

AR-BIC 2023: Bioinformatics, Big Data, AI (Artificial Intelligence), and Public Health: An Integrated World

The Arkansas Bioinformatics Consortium (AR-BIC) recently held its 9th Annual Conference in Little Rock, Arkansas, on March 13 and 14, 2023. The theme for this year's conference was "Bioinformatics, Big Data, AI (Artificial Intelligence), and Public Health: An Integrated World." Under the leadership of Dr. Weida Tong, Director of NCTR's Division of Bioinformatics and Biostatistics (DBB), the Scientific Program Committee organized a one-and-a-half-day program that included a wide range of presentations by speakers from diverse fields. The conference featured 2 keynote speakers, 6 breakout sessions, 2 AI workshops, over 30 platform presentations, and more than 40 poster presentations. The in-person attendance for this year's conference was the largest in AR-BIC's nine-year history at over 180 participants. A special issue of the journal *Experimental Biology and Medicine* has been planned for research performed by this year's AR-BIC participants. By the end of the conference, over 10 papers had been confirmed for submission to the special issue, which is intended to become an annual tradition. This conference successfully provided a platform to share and discuss current state-of-the-art practices and ongoing efforts to apply AI in healthcare and to enable efficient data mining to promote public health.

In the era of big data, bioinformatics and artificial intelligence (AI) are effective tools that can be utilized to improve public Health. The integration of bioinformatics, big data, AI, and public health has the potential to transform healthcare by providing personalized medicine, improving disease surveillance, and enabling early detection and intervention. With the vast amount of health-related data generated, bioinformatics and big data techniques can be used to identify disease patterns and risk factors, predict outbreaks, and inform public health policies. AI is actively being utilized to develop predictive models that may help clinicians make more accurate diagnoses, develop treatment plans, and monitor patient progress. Thus, AR-BIC 2023 aimed to provide most recent updates on the application of bioinformatics, datamining, and AI to address a variety of pressing public health issues.

The AR-BIC 2023 program was planned to cater a board range of attendees with various expertise and at different career stages. It provided the participants with access to a wealth of knowledge and resources that were relevant to their specific needs and interests in implementing

AI- and bioinformatics-based applications in their current fields. The discussions and presentations were designed to be accessible and engaging, regardless of the attendees' background or level of experience. Furthermore, the in-person setting of AR-BIC 2023 allowed the attendees to easily interact with other researchers, developers, and healthcare professionals, enabling them to form valuable partnerships and exchange ideas. This facilitated cross-disciplinary collaboration, which is crucial for the development of innovative solutions that can address complex healthcare challenges.

The first day of the meeting started with two informative workshops. Dr. Xiaowei Xu, Professor of Department of Information Science, University of Arkansas at Little Rock, spoke about AI for Natural Language Processing (NLP) and provided an overview of language models, causal inference, and the basic concept and application of AI for NLPs. In the other workshop, Dr. Joe Zhang, Professor of School of Computing Sciences and Computer Engineering, University of Southern Mississippi, introduced deep learning based analysis of histopathological images of breast cancer to address some of the important questions regarding image data processing, model selection and development, and performance evaluation and improvement.

FDA's Chief Scientist, Dr. Namandjé Bumpus, provided opening remarks for the conference by highlighting the current research goals of pursuing new alternative methodologies and helping the public to better understand science. Following the opening remarks were the first keynote presentation titled "Algorithmic Medicine: New Opportunities to Increase Patient Trust," in which Drs. Joseph Sanford and Kevin W. Sexton of University of Arkansas for Medical Sciences (UAMS) spoke about the implementation of machine learning (ML) and AI technologies into an enterprise medical record system focusing on the challenges, opportunities, and common missteps. They provided an introduction to clinical informatics, which is considered the newest medical specialty, and discussed its future role in the continued development and implementation of these cutting-edge technologies.

Dr. Ruth Roberts, Chair and Director of Drug Discovery at Birmingham University, UK, Cofounder of ApconiX, gave another exciting keynote presentation. Dr. Roberts talked about the importance of early derisking of drug targets in drug development. Considering the time-consuming and costly process with high attrition rates in drug development, regulatory guidelines are essential to ensure drug safety and efficacy. Despite these guidelines, Dr. Roberts pointed out

that toxicity issues still arise and cause drugs to halt in discovery and development. However, innovative approaches can improve success rates and reduce animal testing. Dr. Roberts discussed Target safety assessments (TSAs) and a data science-based approach has been developed to gather relevant evidence for an optimal TSA. Further, Dr. Roberts presented a deep generative adversarial network (GAN)-based framework that has been created to generate new animal results from existing animal studies without additional experiments. Dr. Roberts emphasized that by taking advantage of the latest developments in data science, these approaches can identify and mitigate risks, aid decision making, and optimize resource management in drug development.

Mr. Bryan Barnhouse, President & CEO of Arkansas Research Alliance (ARA), led a breakout session that provided industry views and tools for improving quality of care with health care data analytics. The session included speakers from three Arkansas health organization who talked about how they utilized data analytics to improve revenue, reduce costs, increase operational efficiencies, and enhance overall performance. The presentations covered the challenges, opportunities, tools, and applications to convert large data into innovative solutions for improved patient and client outcomes.

Dr. Douglas Rhoads of University of Arkansas and Dr. Steven Foley of National Center for Toxicological Research, US FDA, hosted a session on bioinformatics and AI in pathogen surveillance and microbial genomics, which explored different aspects of bioinformatics and the possible application of AI for surveillance, prevention, and treatment for microbial pathogens. In this session, Dr. Arvind Ramanathan of Argonne National Labs gave a feature presentation on using foundation models to track, monitor and predict SARS-CoV-2 variants of concern. Specifically, Dr. Ramanathan discussed the development of genome-scale language models (GenSLMs) to identify and classify new and emergent variants of SARS-CoV-2. By pre-training on 110 million prokaryotic gene sequences and fine-tuning on 1.5 million SARS-CoV-2 genomes, the GenSLMs was reported to accurately and rapidly identify variants of concern and can generalize to other prediction tasks. In this pioneering work, Dr. Ramanathan demonstrated the scalability of GenSLMs on GPU-based supercomputers and AI-hardware accelerators and revealed novel insights from tracking the evolutionary dynamics of SARS-CoV-2 using GenSLMs.

A session chaired by Drs. David Ussery and Jonathan Bona of UAMS focused on AI in healthcare and included 6 presentations given by UAMS professors and graduate students. The presentations showed how AI is utilized to help in research in cardiovascular medicine (Dr. Subhi J. Al'Aref), cancer imaging (Dr. Fred Prior), and drug design (Dr. Grover Miller). Also, the speakers discussed their application of ML and AI methodologies to mining social media data for adverse drug reactions (Catherine Shoults), public health (Brian Delavan), and for 'crowdsourcing AI solutions' to challenges in healthcare (Jennifer Fowler).

Dr. Joshua Xu of National Center for Toxicological Research, US FDA, led a session on integrated genomics for precision oncology and covered five presentations on topics ranging from single cell sequencing and best practices in bioinformatics to real-world data from clinical studies. The feature speaker of the session was Dr. Leming Shi of Fudan University, Shanghai, China, who spoke about quality control and standardization of multiomics for precision medicine. Dr. Shi presented his work in developing publicly available multiomics reference materials for DNA, RNA, proteins, and metabolites to improve the objective assessment of multiomics profiling. Dr. Shi showed that the "ratio"-based omics profiling data were capable to empower the horizontal and vertical data integration in multiomics studies by improving reproducibility across batches, labs, platforms, and omics types.

Dr. Samantha Robinson of University of Arkansas, Fayetteville, hosted a session on drivers of public health in the US population, exploring why valid self-report measures matter. Dr. Todd Shields, Arkansas State University kicked off the session with his presentation on a validation study of fatalism scales. Dr. Shields pointed out the lack of evaluations of the psychometric properties of the most widely used scales and the limitations of existing investigations. Dr. Shields showed their results from evaluating the psychometric properties of four of the most commonly used measures of Fatalism, which demonstrated the most appropriate scales and items for use in the Fatalism research and its influence on behavior, marketing, and public health. The session highlighted a multi-institutional, multi-disciplinary work and the collaborative effort the collaborative power of Arkansas researchers.

Drs. Huixiao Hong & Shraddha Thakkar of US FDA held a session discussing ML and deep learning (DL) for big data analysis and drug development. The session consisted of a series of presentations from both well-established scientists and graduate students. The speakers reported

the most recent advances in applications of ML/DL in different areas including medical image processing, drug development, and proteoform detection. For example, Dr. Jie Liu of US FDA presented her work on developing a predictive model using a random forest algorithm to identify potential drug candidates for repurposing to treat COVID-19. The presentation showcased the potential of ML in accelerating drug development for managing COVID-19.

As a tradition, AR-BIC 2023 hosted a poster session which consisted of 40 poster presentations, including 30 from postdocs, graduate, undergraduate and high school students that participated in the poster award competition. Twenty-four poster judges from various institutes and organization served on the Poster Committee to evaluate the posters for overall presentation, scientific excellence and presenter's contribution. Drs. Huyen Le, Shivangi Shrimali, and Chandra Mohan Reddy Muthumula of National Center for Toxicological Research, US FDA, have won the first, second, and third place of Postdoctoral Fellow Poster Presentation Award. Michael Rutherford of UAMS, Aishat Lawal of University of Arkansas, Fayetteville, and Anu Iyer Little Rock Central High School have won the the first, second, and third place of Student Poster Presentation Award.

The AR-BIC 2023 meeting was centred around the theme of using AI, big data, and bioinformatics to enhance public health. Experts and researchers from various disciplines came together to share their latest findings and insights, with a particular emphasis on how these advanced technologies can be leveraged to improve disease prevention, diagnosis, and treatment. The meeting served as a platform for attendees to network and collaborate, highlighting the important role of interdisciplinary collaboration in addressing complex health challenges.