Solutions Overview

BIORANKINGS

EXPERT STATISTICIANS FOR DATA ANALYTICS CHALLENGES

Primary Contact Information

Bill Shannon, PhD, MBA
Cofounder
bill@biorankings.com
(314) 704-8725

Secondary Contact Information

Ally McClure
Cofounder
ally@biorankings.com
(314) 633-1821

Executive Summary

About Us

BioRankings provides virtual collaborative data analytics services and customized software to clients in the medical and agriculture life sciences space. The company spun out of Washington University School of Medicine in 2014 by Bill Shannon and Ally McClure and now has seven full-time staff members with skills in biostatistics, software engineering, database development, writing and communications, and project management. The BioRankings team has experience working in many areas of biology and research with clients representing small biotech companies, start-ups, large pharmaceutical, plant, and animal (food and companion) health science companies, and academic research institutions.

Areas of Specialization

- Classical biostatistics
- Multi-omics
- Databases
- Cloud-computing
- AI/ML

Key Differentiators

- Deep domain expertise in statistical approaches for omics data
- Focus on results interpretation for all stakeholders
- Works collaboratively with R&D teams
- Provides customized analyses, not canned reports
- Commitment to knowledge transfer of our expertise to clients
- Focus on analytics pipelines and software to streamline analyses
- Experience in creating custom databases for client needs
- Partnerships with cloud computing companies (AWS and Microsoft)

Introduction to BioRankings

Analytics and Consulting

Research

Software and Pipelines

Knowledge Transfer

The BioRankings team brings decades of experience working with data in diverse subject areas in medicine and agriculture. We've helped Fortune 50 companies, startups, and research institutions navigate their data and get accurate, actionable analytics results.

Spun out of Washington University School of Medicine in St. Louis in 2014, the BioRankings team of statisticians and software engineers collectively has contributed to over 200 publications and has received significant National Institutes of Health (NIH) funding over the last 30 years, including over $4 million in NIH Small Business Innovation Research (SBIR) grants to BioRankings to develop and commercialize innovative statistical approaches for omics data.

Our four pillars are the foundation of BioRankings:

Analytics and Consulting

Expert statisticians focus on understanding client data and use appropriate statistical methods for interpreting results to inform decision making.

Research

Our researchers develop analytics approaches for new, challenging data.

Software and Pipelines

Software and code can be custom developed and fit into company analytics pipelines.

Knowledge Transfer

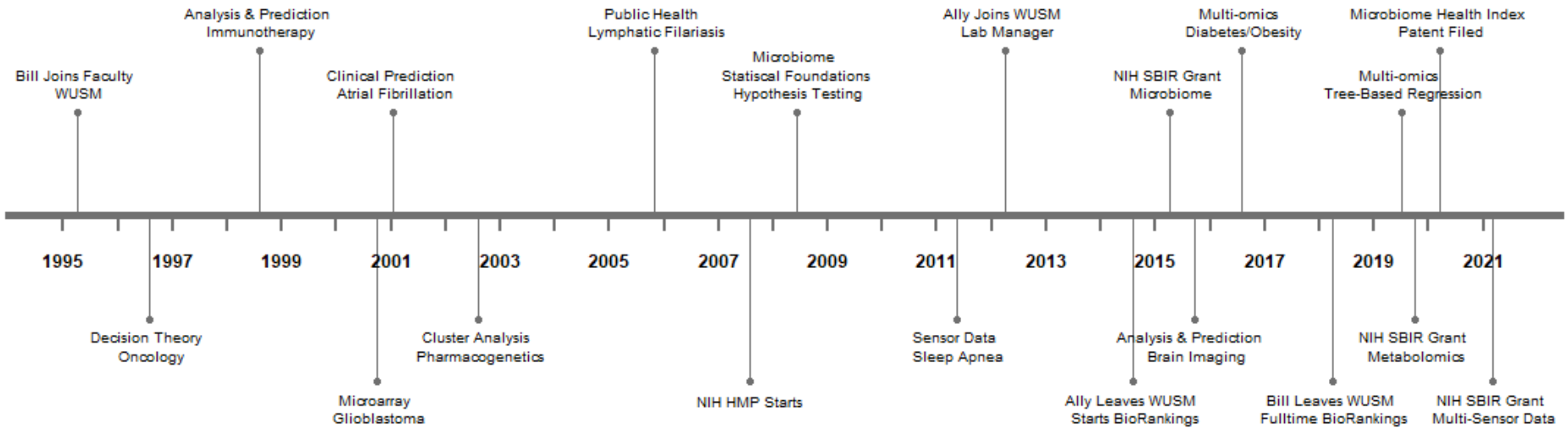
Our team works collaboratively with clients, ensuring all stakeholders understand our approach and results.

BioRankings acts similar to an academic biostatistics core in providing classical statistical planning and analyses to corporate researchers, including:

- Experimental planning and design
- Sample size / power calculations
- Survival analysis
- Hypothesis testing
- Data management
- Statistical analyses
- Manuscript, report, and presentation support

We approach projects as collaborators and not consultants. In the Background Section we describe some of our methods development for microbiome, metabolome, and sensor data.

BioRankings Key Dates and Contributions



Background

Scientific Background of Co-founder

With an MS in Zoology and a PhD in Biostatistics, Dr. Shannon spent 20 years on the faculty at Washington University School of Medicine (WUSM) rising from Assistant to Full Professor of Biostatistics in Medicine with Tenure. After leaving the university to work full-time at BioRankings he was named a Professor Emeritus. Dr. Shannon's academic credentials, and some publications and grant information by him and BioRankings staff are presented in the Documents and Publications Table on page 7.

In academia, Dr. Shannon was Director of the Department of Medicine's Biostatistical Consulting Center supervising the work of anywhere from 5-10 biostatisticians, software engineers, and post-doctoral students. The center provided biostatistical support to faculty from all departments. As an NIH funded investigator, his team also developed new methods for analyzing complex and high-dimensional data in microarray, microbiome, untargeted metabolomics, brain imaging, and other areas.

As an academic biostatistics lab his team was highly collaborative working closely with subject matter experts from all areas of biomedical R&D. Several of these areas are highlighted in the Timeline. By learning the biology being studied the statistical analyses proposed by the team could be optimized, and by teaching details about the statistics to the researchers, better interpretation of the results were obtained.

In 2012 Dr. Shannon earned his MBA. In 2015 he and his business partner, Alexandra McClure who was managing the WUSM consulting center, developed a business model and started BioRankings to provide the same collaborative consulting and statistical R&D services to corporate partners that were provided to academic investigators.

Work in Microbiome

Omics data is hard to analyze statistically because of the large number of variables measured on a few samples. Known as the Large-P-Small-N, or Curse of Dimensionality problem, this is known to increase false positives. The results often can't be replicated and the costs of doing these additional experiments on false positives can be huge.

To avoid this, we work with collaborators to identify subsets of the variables based on biological knowledge, and analyze these subsets independently to generate testable hypotheses. In one study, a subset of 13 cytokines thought to be associated with diabetes was selected and analyzed with microbiome taxa counts. We developed a multi-omics statistical regression model and showed a clear association between the cytokines and microbiome populations. By focusing on subsets, we avoid the Large-P-Small-N problem and leverage classical statistical methods. This uses biological thinking and hypotheses instead of fitting a model to 1,000s of variables in the hope of finding patterns.

As the first biostatisticians working on the NIH Human Microbiome Project, we were involved with establishing the sequencing protocols, bioinformatics pipelines, and data analyses. Some of our early work laid the statistical foundations for hypothesis testing and sample size/power calculation for microbiome data, as well as the analysis of data for the HMP, necrotizing enterocolitis, and other clinically focused microbiome studies.

Recently, in collaboration with Rebiotix (a Ferring Pharmaceutical Company), we developed and jointly patented the Microbiome Health Index, MHI for Clostridium difficile infection health status. A poster on this index is found in the Documents and Publications Table, and new results are under review for publication. Rebiotix

Background

has been a long-time client and collaborator of BioRankings, where we have served as their Virtual Biostatistics Core from startup through acquisition and FDA trials.

We have also collaborated with clients in agriculture on soil microbiome and animal growth.

Work in Metabolome

BioRankings has been analyzing and developing new methods for untargeted metabolomics data. To date we have received ~$2M in National Institutes of Health Small Business Innovation Research (NIH SBIR) grants to create a cloud-based automatic peak detection software platform. Developed in collaboration with research partners at some of the top universities, the approach uses a statistical model we developed for the unbiased detection of metabolite peaks. We are now focused on solving the 'major road-block' of annotating the discovered metabolites using mathematical consensus theory. The cloud-based software is being tested with data from academic metab-olomics cores and is being prepared for licensing.

Work in Sensor Data (Internet-of-Medical-Things)

The use of multi-sensor data in medicine and agriculture is increasing, and the need for better analytics is high. In 2011 we published work using functional data analysis for analyzing actigraphy (movement) data in patients with sleep apnea, and showed different patterns between severe and mild cases.

In the last month we received an NIH SBIR grant to further develop these methods that we expect to seamlessly extend to any streaming sensor data.

Work in Integrative Multi-omics

Seeing the need for incorporating multiple omics into a single analysis, BioRankings has developed new statistical methods and modified existing methods designed for non-omics data to work with omics data. In 2019 we published a new tree-based model to regress microbiome taxa count data onto cytokines to study the association between these datasets in diabetes. Another approach we developed was to modify the RLQ method from ecology to measure the impact of specific taxa on pathway metabolites. This led to potential biomarkers and potential treatments (i.e., increase specific taxa to drive the pathway) for diabetes and obesity.

Documents and Publications

Topic	Type	Title	Subject	Link
Background	CV	NIH Biosketch		Link
High Dimensional Data	Blog	Curse of Dimensionality	Mathematical problems this type of data produces	Domino Data Lab
	Blog	Analyzing Large-P-Small-N Data with Examples from Microbiome and Metabolomics	Approaches for analyzing this type of data	Domino Data Lab
Omics (Microbiome)	Manuscript	New statistical method identifies cytokines that distinguish stool microbiomes (Senior Author)	Regression of microbiome taxa count data onto other covariates	Nature
	Manuscript	Hypothesis testing and power calculations for taxonomic-based human microbiome data (Senior Author)	Statistical foundation for microbiome taxa count data	Pubmed
	Manuscript	Biogeography of the ecosystems of the healthy human body	Analysis of HMP data and comparison across body sites	Pubmed
	Manuscript	Gut bacteria dysbiosis and necrotising enterocolitis in very low birth-weight infants: a prospective case-control study (Co-Senior Author)	Analysis of microbiome in a clinical setting	Pubmed
	Manuscript	Patterned progression of bacterial populations in the premature infant gut (Co-Senior Author)	Analysis of microbiome in a clinical setting	PNAS
	Poster	Evaluating a Prototype MICROBIOME HEALTH INDEX™ (MHI™) as a Measure of Microbiome Restoration Using Data Derived from Published Studies of Fecal Microbiota Transplant to Treat recurrent Clostridium difficile Infections (rCDI)	Description of a microbiome index for quantifying health	Link
Omics (Metabolome)	Project Summary	Object Oriented Data Analysis for Untargeted Metabolomics	SBIR funded project for automating peak detection	Link
	Project Summary	Consensus Theory for Annotating Untargeted Metabolomics	SBIR project under review for annotating metabolites	Link
Multi-Sensor Data (IoT/IoMT)	Manuscript	Measuring the impact of apnea and obesity on circadian activity patterns using functional linear modeling of actigraphy data	Use of functional data analysis for analyzing actigraphy data	NCBI
	Project Summary	Analyzing Streaming Multi-Sensor Data to Predict Stroke in Preterm Infants	SBIR project under review for IoT / IoMT methods	Link