase 2 (SEQC2) project focusing on onco-panel sequencing, genetic variant detection, genomics, and data analysis
- Conducting studies to evaluate the performance of emerging algorithms for next-generation sequencing data analysis on quality control purpose
- Developing machine learning and bioinformatics models to identify drug and human disease related genetic mutations and discover the underlying mechanisms
- Developing innovative algorithms for next-generation sequencing data analysis, such as single-cell RNA-seq data



Dr. Dongying Li

Dr. Dongying Li is a staff fellow in the Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR). Dr. Li is invested in leveraging big data and integrating computational, in vitro, and in vivo analyses to identify novel, actional biomarker candidates for early and accurate prediction of liver toxicity. Her research also focuses on investigating epigenetic mechanisms that contribute to drug-induced liver injury and inter-individual differences in drug response. Specifically, she is interested in understanding the role of noncoding ribonucleic acids (RNAs), such as microRNAs and long noncoding RNAs, in regulating the expression of drug metabolizing enzymes, which are essential to drug efficacy and toxicity. Her research activities are highly in line with FDA's priority in predictive toxicology and using alternative methods for toxicological studies. She has authored/co-authored over 20 peer-reviewed publications, including reviews, research articles, and book chapters.



Dr. Ting Li

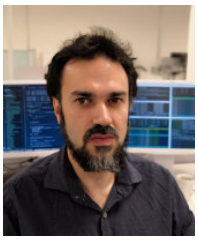
Dr. Ting Li is a research scientist of the Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR). Her primary research interest is to leverage machine learning and artificial intelligence (AI) for drug safety.

Dr. Li works on two FDA mission-critical projects at NCTR:

1. Leveraging the AI methodology from image application to transcriptomic application to construct an AI solution for the system toxicity filed, where the AI model enables to infer the transcriptomic profiles from one organ to other eight organs. The novel model architecture could extend to other studies, such as infer the in vivo transcriptomic results through in vitro transcriptomic profiles.
2. Developing AI models for various toxicological endpoints, such as drug-induced liver injury (DILI), carcinogenicity and mutations in the DNA.

Dr. Jie Liao

Jie Liao is currently a postdoctoral fellow and research assistant at College of Pharmaceutical Sciences, Zhejiang University. He received his bachelor and Ph.D experience in pharmaceutical analysis at Zhejiang University. His main research areas are technologies and data science for spatially resolved transcriptomics, reconstruction of the pseudo-space of single-cell data, and spatial omics-based network pharmacology. As the first or co-first author, his relevant research has been published or accepted in Trends in Biotechnology, Nature Communications, Briefings in Bioinformatics, and so on. He also applied for two patents, and obtained 11 copyrights. He chaired the national science foundation for young scientists of China and the China postdoctoral Science Foundation.



Dr. Carlos Loucera

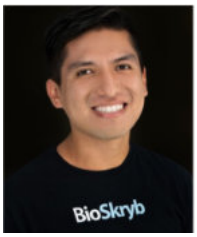
Dr. Carlos Loucera is a Postdoctoral Researcher in the Machine Learning Group of the Bioinformatics Area at the Computational Medicine Platform (FPS-Seville). His current areas of research/work are:

1. Interpretable Machine Learning for Drug Repurposing in Rare Diseases;
2. Causal inference in Real World Data & Evidence;
3. Interpretable Machine Learning in Metagenomics;
4. Data Science Scientific Leader across the Computational Medicine Platform clinical and pharmaceutical collaborations;
5. Interpretable Manifolds for transcriptomic data.

Dr. Loucera has published over 20 peer-reviewed papers and book chapters.

Dr. Oswaldo (Ozzy) Lozoya

Oswaldo (Ozzy) Lozoya, Ph.D. is a NIH Special Volunteer in the Environmental Epigenomics and Disease Group at NIEHS. Dr. Lozoya is a biomedical engineer with expertise on integrative multi-omics technology research and development. His research interests lie at the intersection of metabolism, mitochondrial biology, biophysical stimuli, and epigenetics on stem cell physiology, with emphasis on the effects of environmental exposures and infectious pathogens. Dr. Lozoya's mission is to accelerate extraction of biological insight from "Big Data" by deploying new analytical paradigms in collaboration with government, industry, and academic stakeholders to the benefit of public health.



Dr. Isai Salas-González

Dr. Salas-González is a Senior Computational Biologist at BioSkryb Genomics where he has developed the computational tools to analyze data from ResolveOME™, a flagship workflow that permits to amplify the genome and transcriptome from the same cell simultaneously with unprecedented fidelity. In addition, he has contributed to the deployment of BaseJumper™, a state-of-the-art analytical platform to analyze and intuitively visualize whole genome sequencing and full-length transcriptome data from single cells

He has authored over 20 research articles, his doctorate research deconvoluting microbial heterogeneity on the light of plant-microbiome interactions was published in Science, Nature and Nature Genetics.

Dr. Susanna-Assunta Sansone

Prof. Susanna-Assunta Sansone is a Full Professor in Data Readiness at the University of Oxford, an Associate Director and Principal Investigator at the Oxford e-Research Centre, and the newly appointed University-wide Academic Lead for Research Practice. She is also a Consultant for Springer Nature, and Founding Honorary Academic Editor of their Scientific Data journal; also an author of the FAIR Principles. Susanna leads on the Data Readiness Group, in the Oxford's Department of Engineering Science, and focuses on data sharing, reproducibility and the evolution of scholarly publishing. In the Life Science she is the driver behind many flagship resources in the European research data infrastructure, such as FAIRsharing and the FAIR Cookbook, co-developed with pharmas.

Dr. Xin Shao

Xin Shao, Ph.D., research associate of Zhejiang University, mainly engaged in the research of single-cell omics, cell-cell communication, and liver transplantation. As the first author or co-author, publish 15 SCI papers in several journals such as Nature Communications, Nucleic Acids Research, Protein & Cell, iScience, Briefings in Bioinformatics. Member of the Intelligent Pharmacy Society of Zhejiang Bioinformatics Society and the Network Pharmacology Professional Committee of the World Federation of Chinese Medicine Societies. Preside over the National Natural Science Foundation of China Youth Program and China Postdoctoral Natural Science Foundation.

Dr. Weida Tong

Dr. Weida Tong is Director of the Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR). He has been:

1. Developing machine learning and AI for digital health and drug repositioning;
2. Supervising and leading an FDA-led community wide consortium to analyze technical performance and utility of emerging genomics technologies with an emphasis on regulatory science and precision medicine;
3. Developing Liver Toxicity Knowledge Base to address drug safety concerns related to drug-induced liver injury;
4. Designing and developing computer-based technology to support FDA's effort on bioinformatics and scientific computing; and
5. Conducting molecular modeling and structure-activity relationships on various toxicological endpoints, such as endocrine disruption and carcinogenicity.

Dr. Tong has published over 300 peer-reviewed papers and book chapters.

Dr. Leihong Wu

Dr. Leihong Wu is a Staff fellow in the Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR). Dr. Wu's research interest is to apply bioinformatics — particularly, Artificial Intelligence (AI) and Machine Learning (ML) — to biomedical research and informatics. Specifically, Dr. Wu's work has focused on the development of algorithms for biological and pharmaceutical research tasks such as drug safety, QSAR modeling, and genomics.

Dr.Wu's current research interests include:

- Developing innovative AI algorithms for big data analysis, including multi-platform biological data such as gene expression, sequencing, and bioassays
- Designing and developing AI/machine learning framework to facilitate regulatory science
- Developing advanced predictive models using deep learning for biomarker identification of DILI and predictive toxicology research
- Developing innovative machine learning algorithms for text-mining in massive FDA regulatory documents
- Developing convolutional neural network architectures for advanced imaging analysis in drug and food safety assessment
- Designing and developing databases and visualization tools that promoting AI for regulatory use, including FDALabel
-

Dr. Joshua Xu

Dr. Xu is the Branch Chief for Research-to-Review (R2R) at the Division of Bioinformatics and Biostatistics of FDA's National Center for Toxicological Research (NCTR). He specializes in genomics, big data, image analysis, and machine learning. His recent endeavor has been with the FDA-led Sequencing Quality Control Phase 2 (SEQC2) project to evaluate the technical reliabilities and scientific applications of the next generation sequencing (NGS) technologies. He leads the Oncopanel Sequencing Working Group to assess the reproducibility and detection sensitivity of onco-panel sequencing including liquid biopsy. He is also the executive secretary for MAQC Society.

The MAQC Society would like to thank the U.S. Food and Drug Administration's National Center for Toxicological Research (NCTR) for sponsoring this conference. We thank the precisonFDA team for helping organize the conference and the FDA conference grant for support. Also, we would like to thank the FDA for the use of their facility and also for the IT/AV personnel for administering the technical aspects of the conference.