JUNRUI DI

510-809-5425 ⋄ dijunrui@gmail.com ⋄ https://junruidi.github.io/

PROFESSIONAL PROFILE

Initiative-taking and research-oriented statistician/data scientist dedicating to transform disease monitoring and assessments with digital health technologies. Six years of research experience in modeling data generated by wearable devices and mobile sensors (e.g. accelerometers and actigraphy), and two years of pharmaceutical industry experience in deploying digital health technologies into clinical studies across multiple therapeutic areas. Proven ability to work in a matrix environment with cross-functional teams which can be demonstrated by multiple Pfizer value recognition awards. Solid technical background with track records of statistical and scientific publications, patent application, and peer-review activities.

TECHNICAL AND SCIENTIFIC INTERESTS

· Feature engineering of accelerometry signals

· Functional data analysis

· Fusion and integration of multiple data modalities (e.g. multiple devices)

· Machine learning (e.g. activity types prediction and classification)

· R programming, package and Shiny app development

· Study design with wearable devices for disease monitoring and assessment

· Development and validation of digital endpoints and digital biomarkers from wearable devices

EDUCATION

Johns Hopkins Bloomberg School of Public Health *Mar 2019*
Ph.D. in Biostatistics

Thesis: *Feature Engineering and Statistical Modeling of Wearable Devices Data*

Georgetown University *Dec 2013*
M.S. in Biostatistics

Thesis: *Robust Integrative Analysis of Multi-Block Contaminated Datasets*

University of California, Berkeley *May 2012*
B.A. in Applied Mathematics

High Distinction General Scholarship (roughly equivalent to Magna Cum Laude), Phi Beta Kappa

EXPERIENCE

Pfizer *Cambridge, MA*
Associate Director, Digital Medicine Statistics (Clinical and Nonclinical) *Dec 2021 - Present*
Manager, Digital Medicine Statistics (Clinical and Nonclinical) *Apr 2019 - Nov 2021*

· Design and execute multiple Pfizer internal and external collaborative methodological studies to develop and validate novel digital and wearable biomarkers while ensuring sound statistical thinking and methods are utilized

- Developed protocol and SAP for three methodological studies to collect training and independent test datasets, and built an accelerometry-based nocturnal scratch detection algorithm for patients with Atopic Dermatitis together with data scientists and clinicians, which has led to 1 algorithm publication, 1 provisional patent application, 10 clinical presentations and publications (ready for submission), 2 briefing documents submitted to FDA/EMA.
    - Protocol and SAP developments for Pfizer Innovation Research (PfiRE) Lab studies to detect and monitor gait and physical activity using accelerometers for adult and pediatric populations in both in-lab and at-home environments.
    - Protocol and SAP developments for a natural history study to detect differences of accelerometry-measured physical activity between patients with heart failure with preserved ejection fraction (HFpEF) and matched controls.
    - Built signal processing and machine learning based models to classify sleep stages with both accelerometry-measured motions and electrocardiogram (ECG) measured heart rate.
- Provide subject matter expertise to asset studies across multiple therapeutic areas by offering guidance on devices/endpoints selection, drafting protocols and SAPs for the device portion, bringing analytic expertise for digital endpoints, and implementing industry-leading mobile/digital technologies into clinical trials.

    - Supported 6 Phase 3 studies and 3 Phase 2B studies for various dermatological conditions including Atopic Dermatitis, Chronic Hands Eczema, and Psoriasis to implement accelerometry detected sleep and nocturnal scratch.
    - Supported 1 Phase 3 study on dilated cardiomyopathy and 1 Phase 2A study on heart failure with reduced ejection fraction (HFrEF) to quantify physical function using actigraphy measured physical activity.
- Actively engage with external experts in the digital health community to drive its broad acceptance by leading a publication on handling missing data on digital health technologies and co-authoring a publication on decentralized trials with multiple DIA working groups.
- Supervise internal and external statistical programmers to ensure sound statistical methods are utilized to generate TLFs for both methodological and asset studies and that they are in compliance with Pfizer standards.
- Co-found the Pfizer Artificial Intelligence Network (PAIN) to enable Pfizer AI practitioners to identify specific skill sets, share best practices, discuss project work, and increase resource reuse, which has resulted in a core group of 30+ people from multiple AI-related functional teams.
- Co-found and co-organize the biweekly Pfizer Artificial Intelligence Network Talk Series (PAINTS) with topics ranging from technical aspects such as development and application of AI methods applied to Pfizer internal problems (R&D, commercial, manufacture, etc), and nontechnical talks such as legal considerations for patenting AI products, which has 80+ audiences on average for each talk and inspired further in-depth discussion and collaboration across multiple groups.

Johns Hopkins Bloomberg School of Public Health *Baltimore, MD*
Research Assistant *Jun 2015 - Apr 2019*
- Developed statistical methods to extract features from noisy accelerometry signal, to reduce dimensionality of multivariate accelerometry features, and to integratively analyze features representing multiple physiological domains or extracting from multiple devices/sensors.
- Conducted data processing and statistical analyses in collaboration with researchers for studies with measurements of sleep, activity, and circadian rhythms using wearable devices.
- Constructed data processing and analytic pipeline that converts epoch level activity counts data acquired from actigraphy to summary level digital endpoints for statistical modeling which is currently used by multiple study teams.

Takeda Pharmaceuticals *Cambridge, MA*
Biostatistician Intern *Jun 2018 - Aug 2018*

· Provided theoretical and analytic evidence that digital endpoints measured by wearable devices are more sensitive than traditional measurement.
· Developed an R Shiny App to analyze/visualize clinical data from a chest-worn device used to measure physical activity and cardiovascular behavior to assist clinicians to identify false positive events reported.

Georgetown University   *Washington, DC*
Research Assistant, Department of Medicine   *May 2013 - Apr 2014*

· Supported the Multicenter AIDS Cohort Study (MACS) site PI by providing data analysis, visualization, and tabulation on data generated from the 30-year nationwide longitudinal cohort study involving four sites, and over two hundred tested variables.
· Quantified satisficing in MACS online survey based on the group-based trajectory analysis of average response time across visits.

Research Assistant, Department of Biostatistics   *Sep 2012 - May 2013*

· Refined "Joint and Individual Variation Explained" (i.e. JIVE), a recently developed integrative analysis method of multi-block datasets of different types, to be robust to outliers using low-rank approximation based on robust singular value decomposition.
· Selected metabolomics biomarkers to develop clinical assays for diagnosis and prognosis of Pancreatic Ductal Adenocarcinoma (PDAC) by applying various variable selection techniques.

PUBLICATIONS

Published / In Press

1. **Di, J.**, Demanuele, C., Kettermann, A., Karahanoglu., FI., Capperlleri, JC., Potter, A., Bury, D., Cedarbaum, J., Byrom, B.. Considerations to address missing data when deriving clinical trial endpoints from digital health technologies (DHTs). *Contemporary Clinical Trials.* 2022; 113: 106661.

2. Spira, A. P., Zipunnikov, V., Raman, R., Choi, **Di, J.**, Bai, J., Carlsson, C. M., Mintzer, J. E., Marshall, G. A., Porsteinsson, A. P., Yaari, R., Wanigatunga, S. K., Kim, J., Wu, M. N., Aisen, P. S., Sperling, R. A., and Rosenberg, P. B. Brain amyloid burden, sleep, and 24-hour rest/activity rhythms: screening findings from the Anti-Amyloid Treatment in Asymptomatic Alzheimer's and Longitudinal Evaluation of Amyloid Risk and Neurodegeneration Studies. *SLEEP Advances.* 2021; 2(1): zpab015.

3. Malone, SK., **Di, J.**, Leroux, A., Riegel, B., Melkis, GD., Rickels, MR., Punjabi, NM., Pack, A., Crainiceanu, C., and Urbanek, J.. Bedtime habits in adults with and without type 2 diabetes. *Journal of Articles in Support of the Null Hypothesis.* 2021; 18(1): 33-40.

4. Li, J., Somers, VK., Lopez-Jimenez, F., **Di, J.**, and Covassin, N.. Demographic characteristics associated with circadian rest-activity rhythm patterns: A cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity.* 2021; 18: 107.

5. Kuo, PL., **Di, J.**, Ferrucci, L., and Lin, FR.. Analysis of hearing loss and physical activity among US adults, 60 - 69 years old. *JAMA Network Open.* 2021; 4(4): e215484.

6. Mahadevan, N., Christakis, Y., **Di, J.**, Bruno, J., Zhang Y., Dorsey, ER., Pigeon, WR., Beck, LA., Thomas, K., Liu, Y., Wicker, M., Brooks, C., Kabiri, NS., Bhangu, J., Northcott, C., and Patel, S.. Development of digital measures for nighttime scratch and sleep using wrist-worn wearable devices. *npj Digital Medicine.* 2021; 4: 42.

7. Chen O.Y., Cao, H., Phan H., Nagels, G., Reinen, J., Gou, J., Qian, T., **Di, J.**, Prince, J., Cannon, T., and De Vos, M.. Identifying neural signatures mediating behavioral symptoms and psychosis onset: high-dimensional whole brain functional mediation analysis. *NeuroImage.* 2021; 226: 117508.

8. Schmidt, A., **Di, J.**, Zipunnikov, V., Frey, K., Scharfstein, D., O'Toole, R., Bosse, M., Obremskey, W., Stinner, D., Hayda, R., Karunakar, R., Hak, M., Carroll, E., Collins, S., MacKenzie, E.. Perfusion

pressure lacks diagnostic specificity for the diagnosis of acute compartment syndrome. *Journal of Orthopaedic Trauma.* 2020; 34(6): 287-293.

9. Wanigatunga, A., **Di, J.**, Zipunnikov, V., Urbanek, J., Kuo, P., Simonsick, EM., Ferrucci, L., and Schrack, J.. Association of total daily physical activity and fragmented physical activity with mortality in older adults. *JAMA Network Open.* 2019; 2(10): e1912352.

10. **Di, J.**, Spira, A., Bai, J., Urbanek, J., Leroux, A., Wu, M., Resnick, S., Simonsick, E., Ferrucci, L., Schrack, J., and Zipunnikov, V.. Joint and individual representation of domains of physical activity, sleep, and circadian rhythmicity. *Statistics in Biosciences.* 2019; 11(2): 371-402.

11. Leroux, A., **Di, J.**, Smirnova, E., Mcguffey, E., Cao, Q., Bayatmokhtari, E., Tabacu, L., Zipunnikov, V., Urbanek, J., Crainiceanu, C.. Organizing and analyzing the activity data in NHANES. *Statistics in Biosciences.* 2019; 11(2): 262-287.

12. Schrack, J., Kuo, P., Wanigatunga, A., **Di, J.**, Simonsick, EM., Spira, A., Ferrucci, L., and Zipunnikov, V.. Active-to-Sedentary behavior transitions, fatigability, and physical functioning in older adults. *Journal of Gerontology: Series A.* 2019; 74(4): 560-567.

13. Johns, J., **Di, J.**, Merikangas, K., Cui, L., Swendsen, J., and Zipunnikov, V.. Fragmentation as a novel measure of stability in normalized trajectories of mood and attention measured by ecological momentary assessment. *Psychological Assessment.* 2019; 31(3): 329-339.

14. Schmidt, A., Bosse, M., Obremskey, W., O'Toole, R., Carroll, E., Stinner, D., Hak, D., Karunakar, M., Hayda, R., Frey, K., **Di, J.**, Zipunnikov, V., MacKenzie, E., and METRC. Continuous near-infrared spectroscopy demonstrates limitations in monitoring the development of acute compartment syndrome in patients with leg injuries. *The Journal of Bone & Joint Surgery.* 2018; 100(19): 1645-1652.

15. Grigsby, M., **Di, J.**, Leroux, A., Xiao, L., Zipunnikov, V., Crainiceanu, C., and Checkley, W.. Novel metrics for growth model selection. *Emerging Themes in Epidemiology.* 2018; 15(1): 4.

16. Urbanek, J., Spira, A., **Di, J.**, Leroux, A., Crainiceanu, C., and Zipunnikov, V.. Epidemiology of objectively measured bedtime and chronotype in US adolescents and adults: NHANES 2003-2006. *Chronobiology International.* 2018; 35(3): 416-434.

17. Varma, V., Dey D., Leroux A., **Di, J.**, Urbanek, J., Xiao, L., and Zipunnikov, V.. Total volume of physical activity: TAC, TLAC or TAC(λ). *Preventive Medicine.* 2018; 106: 233-235.

18. Varma, V., Dey D., Leroux A., **Di, J.**, Urbanek, J., Xiao, L., and Zipunnikov, V.. Re-evaluating the effect of age on physical activity over the lifespan. *Preventive Medicine.* 2017; 101: 102-108.

19. **Di, J.**, Li, Y., Friedman, MR., Reddy, S., Surkan, PJ., Shoptaw, S., and Plankey, M.. Determining survey satisficing of online longitudinal survey data in the Multicenter AIDS Cohort Study using a Group-Based Trajectory Analysis. *Journal of Medical Internet Research Public Health and Surveillance.* 2016; 2(2): e150.

Preprints / Technical Reports

20. **Di, J.**, Leroux, A., Urbanek, J., R., Varadhan, Spira, A., Schrack, J., and Zipunnikov, V.. Patterns of sedentary and active time accumulation are associated with mortality in US adults: The NHANES study. *bioRxiv: 182337.*

Under Review / Revision

21. E JY., Schrack, JA., Zipunnikov, V., Whitman, G., Lawton J., Goeddel L., Bush, B., **Di, J.**, Tsay, T., and Brown CH.. Machine learning prediction for two hospital-based outcomes using accelerometer-derived metrics of physical activity. Under review *Annals of Internal Medicine.*

22. **Di, J.**, Leroux, A., Urbanek, J., R., Varadhan, Spira, A., Schrack, J., and Zipunnikov, V.. Two-state transition analysis of within-day active-sedentary profiles provides new insights into mortality risk of US adults. Under review *Preventive Medicine.*

23. Chen, OY., Saraiva, RG., **Di, J.**, Phan, H., Nagels, G., Schwantje, T., Cao, H., Gou, J., Reinen, JM., Xiong, B., and de Vos, M.. The roles and challenges of the P-value. Under review *JASA.*

24. Li, J., Vungarala, S., Somers, VK., **Di, J.**, Lopez-Jimenez,, F., Covassin, N.. Association between rest-activity rhythm and obesity in adults. Under review *EBioMedicine.*

25. Wanigatunga, A., Liu, F., Urbanke, J., Wang, H., **Di, J.**, Zipunnikov, V., Cai, Y., Doughtery, R., Simonsick. E., Ferrucci, L., and Schrack, J.. Physical activity patterns, aging and gait speed in the Baltimore Longtitudinal Study of Agiing. Under review *Journal of Gerontology: Series A.*

26. Brown, C. Yanek, L., Healy, R., Tsay, T., **Di, J.**, Whitman, G., Mandal, L., Madeira T., Goddel, L., Grant, MC., Young, D., Zipunnikov, V., Hoyer, EH., and Schrack, J.. A comparison of three wearable accelerometers and an observation scale to measure early physical activity and mobility after cardiac surgery. Under review *The Journal of Thoracic and Cardiovascular Surgery.*

PATENTS

1. Computer decision support tool for scratch detection and flare prediction. *Provisional patent* filed June 23, 2020. US: 63/043,108.

2. Digital medicine companion for treating and managing skin diseases. *Provisional patent* filed December 20, 2021. US: 63/291,798.

SOFTWARES

1. **cdHOTS** (R package). Construct and decompose higher order tensorian statistics.

   https://github.com/junruidi/cdHOTS

2. **actigraphy-profiles** (R Shiny app). A visualization tool for displaying minute-level actigraphy profiles.

   https://github.com/junruidi/actigraphy-profiles

3. **ActFrag** (R package). Extract activity fragmentation metrics from minute level actigraphy data.

   https://cran.r-project.org/web/packages/ActFrag/index.html
   https://github.com/junruidi/ActFrag

4. **ActCR** (R package). Extract circadian rhythms metrics from actigraphy data.

   https://cran.r-project.org/web/packages/ActCR/index.html
   https://github.com/junruidi/ActCR

PRESENTATIONS / PANELS

1. Extracting Joint Information from Wearable Sensors at Multiple Body Locations (poster). *2021 Global Statistical Conference, Online.*

1. Beyond Step Counting - Measuring Nighttime Scratch andSleep with Wearable Devices (contributed). *2021 ICSA Applied Statistics Syposium, Online.*

2. Development of Digital Measures for Night-Time Scratch and Sleep Using Wrist-Worn Wearable Devices (contributed). *2021 ENAR, Baltimore, MD.*

3. The Future of Statistics and What is at Stake (**panelist**). *2020 Global Statistical Conference, Online.*

4. Wearable Technology to Quantitatively Assess Nighttime Scratch and Sleep in Atopic Dermatitis (**plenary**). *2020 Global Statistical Conference, Online.*

5. Integrative Analysis of Domains of Physical Activity, Sleep, and Circadian Rhythmicity Collected by Wearables (**invited session**). *2019 JSM, Denver, Colorado.*

6. Joint and Individual Representation of Domains of Physical Activity, Sleep, and Circadian Rhythmicity Collected by Wearables (**invited session**). *2018 BigDIA, Houston, TX.*

7. Capturing Enhanced Information with Higher-Order Tensorian Statistics and Predicting Mortality from Accelerometry-Measured Physical Activity (contributed). *2018 JSM, Vancouver, BC, Canada.*

8. Analysis of Tensor Cumulants and Its Application to NHANES (contributed). *2018 ENAR, Atlanta, GA.*

9. Fragmentation of Daily Physical Activity: Prediction of Mortality in NHANES 2003-2006 (oral). *2017 IAGG, San Francisco, CA.*

10. Fragmentation of Physical Activity and Its Application (oral). *2017 ICAMPAM, Bethesda, MD.*

11. Novel Statistical Framework to Quantify Fragmentation of Physical Activity (contributed). *2017 ENAR, Washington, DC.*

12. Integrative Analysis of Multi-Block Contaminated Datasets (topic contributed). *2013 JSM, Montreal, Canada.*

EDITORIAL ACTIVITIES

Referee for:

Journal of Statistical Software (JSS) [1]

International Association of Gerontology and Geriatrics 2017 World Congress (IAGG) [1]

Journal of Medical Internet Research Cardio (JMIR Cardio) [1]

Journal of Medical Internet Research Mental Health (JMIR Mental Health) [1]

Journal of Medical Internet Research mHealth and uHealth (JMIR mHealth and uHealth) [5]

Journal of Medical Internet Research Biomedical Engineering (JMBE) [1]

Interactive Journal of Medical Research [1]

Statistics in Biosciences (SIB) [1]

Journal of Medical Internet Research Research [1]

Journal of Medical Internet Research Research Protocols (JRP) [1]

Journal of Medical Internet Research Serious Games (JSG) [1]

BMC Geriatrics [2]

Medicine & Science in Sports & Exercise (MSSE) [1]

PLOS One [2]

PROFESSIONAL ACTIVITIES

Co-founder of Pfizer Artificial Intelligence Network   May 2020 - Present

Co-founder and co-organizer of Pfizer Artificial Intelligence Network Talk Series   June 2020 - Present

Session chair, JSM   Aug 2018

Organizer of the JHSPH Biostatistics Computing Club   Aug 2015 - May 2016

HONORS & AWARDS

Early Clinical Development Cross Functional Team Courage Award   *Dec 2021*

ECD Courage Award recognizes colleagues who have made courageous decisions through data generation, analysis, synthesis, communication, and behavior to ensure Pfizer's success.

For: "Designing, planning, and conducting the PfiRe Lab studies to collect at-home gait and physical activity data from both adults and pediatric populations for the purpose of the master protocol development with wearable devices."

Recognizing our Values Award   *Jan 2021*

This award is for recognizing colleagues whose special accomplishment furthers Pfizer's achievement of its purpose - Breakthroughs that change patients' lives. The program is designed to recognize and reward colleagues for an exceptional contribution to the business while demonstrating the Pfizer values and culture.

For: "Publishing a high impact peer-reviewed paper helping Pfizer lead the conversation pertaining to digital endpoints, which contributes to advancing the field, in an effort to ultimately reach a point where digital clinical measurements are utilized in a 'business as usual' manner."

The W.E. Upjohn Prizes Award   *Dec 2020*

This award recognizes and rewards colleagues for an exceptional contribution to the business while demonstrating Pfizer values and culture. Prizes should be awarded to colleagues for contributing above and beyond ones' job & responsibility to an outcome that helps achieve a business goal.

For: "Being a critical player in the design, implementation, conduct, and close of the SPA internship program by demonstrating Pfizer value of excellence, and helped to achieve the end goal and complete the hiring targets for 2020."

WRDM Team Achievement Award Finalists   *Jul 2020*

The WRDM Achievement Awards program recognizes colleagues who demonstrate exceptional commitment to our Values; make extraordinary scientific, strategic, and/or operational contributions to our business; and whose innovations tangibly impact Pfizer's ability to achieve our Purpose: Breakthroughs that change patients' lives.

The W.E. Upjohn Prizes Award   *Sep 2019*

This award recognizes and rewards colleagues for an exceptional contribution to the business while demonstrating Pfizer values and culture. Prizes should be awarded to colleagues for contributing above and beyond ones' job & responsibility to an outcome that helps achieve a business goal.

For: "Being an excellent team member within Pfizer by dropping all his ongoing workload to assist in completing the QC task for an external conference presentation in light of the short timelines to ensure confidence and the quality of the data being produced to the level necessary."

BigDIA Travel Award   *Sep 2018*

This is the travel award for the 4th International Conference on Big Data and Information Analytics.

The June B. Culley Award   *Dec 2017*

This award honors outstanding achievement by a Biostatistics student on his or her schoolwide examination paper.

The Louis I. and Thomas D. Dublin Award   *Mar 2017*

This award, which is for the Advancement of Epidemiology and Biostatistics, supports those students whose research focuses on the effective use of statistical reasoning and methods in epidemiology.

Washington Statistical Society Outstanding Graduate Student Award   *Jun 2013*

This award is presented by the Washington Statistical Society to the outstanding full-time graduate student of statistics/biostatistics at each university in the Washington metropolitan area with a graduate statistics/biostatistics program.

CERTIFICATIONS

AI for Medical Prognosis (Coursera Certificates)   *May 2020*

AI for Medical Diagnosis (Coursera Certificates)   *May 2020*

Deep Learning Specialization (Coursera Certificates)   *Dec 2019*
- Neural Networks and Deep Learning
- Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization
- Structuring Machine Learning Projects
- Convolutional Neural Networks
- Sequence Models

SAS Certified Advanced Programmer for SAS 9   *Aug 2013*

SAS Certified Base Programmer for SAS 9   *Jul 2013*