

# JUNRUI DI

dijunrui@gmail.com ♦ <https://junruidi.github.io/>

(Updated December 8th 2020)

## STATISTICAL METHODS RESEARCH INTERESTS

---

feature engineering from accelerometry signals, dimension reduction, functional data analysis, integration of multiple modalities, machine learning.

## SCIENTIFIC RESEARCH INTERESTS

---

wearable devices and their applications in public health (e.g. mental health and aging) and clinical trials (e.g. disease monitoring and assessment), physical activity assessment, sleep, circadian rhythmicity.

## EDUCATION

---

**Johns Hopkins Bloomberg School of Public Health**

*Mar 2019*

Ph.D. in Biostatistics

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

Advisor: Vadim Zipunnikov, Ph.D.

**Georgetown University**

*Dec 2013*

M.S. in Biostatistics

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

Advisor: Valeriy Korostyshevskiy, Ph.D.

**University of California, Berkeley**

*May 2012*

B.A. in Applied Mathematics

*High Distinction General Scholarship (roughly equivalent to Magna Cum Laude)*

## EXPERIENCE

---

**Pfizer**

*Cambridge, MA*

*Manager, Digital Medicine Statistics*

*Apr 2019 - Present*

- Responsible for ensuring sound statistical thinking and methods are utilized in the discovery and development of new digital and wearable biomarkers by bringing the principles of objective decision-making into the development, validation and implementation of industry-leading mobile/digital technology programs for incorporation into clinical trials.
- Interact with internal and external experts to assure sound quantitative approaches are applied to collection and analysis of a wide variety of digital and wearable data types.
- Apply statistical expertise to prepare both internal and external reports, presentations, manuscripts.
- Co-lead and co-organize the Pfizer Artificial Intelligence Networking group.

**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.
- Conducted data processing and statistical analysis of multiple studies with measurements of sleep, activity, and circadian rhythms using wearable devices.

**Takeda Pharmaceuticals***Biostatistician Intern**Cambridge, MA**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***Research Assistant, Department of Medicine**Washington, DC**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. 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.
2. Schmidt, A., **Di, J.**, Zipunnikov, V., Frey, K., Scharfstein, D., O’Toole, R., Bosse, M., Obrebsky, 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.
3. 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.
4. **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.
5. 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.
6. 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.
7. 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.

8. 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.
9. 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.
10. 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.
11. Varma, V., Dey D., Leroux A., **Di, J.**, Urbanek, J., Xiao, L., and Zipunnikov, V.. Total volume of physical activity: TAC, TLAC or TAC( $\lambda$ ). *Preventive Medicine*. 2018; 106: 233-235.
12. 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.
13. **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**

14. **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**

15. Spira, AP., Zipunnikov, V., Raman, R., Choi, J., **Di, J.**, Bai, J., Carlsson, CM., Mintzer, JE., Marshall, GA., Porsteinsson, AP., Yaari, R., Wanigatunga, SK., Kim, J., Wu, MN., Aisen, PS., Sperling, RA., and Rosenberg, PB.. Brain amyloid burden, sleep, and 24-hour rest/activity rhythms: screening findings from A4 and LEARN. Under review *Sleep*.
16. 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. Under review *npj Digital Medicine*.
17. Kuo, PL., **Di, J.**, Ferrucci, L., and Lin, FR.. Hearing loss and physical activity: A cross-sectional study from National Health and Nutrition Examination Survey 2003 – 2004. Under review *JAMA Internal Medicine*.
18. 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*.
19. Lin, Z., Rojo-Wissar, DM., Nestadt, P., Urbanek, J., **Di, J.**, Zipunnikov, V., and Spira, Ap.. Actigraphy-measured circadian factors and mortality in US adults: Results from NHANES. To be submitted to *Sleep*.
20. 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. To be submitted to *JASA*.

#### **In Preparation**

21. **Di, J.**, and Zipunnikov, V.. Two-state transition analysis of accelerometry-estimated active-sedentary patterns for modelling mortality risk in US adults. To be submitted to *Preventive Medicine*.

22. **Di, J.**, and Zipunnikov, V.. Capturing enhanced information with higher-order tensorial statistics and its application in accelerometry-measured physical activity.
23. **Di, J.**, and Zipunnikov, V.. Multilevel analysis of multivariate data from multiple domains and its application to model accelerometry-measured physical activity, sleep, and circadian rhythmicity.
24. **Di, J.**, and Zipunnikov, V.. Longitudinal joint and individual variance explained to model longitudinal effect of cognitive and physical functions and their codependency on each other.

## PATENTS

---

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

## 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>
4. **ActCR** (R package). Extract circadian rhythms metrics from actigraphy data.  
<https://github.com/junruidi/ActCR>

## PRESENTATIONS

---

1. The Future of Statistics and What is at Stake. (**panelist**). *2020 Global Statistical Conference, Online*.
2. Wearable Technology to Quantitatively Assess Nighttime Scratch and Sleep in Atopic Dermatitis. (**plenary**). *2020 Global Statistical Conference, Online*.
3. Integrative Analysis of Domains of Physical Activity, Sleep, and Circadian Rhythmicity Collected by Wearables. (**invited**). *2019 JSM, Denver, Colorado*.
4. Joint and Individual Representation of Domains of Physical Activity, Sleep, and Circadian Rhythmicity Collected by Wearables (**invited**). *2018 BigDIA, Houston, TX*.
5. Capturing Enhanced Information with Higher-Order Tensorian Statistics and Predicting Mortality from Accelerometry-Measured Physical Activity (contributed). *2018 JSM, Vancouver, BC, Canada*.
6. Analysis of Tensor Cumulants and Its Application to NHANES (contributed). *2018 ENAR, Atlanta, GA*.
7. Fragmentation of Daily Physical Activity: Prediction of Mortality in NHANES 2003-2006 (oral). *2017 IAGG, San Francisco, CA*.
8. Fragmentation of Physical Activity and Its Application (oral). *2017 ICAMPAM, Bethesda, MD*.
9. Novel Statistical Framework to Quantify Fragmentation of Physical Activity (contributed). *2017 ENAR, Washington, DC*.

10. 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 [1]

## PROFESSIONAL ACTIVITIES

---

Co-founder and organizer of Pfizer Artificial Intelligence Network Talk Series	May 2020 - Present
Session chair, JSM	Aug 2018
Organizer of the JHSPH Biostatistics Computing Club	Aug 2015 - May 2016

## HONORS & AWARDS

---

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

*This award recognizes and rewards colleagues for an exceptional contribution to the business while demonstrating the 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.*

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 the 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.*

BigDIA Travel Award	Sep 2018
<i>This is the travel award for the 4th International Conference on Big Data and Information Analytics.</i>	
The June B. Culley Award	Dec 2017
<i>This award honors outstanding achievement by a Biostatistics student on his or her schoolwide examination paper.</i>	
The Louis I. and Thomas D. Dublin Award	Mar 2017
<i>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.</i>	
Washington Statistical Society Outstanding Graduate Student Award	Jun 2013
<i>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.</i>	
Phi Beta Kappa Honor Society Inductee	May 2012
<i>Phi Beta Kappa Honor Society</i>	

## TEACHING EXPERIENCE

---

PH.140.623 - <b>Lab Instructor</b> Statistical Methods in Public Health III	Spring 2019
PH.140.621 - <b>Lab Instructor</b> Statistical Methods in Public Health I	Fall 2018
PH.140.623 - <b>Lab Instructor</b> Statistical Methods in Public Health III	Spring 2018
PH.140.621 - <b>Lab Instructor</b> Statistical Methods in Public Health I	Fall 2017
PH.140.623-4 - <b>TA</b> Statistical Methods in Public Health III-IV	Spring 2017
PH.140.621-2 - <b>TA</b> Statistical Methods in Public Health I-II	Fall 2016
PH.140.753-4 - <b>TA</b> Advanced Methods in Biostatistics III-IV	Spring 2016
PH.140.751-2 - <b>TA</b> Advanced Methods in Biostatistics I-II	Fall 2015
BIST 514 - <b>TA</b> Linear Modeling & Multivariate Analysis	Spring 2014

## CERTIFICATIONS

---

AI for Medical Prognosis (Coursera Certificates)	May 2020
AI for Medical Diagnosis (Coursera Certificates)	May 2020
Deep Learning Specialization (Coursera Certificates)	Dec 2019
<ul style="list-style-type: none"> <li>- Neural Networks and Deep Learning</li> <li>- Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization</li> <li>- Structuring Machine Learning Projects</li> <li>- Convolutional Neural Networks</li> <li>- Sequence Models</li> </ul>	
SAS Certified Advanced Programmer for SAS 9	Aug 2013
SAS Certified Base Programmer for SAS 9	Jul 2013