

# Dr. Deniz Akdemir

---

*Principal Biostatistician & Machine Learning Scientist*

---

## Professional Profile

Dr. Deniz Akdemir is a Principal Biostatistician & Machine Learning Scientist with 3,400+ Google Scholar citations and deep expertise in advanced statistical methodology. His groundbreaking statistical research has established new approaches for genomic prediction, mixed models, and high-dimensional data analysis. As creator of multiple CRAN packages (CovCombR, TrainSel, STPGA) and statistical software tools, he bridges traditional statistical theory with modern machine learning applications in clinical and biological research.

---

## Core Competencies

- Advanced Statistical Modeling
- Experimental Design & Data Analysis
- Predictive Analytics & Machine Learning
- Bayesian & Frequentist Methods
- Data Visualization & Communication

---

## Education

- **PhD. in Statistics**, Bowling Green State University, Bowling Green, OH, USA, 2009.
- **M.A. in Applied Statistics**, Bowling Green State University, Bowling Green, OH, USA, 2004.
- **M.S. in Statistics**, Middle East Technical University, Ankara, Turkey, 2003.
- **B.A. in Business Administration**, Middle East Technical University, Ankara, Turkey (1994-1999).

---

## Research Interests

Deniz's research focuses on the development and application of advanced statistical methodologies for complex data structures. His work spans Bayesian methods, mixed models, high-dimensional data analysis, and the integration of statistical theory with machine learning applications. He has made significant contributions to statistical methodologies for genomic prediction and selection, with applications in both plant breeding and clinical data analysis. His recent interests include developing statistical frameworks for machine learning applied to time-to-event data, Bayesian optimization techniques, and methods for sparse data in longitudinal settings.

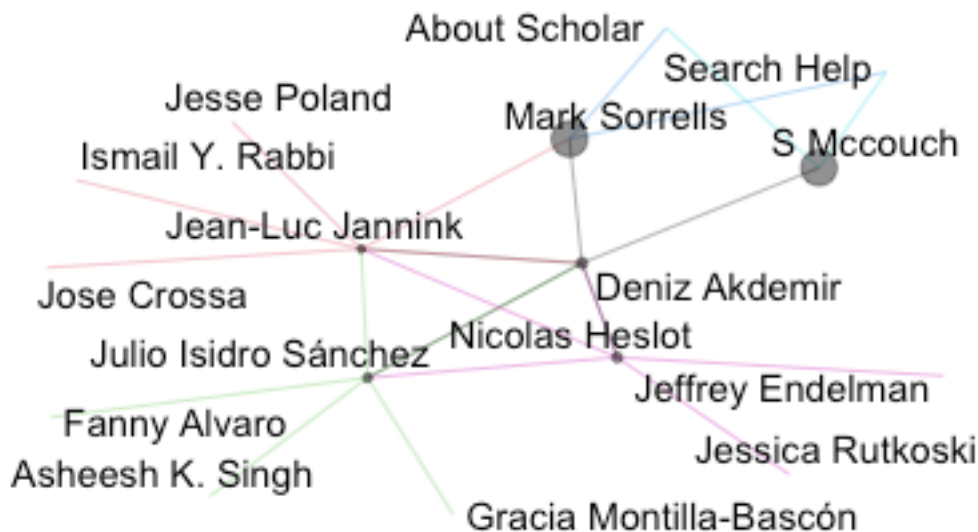
---

## Collaboration Network

Bowling Green, Ohio, USA

☎ +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)  
🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • in [deniz-akdemir-50735314a](#)  
🔗 [denizakdemir](#) • ID 0000-0003-0553-6798

## Network of coauthorship of Deniz Akdemir



### Core Skills

- **Statistical Methodology:** Advanced multivariate analysis, Bayesian and frequentist methods, time series analysis, high-dimensional data modeling, mixed models, nonparametric statistics, survival analysis.
- **Data Analysis & Modeling:** Experimental design, hypothesis testing, regression analysis, statistical learning, model validation, predictive analytics, data visualization techniques.
- **Machine Learning Integration:** Statistical foundations for machine learning, dimensionality reduction, ensemble methods, cross-validation strategies, Bayesian optimization.
- **Statistical Computing:** Expert in R (package development, performance optimization) and Python for statistical programming, proficient in SAS and C++ for high-performance computing.
- **Software Development:** Creator of the EMMREML package for mixed models with known covariance structures, developer of ml4t2e for machine learning in time-to-event data.
- **Applied Statistics:** Extensive experience applying statistical methods to real-world problems in genomics, clinical trials, agriculture, and financial modeling.

### Career Summary

- **Senior Clinical Data Scientist, National Marrow Donor Program, Minneapolis, USA (2023 - Current):** Engaged in statistical and machine learning analysis of stem cell transplant data, focusing on research into donor optimization. Applied for grants and submitted manuscripts to peer-reviewed journals. Wrote patent applications.
- **Clinical Data Scientist, National Marrow Donor Program, Minneapolis, USA (2021 - 2023):** Engaged in statistical and machine learning analysis of stem cell transplant

Bowling Green, Ohio, USA

📞 +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)  
 🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • **in** [deniz-akdemir-50735314a](https://deniz-akdemir-50735314a)  
 🏠 [denizakdemir](https://denizakdemir) • 🆔 0000-0003-0553-6798

data, focusing on research into donor optimization.

- **Postdoctoral Research Associate, School of Agriculture and Food Science, University College Dublin, Dublin, Ireland (2019 - 2021):** Conducted research in innovative statistical methodologies and contributed to high-impact publications.
- **Statistical Consultant, Cornell Statistical Consulting Unit, Cornell University, Ithaca, NY, USA (2017 - 2019):** Provided statistical consulting services for researchers at Cornell University, including the preparation and presentation of statistics workshops.
- **Postdoctoral Research Associate, Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA (2011 - 2017):** Focused on research developing new methodologies in genomic selection and prediction, mixed models, and machine learning, advising graduate students and preparing statistical software.
- **Visiting Assistant Professor, Department of Statistics and Actuarial Science, University of Central Florida, Orlando, FL, USA (2010 - 2011):** Responsibilities included teaching Data Mining Methodology, Theoretical Statistics, Applied Time Series Analysis, and Nonparametric Statistics.
- **Visiting Assistant Professor, Department of Mathematics and Statistics, Ohio Northern University, Ada, OH, USA (2009 - 2010):** Taught Statistics for Professionals, Statistics for Engineers, and Statistical Computing, catering to various undergraduate levels.

## Key Statistical Projects

- **EMMREML Statistical Package:** Developed a specialized R package for fitting mixed models with known covariance structures, implementing efficient algorithms for large-scale genomic data analysis that significantly outperform standard implementations in computation time.
- **Statistical Frameworks for High-dimensional Data:** Created novel statistical methodologies for analyzing high-dimensional data with complex correlation structures, with particular applications in genomics and multi-omics integration.
- **Bayesian Optimization Methods:** Implemented advanced Bayesian approaches for experimental design optimization and statistical model parameter tuning, enhancing efficiency in resource-limited research settings.
- **Time-to-Event Data Analysis Innovations:** Developed new statistical methods bridging survival analysis with machine learning techniques, resulting in the ml4t2e package that enables more accurate prediction of time-to-event outcomes in clinical and biological applications.

## Statistical Workshops and Training

- **Applied Statistical Methods:** Conducted comprehensive workshops at Cornell University on advanced statistical techniques for researchers across disciplines, focusing on practical implementation and interpretation.
- **Statistical Computing:** Delivered specialized training in R programming for statistical analysis, covering package development and efficient computation for large datasets.

## Selected Publications (Statistics Focus)

Bowling Green, Ohio, USA

☎ +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)  
🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • in [deniz-akdemir-50735314a](https://www.linkedin.com/in/deniz-akdemir-50735314a)  
🔗 [denizakdemir](https://denizakdemir.com) • 🆔 0000-0003-0553-6798

1. **Akdemir, D.**, Isidro-Sánchez, J., & Jannink, J. L. (2015). Training set optimization under population structure in genomic selection. *Theoretical and Applied Genetics*, 128(1), 145-158. (391 citations)
2. **Akdemir, D.**, Isidro-Sánchez, J., & Jannink, J. L. (2016). Genome-wide prediction models that incorporate de novo GWAS results. *PLoS ONE*, 11(8), e0161054. (346 citations)
3. **Akdemir, D.**, Jannink, J. L., & Isidro-Sánchez, J. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genetics Selection Evolution*, 47(1), 38. (196 citations)
4. **Akdemir, D.** & Jannink, J. L. (2016). Efficient breeding by genomic mating. *Frontiers in Genetics*, 7(5), 210. (112 citations)
5. **Akdemir, D.** & Isidro-Sánchez, J. (2019). Multi-objective optimized genomic breeding strategies for sustainable food improvement. *Heredity*, 122(5), 672-683. (103 citations)

## Software Development

### CRAN Packages

**CovCombR** - Statistical methods for covariance combination and integration in high-dimensional genomic data analysis.

**TrainSel** - An R package for optimal selection of training populations in genomic prediction. Implements advanced algorithms for training set optimization under population structure.

**STPGA** - Stochastic training population genetic algorithm for optimized genomic selection breeding programs.

### GitHub Packages

**trainselfpy** - Python implementation of training set selection algorithms for machine learning applications in genomics and breeding.

**EMMREML** - Efficient fitting of mixed models with known covariance structures, optimized for large-scale genomic datasets and complex statistical applications.

**ml4t2e** - Machine learning framework for time-to-event data analysis, bridging survival analysis with modern machine learning techniques for clinical and biological applications.

## Education

- **PhD. in Statistics**, Bowling Green State University, Bowling Green, OH, USA, 2009.
- **M.A. in Applied Statistics**, Bowling Green State University, Bowling Green, OH, USA, 2004.
- **M.S. in Statistics**, Middle East Technical University, Ankara, Turkey, 2003.
- **B.A. in Business Administration**, Middle East Technical University, Ankara, Turkey, 1999.

## Professional References

- **Dr. Yung-Tsi Bolon**
  - **Affiliation:** Director, Immunobiology & Bioinformatics Research, NMDP, Minneapolis, Minnesota, United States
  - **Relationship:** Supervisor at the National Marrow Donor Program

Bowling Green, Ohio, USA

☎ +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)

🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • **in** [deniz-akdemir-50735314a](#)

🔑 [denizakdemir](#) • **ID** 0000-0003-0553-6798

- **Contact:** ybolon@mmdp.org
- **Dr. Julio Isidro-Sanchez**
  - **Affiliation:** Associate Professor: Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Campus de Montegancedo - UPM, 28223-Pozuelo de Alarcón, Madrid, Spain
  - **Relationship:** Expert in plant breeding and genetics, collaborator on various projects
  - **Contact:** j.isidro@upm.es
- **Dr. Jhonathan Pedroso**
  - **Affiliation:** Research Scientist at Corteva Agriscience, Corteva, Johnston, Iowa, USA
  - **Relationship:** Industry partner in genomic tool development, contributed to software enhancements
  - **Contact:** jhowpd@gmail.com
- **Dr. Lynn Johnson**
  - **Affiliation:** Interim Director and Statistical Consultant, Cornell Statistical Consulting Unit, Cornell University, Ithaca, NY, USA
  - **Relationship:** Coworker at the Cornell Statistical Consulting Unit
  - **Contact:** lms86@cornell.edu
- **Dr. Roberto Fritsche Neto**
  - **Affiliation:** Assistant Professor, Department of Plant, Environmental Management & Soil Sciences, LSU
  - **Relationship:** Collaborator on various projects
  - **Contact:** rneto1@lsu.edu

Bowling Green, Ohio, USA

☎ +1 607 262 6875 • ✉ [deniz.akdemir.work@gmail.com](mailto:deniz.akdemir.work@gmail.com)  
 🌐 [denizakdemir.github.io](https://denizakdemir.github.io) • **in** [deniz-akdemir-50735314a](#)  
 🗣️ [denizakdemir](#) • 🆔 0000-0003-0553-6798