Janick Weberpals, RPh, Ph.D.

Health Data Scientist

Born 1989 in Germany,
Current location: Boston, MA

janickweberpals.github.io

+1 (857) 381-7865

linkedin.com/in/weberpals

janick.weberpals@me.com

About me

I'm a healthcare data scientist with 8+ years of experience in the design and analysis of large *real-world* clinical database studies in both industry and academia. I'm passionate about the integration of various data modalities (EHR, claims, genomics, imaging, NLP) using deep learning to solve problems in healthcare and causal inference. In addition, I have gained deep clinical knowledge mainly in the fields of cancer but also other disease areas. I have (co-) authored 25+ peer-reviewed methodological and clinical publications and received several awards.

Skills

- R, RStudio, git
- Medical coding standards (ICD, etc.)
- Python, Tensorflow, Keras
- LaTeX, Markdown, Quarto
- HPC, SLURM, Unix
- SQL
- SAS

Education

2018-2020 **Postdoctoral Fellowship** Data Science, Roche Innovation Center, Germany
Deep learning on electronic health record (EHR) data

2015-2018 **Ph.D. Epidemiology** Medical Faculty, Heidelberg University, Germany
Graduated with *Summa cum laude* honors

2015-2018 **Board certification** Bavarian Chamber of Pharmacists, Munich, Germany
Specialized Pharmacist in Drug Information

2010-2015 **Registered Pharmacist** College of Pharmacy, Marburg University, Germany
Pharmaceutical Sciences (PharmD)

Professional Experience

2022- **Instructor in Medicine** Harvard Medical School, Boston, MA, USA
Faculty at Harvard Medical School & Brigham and Women's Hospital leading innovative projects together with FDA and other major institutions utilizing advanced analytics to leverage routinely collected healthcare data (EHR, imaging, claims) to generate high-quality real-world evidence studies of medical interventions. Co-PI of a Sentinel Innovation Center project developing principled methods and *smdi R package* to handle missing covariates in causal inference.

2020-2022 **Data Scientist** Hoffmann-La Roche/Genentech, Basel, Switzerland
Awarded "Exceptional Performance" in 2021. Led and collaborated in cross-functional teams to implement and validate NLP models and real-world database studies which supported and expedited clinical teams with regulatory drug applications in oncology and neuroscience. Contributed to R package development which increased speed, transparency and validity of RWE projects.

2018-2020 **Postdoctoral Fellow in Deep Learning on EHR** Roche, Munich, Germany
Implemented *deep learning methods* to analyze large EHR databases which supported protocol design and strategic decision making in early-stage single-arm clinical trials.

2015-2018 **Doctoral Researcher** German Cancer Research Center, Heidelberg, Germany
Managed, QC'ed and analyzed oncological databases. Partnered with (inter)national cancer registries which resulted in 14 peer-reviewed publications impacting *public cancer survival surveillance*.

2014-2015 **Research Scholar** University of Florida, Gainesville, FL, USA
Contributed to a multidisciplinary project to develop an EHR-based *predictive risk model* to prevent adverse events among hospitalized patients. The model was implemented in select US hospitals.

Selected Awards & Honors (full list: *janickweberpals.github.io/awards*)

2018 Stephan-Weiland Award (German Society for Epidemiology)
2018 Advancement Award for best Ph.D. thesis in Epidemiology (German Association for Medical Informatics, Biometry and Epidemiology)
2017 Poster Award (Helmholtz International Graduate School)
2016 & 17 International Society for Pharmacoepidemiology (ISPE) Scholarship

Selected Talks & Outreach (full list: *janickweberpals.github.io/talks*)

Issues and Solutions When Estimating Treatment Effects Using US Electronic Health Record Data (invited panelist). International Society for Health Economics and Outcomes Research (ISPOR) Annual Meeting, Boston, MA (2023).

Selected publications (full list: *janickweberpals.github.io/publications*)

Weberpals J, Becker T, Schmich F, Ruettinger D, Theis FJ, Bauer-Mehren A. Deep learning-based propensity scores for confounding control in comparative effectiveness research: a large-scale real-world data study. **Epidemiology** (2021).

Loureiro H, Becker T, Bauer-Mehren A, Ahmidi N, Weberpals J. Artificial Intelligence for Prognostic Scores in Oncology: a benchmarking study. **Frontiers in Artificial Intelligence** (2021).