

Colin Pawlowski, Ph.D.

East Arcade, One Main Street, Cambridge, MA 02142
colinpawlowski78@gmail.com | 910-617-9317

Summary

- Data scientist and AI/ML researcher with 5+ years of experience with expertise in electronic health record (EHR) data, large language models (LLMs), deep learning, and real-world evidence (RWE) studies.
- Proven research track record with 25+ peer-reviewed publications (1500+ citations) leveraging AI-based methods to analyze healthcare data.
- Experience leading high-functioning teams, promoting an active and collaborative research culture, and driving projects to completion.

Education **Massachusetts Institute of Technology**, Cambridge, MA
Ph.D. in Operations Research, June 2019. GPA: 5.0/5.0
Supported by National Science Foundation (NSF) Graduate Research Fellowship.

Yale University, New Haven, CT
B.S. in Mathematics (Intensive), May 2014.
GPA: 3.93/4.00; Magna Cum Laude, Phi Beta Kappa Society.

Experience

2020-2024 **nference**, Cambridge, MA
(Nov-present) *Director & Head of Data Sciences*
Leading RWE studies and client engagements in disease areas including cardiovascular disease, type 2 diabetes, and AL amyloidosis. In addition, contributing to the development of the nSights product for the analysis of de-identified EHR data from multiple health systems including Mayo Clinic, Duke Health, Mercy, Vanderbilt University Medical Center, and Banner Health. Contributing to the development of features including: propensity score matching, Cox proportional hazards models, federated learning, AI-ECG models, and generative AI models to extract information from unstructured clinical notes.

2020 **nference**, Cambridge, MA
(Jul-Nov) *Head of Data Sciences*
Led COVID-19 data science research projects and conducted RWE studies on EHR data from the Mayo Clinic. Co-author on 20+ publications, including the first RWE study confirming the effectiveness of mRNA COVID-19 vaccines in a US-based health system. Work cited by multiple government agencies and organizations including the White House, CDC, and WHO.

2020 **nference**, Cambridge, MA
(Jan-Jun) *Data Science Partnerships Lead*
Contributed to the development of DeepModelBuilder, an nferX NLP software platform. Led client engagements to enable use of the nferX AI platform in research and product development-related projects at life sciences organizations.

- 2019** **Conference**, Cambridge, MA
(Jul-Dec) *Translational Scientist*
Trained NLP models to extract information from unstructured biomedical text.
- 2014–2019** **MIT Operations Research Center**, Cambridge, MA
Research Assistant
Developed fast machine learning algorithms to perform statistical inference on healthcare datasets with missing and uncertain values. Worked on applications in personalized medicine using large-scale EHR and genomic data. Research advisor: Dimitris Bertsimas.
- 2013** **Mount Holyoke College REU**, South Hadley, MA
(Summer) *Undergraduate Researcher*
Researched mathematical modeling and epidemiology. Programmed a population-level model for tuberculosis in the USA, with cost analysis for several intervention strategies.
- 2011-2012** **NASA Flight Opportunities Program**, Houston, TX
Microgravity Research Team Leader
Led a team of 6 students; built a prototype of a 3-D cell culture apparatus and tested it aboard NASA’s zero-gravity plane.

Selected Publications

1. **Pawłowski C**, et. al. “AI-ECG is a Cost-Effective Early Screening Tool for Heart Failure With Reduced Ejection Fraction.” AHA Meeting Abstract, 2023 Nov 6.
2. Bertsimas D, **Pawłowski C**. “Tensor Completion with Noisy Side Information.” Machine Learning, 2023 Aug 7.
3. Silvert E, Hester L, Ramudu E, **Pawłowski C**, et. al. “Identifying signs and symptoms of AL amyloidosis in electronic health records using natural language processing, diagnosis codes, and manually abstracted registry data.” American Journal of Hematology, 2023 July 4.
4. Ip W, **Pawłowski C**, et. al. “Augmented curation of disease diagnoses and medications for patients with hepatocellular carcinoma.” ASCO Meeting Abstract, 2023 May 31.
5. Venkatakrisnan AJ, Anand P, Lenehan P, Ghosh P, Suratekar R, Silvert E, **Pawłowski C**, et. al. “Expanding repertoire of SARS-CoV-2 deletion mutations contributes to evolution of highly transmissible variants.” Scientific Reports, 2023 Jan 5.
6. **Pawłowski C**, et. al. “SARS-CoV-2 and influenza coinfection throughout the COVID-19 pandemic: an assessment of coinfection rates, cohort characteristics, and clinical outcomes.” PNAS Nexus, 2022 Jul 1.
7. Puranik A, Lenehan P, O’Horo JC, **Pawłowski C**, et. al. “Durability analysis of the highly effective BNT162b2 vaccine against COVID-19.” PNAS Nexus, 2022 Jul 1.
8. Niesen M, **Pawłowski C**, et. al. “Surveillance of Safety of 3 Doses of COVID-19 mRNA Vaccination Using Electronic Health Records.” JAMA Network Open, 2022 Apr 1.
9. Razonable R, **Pawłowski C**, et. al. “Casirivimab-Imdevimab treatment is associated with reduced rates of hospitalization among high-risk patients with mild to moderate coronavirus disease-19.” EClinicalMedicine, 2021 Oct 1.
10. Ganesh R, **Pawłowski C**, et. al. “Intravenous bamlanivimab use associates with reduced hospitalization in high-risk patients with mild to moderate COVID-19.” Journal of Clinical Investigation, 2021 Oct 1.

11. **Pawlowski C**, et. al. “FDA-authorized COVID-19 vaccines are effective per real-world evidence synthesized across a multi-state health system.” *Med*, 2021 Aug 13.
12. Venkatakrisnan A, **Pawlowski C**, et. al. “Mapping each pre-existing condition’s association to short-term and long-term COVID-19 complications.” *NPJ Digital Medicine*, 2021 Jul 27.
13. **Pawlowski C**, et. al. “Cerebral Venous Sinus Thrombosis is not Significantly Linked to COVID-19 Vaccines or Non-COVID Vaccines in a Large Multi-State Health System.” *Journal of Stroke and Cerebrovascular Diseases*, 2021 Jun 16.
14. **Pawlowski C**, et. al. “Enoxaparin is associated with lower rates of mortality than unfractionated Heparin in hospitalized COVID-19 patients.” *EClinicalMedicine*, 2021 Mar 9.
15. Kirkup C, **Pawlowski C**, et. al. “Healthcare disparities among anticoagulation therapies for severe COVID-19 patients in the multi-site VIRUS registry.” *J Med Virol.*, 2021 Mar 5.
16. **Pawlowski C**, et. al. “Exploratory analysis of immunization records highlights decreased SARS-CoV-2 rates in individuals with recent non-COVID-19 vaccinations.” *Sci Rep.*, 2021 Feb 26.
17. **Pawlowski C**, et. al. “Inference from longitudinal laboratory tests characterizes temporal evolution of COVID-19-associated coagulopathy (CAC).” *eLife*, 2020 Aug 17.
18. Bertsimas D, **Pawlowski C**, Orfanoudaki A. “Imputation of clinical covariates in time series.” *Machine Learning*, 2020 Nov 10.
19. Bertsimas D, Dunn J, **Pawlowski C**, Zhuo Y. “Robust Classification.” *INFORMS Journal on Optimization*, 2018 Oct 19.
20. Bertsimas D, **Pawlowski C**, Zhuo Y. “From predictive methods to missing data imputation: an optimization approach.” *Journal of Machine Learning Research*, 2018 Apr 1.

Complete list of publications: <https://scholar.google.com/citations?hl=en&user=WESfOysAAAAJ>

Patents

1. **Pawlowski C**, et. al. “System to identify size and location information from unstructured inputs.” US Patent App. 18/166,676, 2023.
2. **Pawlowski C**, et. al. “Identifying patient populations vulnerable to viral infection and methods of inducing heterologous immunity in same.” US Patent App. 17/371,555, 2022.

Honors and Awards

2016	athenahealth Hackathon Grand Prize
2015	NSF Graduate Fellowship
2012	Richter Summer Fellowship
2011	NASA Flight Opportunities Program, national research grant
2011	Connecticut Space Grant Consortium Project Grant

Skills

Programming languages: Python, SQL, R, Julia, JavaScript, MongoDB, Bash

Software packages: DeepSpeed, LangChain, MatchIt, OpenAI, PyTorch, React, Scikit-learn, spaCy, TensorFlow, Tidiverse, XGBoost

Project management: Jira, Confluence

Other skills: Adobe Illustrator, Github Co-pilot, ML Flow