U.S. FOOD & DRUG ADMINISTRATION

National Center for Toxicological Research

Bioinformatics and Biostatistics

NCTR's Division of Bioinformatics and Biostatistics (DBB) develops integrated bioinformatics and biostatistics capabilities to address increasing need in areas such as biomarker development, drug safety, drug repositioning, precision medicine, artificial intelligence (AI), rare diseases, endocrine disruptors, and risk assessment.

About the Division

Bioinformatics

Constructs knowledge bases to provide a data-driven decision-making environment for enhanced safety evaluation and precision medicine.

Biostatistics

Conducts research of statistical methods to analyze toxicological and molecular data as well as data-mining techniques for pattern identification and signal detection.

Scientific Computing

Provides IT support to the entire NCTR.

Review-to-Research and Return (R2R)

Translates division research for regulatory application.

By the Numbers

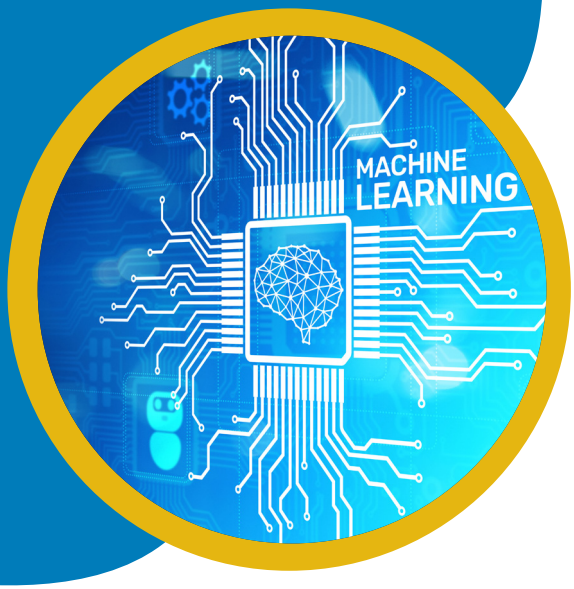
DBB staff were engaged in the following activities in 2023:

38	papers published
35	platform presentations delivered
26	posters presented
26	scientific conferences attended
7	concept papers submitted
6	projects completed/closed

Contact Us
870-543-7538

www.fda.gov/nctrbioinformatics
NCTRBioinformaticsSupport@fda.hhs.gov

3900 NCTR Rd
Jefferson, AR 72079



Outreach

In 2023, DBB scientists contributed to or participated in the following outreach events:

- 9th AR-BIC Annual Conference — Scientific Program Committee (chair)
- 13th Annual Global Summit on Regulatory Science (chair)
- Cheminformatics Resources of U.S. Governmental Organizations — FAIR-ifying and Sharing Chemical-Related Data session (chair); "Machine Learning Models for Rat Multigeneration Reproductive Toxicity Prediction" (presentation)
- FDA 11th Annual Scientific Computing Days (three posters presented and one breakout session hosted)
- Institute of Electrical and Electronical Engineers-USA Innovation, Workforce, and Research Conference (three DBB staff represented NCTR)
- Large Language Model (LLM) workshop hosted by Swissmedic and instrumental in establishing the interagency LLMs Workforce (DBB staff led this first ever workshop)
- "Making Sense of Electronic Health Record (EHR) Race and Ethnicity Data Challenge" (The Biostat Team presented at precisionFDA)
- Society of Toxicology 62nd Annual Meeting/ToxExpo (DBB staff played a pivotal role in preparations and event and presented 18 scientific posters)

2023 Select DBB Accomplishments

- Made significant progress in several focused areas, including the opioid crisis, cannabidiol (CBD), endocrine disruptors, drug safety, COVID-19, toxicogenomics, and precision medicine, and specifically, in Artificial Intelligence (AI)/Machine Learning (ML).
- Developed **DICTrank**, (Drug-Induced Cardiotoxicity Rank), the largest reference list of 1,318 human drugs ranked by risk of drug-induced cardiotoxicity using FDA labeling, which facilitates the development of NAMs (*Drug Discovery Today*).
- Developed **AnimalGAN**, a Generative Adversarial Networks (GAN) model that simulates animal studies to reduce animal use in preclinical studies for drug development (*Nature Communications*).
- Led five projects that were newly funded by FDA intramural grants on the following topics: **AI bias** funded by Office of the Chief Scientist, **drug safety in women's health** by Office of Women's Health, **COVID-19 and its variants** by the Medical Countermeasures Initiative, and **real-world data for minority health** by Office of Minority Health and Health Equity.
- Published approximately 38 manuscripts, some in well-respected journals such as *Nature Biotechnology*, *Nature Communications*, and *Genome Biology*.

Intercenter Support

In 2023, NCTR scientists developed RxBERT, an AI model optimized to better understand human prescription drug labeling. They also supported Center for Drug Evaluation and Research (CDER) regulatory missions by:

- Improving the accuracy and timeliness of the Novel Drug Approvals Dashboard by creating software to identify new and updated information in the Data Analysis Search Host Database of marketing application information derived from CDER-generated regulatory documentation.
- Completing population of a searchable Executive Carcinogenicity Assessment Committee (ECAC) database within the Smart Template System that allows pharmacology/toxicology reviewers to search ECAC meeting minutes from 1995 to present for information relevant to their current reviews.