










# Debangon Dey, Ph.D.

✉ debangan.dey@nih.gov      Ddey07  
 <https://debangandey.rbind.io/>




## Overview

- Summary      A biostatistician focused on developing methods to analyze intensive multilevel multimodal longitudinal data and big geospatial data. Applications focus on Digital Health Technologies (DHTs) and Environmental Sciences.
- Areas of interest      Functional data analysis, longitudinal data, spatial statistics, stochastic processes, semiparametric statistics.

## Work Experience

- 07/14/2025 – ...      **Assistant Professor.** Texas A&M University Department of Statistics.
- 05/20/2022 – 05/20/ 2025      **Postdoctoral fellow.** National Institute of Mental Health. Lead analyst for Motor Activity Research Consortium for Health (mMarch).
- 05/24/2021 – 08/27/2021      **Applied Scientist Intern.** Amazon. Using heterogeneous treatment effects to reduce length of A/B testing trials in Amazon search team.
- 2020      **Analyst.** Baltimore City Health Department - Johns Hopkins Data Analysis Support Team for COVID-19 Response.
- 2015, 2016      **Summer Intern.** Johns Hopkins Bloomberg School of Public Health Dept. of Biostatistics.

## Education

- 2017 – 2022      **Ph.D., Johns Hopkins Bloomberg School of Public Health** in Biostatistics.  
Thesis title: *Topics in Modeling of Multivariate Mixed Data Types and Highly Multivariate Spatial Data*
- 2015 – 2017      **Masters, Indian Statistical Institute** in Statistics.  
Thesis title: *On Investigation of Genotypic Association with Longitudinal Observations of Traits.*
- 2012 – 2015      **Bachelors, Indian Statistical Institute** in Statistics.

## Research Publications and Repositories

### Journal Articles

- 1     **D. Dey**, S. Banerjee, M. A. Lindquist, and A. Datta, “Graph-constrained analysis for multivariate functional data,” *Journal of Multivariate Analysis*, p. 105 428, 2025.
- 2     **D. Dey**, R. Ghosal, K. Merikangas, and V. Zipunnikov, “Functional principal component analysis for continuous non-gaussian, truncated, and discrete functional data,” *Statistics in Medicine*, 2024.
- 3     T. M. Lateef, **D. Dey**, A. Leroux, L. Cui, M. Xiao, V. Zipunnikov, and K. R. Merikangas, “Association between electronic diary-rated sleep, mood, energy, and stress with incident headache in a community-based sample,” *Neurology*, vol. 102, no. 4, e208102, 2024, **Article featured at CNN, Nat Geo, NIH Director’s blog and others.**
- 4     **D. Dey**, A. Datta, and S. Banerjee, “Modeling multivariate spatial dependencies using graphical models,” *The New England Journal of Statistics in Data Science*, pp. 1–13, 2023.

- 5 J. Glaus, S. J. Kang, W. Guo, F. Lamers, M.-P. F. Strippoli, A. Leroux, **D. Dey**, K. J. Plessen, J. Vaucher, P. Vollenweider, *et al.*, “Objectively assessed sleep and physical activity in depression subtypes and its mediating role in their association with cardiovascular risk factors,” *Journal of psychiatric research*, vol. 163, pp. 325–336, 2023.
- 6 E. K. Stapp, V. Zipunnikov, A. Leroux, L. Cui, M. M. Husky, **D. Dey**, and K. R. Merikangas, “Specificity of affective dynamics of bipolar and major depressive disorder,” *Brain and Behavior*, e3134, 2023.
- 7 **D. Dey**, A. Datta, and S. Banerjee, “Graphical gaussian process models for highly multivariate spatial data,” *Biometrika*, vol. 109, no. 4, pp. 993–1014, 2022.
- 8 M. R. Kaufman, **D. Dey**, C. Crainiceanu, and M. Dredze, “# metoo and google inquiries into sexual violence: A hashtag campaign can sustain information seeking,” *Journal of interpersonal violence*, vol. 36, no. 19–20, pp. 9857–9867, 2021.
- 9 **D. Dey** and V. Zipunnikov, “Discussion of “an epidemiological forecast model and software assessing interventions on the covid-19 epidemic in China”,” *Journal of Data Science*, vol. 18, no. 3, pp. 433–436, 2020.
- 10 S. Deb and **D. Dey**, “Spatial modeling of shot conversion in soccer to single out goalscoring ability,” *Journal of Sports Analytics*, vol. 5, no. 4, pp. 281–297, 2019.
- 11 V. R. Varma, **D. Dey**, A. Leroux, J. Di, J. Urbanek, L. Xiao, and V. Zipunnikov, “Total volume of physical activity: TAC, TLAC or TAC ( $\lambda$ ),” *Preventive medicine*, vol. 106, pp. 233–235, 2018.
- 12 A. Spira, V. Zipunnikov, M. Wu, **D. Dey**, E. Simonsick, L. Ferrucci, C. Davatzikos, and S. Resnick, “Association of circadian rest and activity rhythms with brain volumes in cognitively normal older adults,” *Innovation in Aging*, vol. 1, no. suppl\_1, pp. 866–866, 2017.
- 13 V. R. Varma, **D. Dey**, A. Leroux, J. Di, J. Urbanek, L. Xiao, and V. Zipunnikov, “Re-evaluating the effect of age on physical activity over the lifespan,” *Preventive medicine*, vol. 101, pp. 102–108, 2017, **Article featured at TIME, Washington Post, WSJ, BBC, WPYR and others.**
- 14 V. Zipunnikov, **D. Dey**, A. Leroux, J. Di, J. Urbanek, J. Schrack, and C. Crainiceanu, “Total physical activity and its circadian allocation are independent predictors of mortality,” *Innovation in Aging*, vol. 1, no. Suppl 1, p. 1239, 2017.

## Preprints and Under Revision

- 1 **D. Dey** and V. Zipunnikov, *Semiparametric gaussian copula regression modeling for mixed data types (SGCRM)*, 2022. arXiv: 2205.06868 [stat.ME].
- 2 **D. Dey** and V. Zipunnikov, *Connecting population-level auc and latent scale-invariant  $R^2$  via semiparametric gaussian copula and rank correlations*, 2019. arXiv: 1910.14233 [stat.ME].

## Packages and Repositories









- 1 **D. Dey** and R. Ghosal, *R package – SGCTools: Analysis tools for multivariate mixed data and continuous/truncated/discrete functional data using Semiparametric Gaussian Copula (SGC)*, <https://github.com/Ddey07/SGCTools>, 2023.
- 2 **D. Dey** and A. Pita, *The Good, The Bad, and The Ugly of the Beautiful Game — Analyzing FIFA World Cup 2018 data*, <https://ddey07.github.io/open-data/>, 2018.

## Books and Chapters





- 1 S. Deb and **D. Dey**, *Mathematical Techniques for Competitive Examinations*, English, Paperback, ISBN: 978-9393330109.  URL: <https://lead.to/amazon.com/?op=bt&la=en&cu=usd&key=9393330107>.

## Miscellaneous Experience




### Awards and Achievements

- 2022  **Inducted into the Alpha chapter of Delta Omega**, national honorary society in public health.
- 2021  **Selected among top six presenters**, Opta Pro Fourm, 2021, one of the biggest soccer analytics conferences attended by professionals from over 80 clubs and federations worldwide.
-  **Joint Statistical Meetings paper award**, *Section on Bayesian Statistical Science* of American Statistical Association.
- 2020  **Joint Statistical Meetings paper award**, *Section on Survey Research and Methods* of American Statistical Association.
- 2018  **Top five finalist**, US Soccer Hackathon 2018, Chicago, USA, July 14-15, 2018.
- 2012  **INSPIRE scholarship**, Department of Science and Technology, Govt. of India.
- 2011  **Top 5%**, UNESCO Science Olympiad.
- 2011, 2012  **Awardee**, Regional Math Olympiad, India.



### Conferences

- 2023  **JSM 2023**. Organized the session "Recent developments in methods for digital Brain Health data".
- 2022  **24th Annual NIMH IRP Scientific Training Day**. Presented the poster *Influences of mood, energy and sleep on incident headache assessed with prospective real time electronic assessments*.
- 2021  **ICAMPAM 2021**. Presented the talk *Using Mobile Technologies to Investigate Impaired Sleep, Mood, and Energy as Real-Time Triggers of Migraine*.
-  **MASS 2021**. Presented the talk *Modelling of mixed type intensive longitudinal data via Semi-parametric Gaussian Copula and its application to real-time mobile monitoring of daily health behaviors*.


### Teaching

- 2018-2019  Teaching assistant for Probability Theory I-IV, PH.140.721-724.
- 2019-2022  Teaching assistant for Statistical Methods in Public Health I-IV, 140.621-140.624.
- 2015-2017  Instructor at Ramanujan School of Mathematics for two years, giving lessons on advanced high school mathematics.

### Mentoring



- 2020-2022  Mentor for final term Capstone projects for a diverse group of multi-ethnic international students (25 countries) from non-statistical backgrounds, taking the popular MPH course at Johns Hopkins Bloomberg School of Public Health.
- 2022-2024  Mentoring two PostBacs at NIMH - Ananya Swaminathan, Rene Chaudhuri on the statistical components of their projects.

### Reviewer

- 2021-  Journal of Computational and Graphical Statistics, Annals of Applied Statistics, Econometrics and Statistics, Mental Health and Physical Activity, TEST, Environmetrics, Sensors.

## Skills

---

Languages      Strong reading, writing and speaking competencies for English, Bengali, Hindi.  
Coding         R, Python, HTML,  $\text{\LaTeX}$

## References

---

**Dr. Vadim Zipunnikov, *PhD advisor***

Associate Professor  
Johns Hopkins Bloomberg School of Public Health,  
Dept. of Biostatistics  
Email: [vzipunn1@jhu.edu](mailto:vzipunn1@jhu.edu)

**Dr. Kathleen Merikangas, *Postdoc advisor***

Chief of the Genetic Epidemiology Research  
Branch  
National Institute of Mental Health,  
NIH  
Email: [merikank@mail.nih.gov](mailto:merikank@mail.nih.gov)

**Dr. Abhirup Datta, *PhD advisor***

Associate Professor  
Johns Hopkins Bloomberg School of Public Health,  
Dept. of Biostatistics  
Email: [abhidatta@jhu.edu](mailto:abhidatta@jhu.edu)

**Dr. Sudipto Banerjee, *Collaborator***

Professor  
University of California Los Angeles,  
Fielding School of Public Health,  
Dept. of Biostatistics  
Email: [sudipto@ucla.edu](mailto:sudipto@ucla.edu)