

Zachary Butzin-Dozier, PhD, MPH

ButzinDozier.com • zbutzin@berkeley.edu • (302) 437-6262 • 2121 Berkeley Way, Berkeley, CA, 94720

Updated May 12, 2026

Summary Statement

I am a researcher and instructor in epidemiology and biostatistics with a strong ability to communicate complex methods and findings to a broad range of audiences. My research focuses on applying advanced methods in epidemiology, biostatistics, and causal inference to observational data from electronic health records.

Education

UC Berkeley School of Public Health, Berkeley, CA

Ph.D. in Epidemiology, 2022

GPA: 4.0/4.0

UC Berkeley School of Public Health, Berkeley, CA

Master of Public Health in Epidemiology and Biostatistics, 2020

GPA: 4.0/4.0

Tufts University, Boston, MA

Bachelor of Science in Psychology, 2015

Summa Cum Laude

GPA: 3.8/4.0

Funding

K01 Award: 1K01AI182501 (September 2024 – August 2029). Applying a Targeted Machine Learning and Causal Inference Approach to Analyzing Long-Term Sequelae of COVID-19 Infection Through the National COVID Cohort Collaborative; National Institute of Allergy and Infectious Diseases (NIAID); PI: Zachary Butzin-Dozier.

Gates Foundation subcontract (June 2025 – May 2027). The effects of air pollution and temperature on perinatal outcomes. As a postdoctoral scholar, I support the study design, methodologic considerations, and data analyses of this project. PI: Rakesh Ghosh

My previous research (2020-2024) was financially supported by a global development grant (OPP1165144) from the Bill & Melinda Gates Foundation to the University of California, Berkeley. As a postdoctoral scholar, I supported and led several project activities related to study design, data analysis, and manuscript publication.

Honors and Awards

2024: NIH NCATS Clinical and Translational Sciences Awards spotlight article of the month

2024: UC Berkeley School of Public Health Committee on Teaching Excellence Honoree

2023: Placed third in the NIH Long COVID Computational Challenge

2020: Marcus Foster Education Institute Graduate Fellowship

2019: FLAS Academic Year Fellowship for Portuguese language and Latin American studies

2019: Center for Latin American Studies Field Research Grant

2016: Elected Board Member, The Chain Collaborative

2015: Priscilla N. Dunn Award for achievement in psychology

2015: Graduated Summa Cum Laude

2013: Education Abroad Network Merit Scholarship

2013: University of New South Wales Scholarship

2010: Michael J. Ferguson Scholarship

Research

- **Postdoctoral Scholar**, UC Berkeley School of Public Health, Division of Biostatistics, Center for Targeted Machine Learning (September 2022 – Present)
My work applies a Targeted Machine Learning methodology (applied machine learning for causal inference) to evaluate interventions against Long COVID. I am an NIAID K01 scholar (2025-2029), my work was pivotal in maintaining continued funding from the Bill and Melinda Gates Foundation, I facilitated a Data Use Agreement between N3C and UC Berkeley, I led a team to third place in the NIH Long COVID Computational Challenge, and I forged multiple inter-institutional collaborations (UNC-Chapel Hill, Johns Hopkins, University of Washington, University of Minnesota, University of Alabama at Birmingham Medical Center).
- **Graduate Student Researcher**, UC Berkeley (January 2020 – August 2022)
Worked under Professors Jack Colford, Alan Hubbard, and Mark van der Laan. Applied Targeted Learning methodology to data from the WASH Benefits Bangladesh Project. This work culminated in three first-author publications and several supporting-author publications.
- **Graduate Student Researcher**, UC Berkeley/Universidad San Francisco de Quito (May 2019 – August 2019)
Led a mixed-methods research project on antimicrobial resistance in Ecuador. This work culminated in one first-author publication and two supporting-author publications.

Teaching

- **Instructor**, UC Berkeley School of Public Health, Department of Epidemiology (August 2022 – Present)
Instructor for graduate-level courses, including Introductory Methods in Epidemiology (Spring 2023-present), Intermediate Methods in Epidemiology (Spring 2023), and Intervention Trial Design (Fall 2022).
- **Graduate Student Instructor**, UC Berkeley (January 2019 – December 2021)
Instructed courses in Epidemiologic Analysis, Intermediate Epidemiologic Methods, and Research Methods in Psychology. Developed online learning materials and converted residential materials for online use.

Skills

- **Technical:** R, Python, Camtasia, Jupyter Notebooks, Stata, SPSS, Microsoft Office
- **Languages:** Fluent in Spanish, proficient in Portuguese

Peer-Reviewed Journal Articles (Original Research)

1. **Butzin-Dozier, Z.**, Ji, Y., Deshpande, S., Anzalone, A. J., Wang, L., Hurwitz, E., Kumar, M., Patel, R. C., Sun, J., Robertson, M., Mertens, A. N., Camacho-Rivera, M., Colford, J. M., Hubbard, A. E., Madlock-Brown, C., & the National COVID Cohort Collaborative (N3C) Consortium. (2026). Racial bias in Long COVID diagnosis. Accepted for publication at *Health Affairs*.
2. Chen, B., Mertens, A. N., Lin, C.-H., Tan, S. T., Jamshed, F., Figueroa, D., Hemlock, C., **Butzin-Dozier, Z.***, Fernald, L. C. H., Stewart, C. P., Hubbard, A. E., Rahman, M. Z., Ali, S., Arnold, B. F., Dhabhar, F. S., Granger, D. A., Rahman, M., Luby, S. P., Colford, J. M., Jr., & Lin, A. (in press). Associations between micronutrient status, hormones, and immune status during pregnancy and child growth in rural Bangladesh: A prospective cohort study. *Current Developments in Nutrition*.
 - a. *I conducted a literature review, supported study design and hypothesis formation, supported manuscript drafting and editing, and informed interpretation of results.
3. Wong, R., Wiggen, T., Hall, M., Johnson, S. G., Huling, J. D., Turner, L. E., Wilkins, K. J., Yeh, H.-C., Stürmer, T., Bramante, C., **Butzin-Dozier, Z.***, Buse, J., Reusch, J., & the N3C Consortium. (in press). The effect of COVID-19 on incident diabetes in pediatric patients: Findings from the National COVID-19 Cohort Collaborative (N3C). *Pediatric Diabetes*.
 - a. *I was invited to join this project as an expert in causal inference. I helped contextualize subgroup findings, given the risk of false discoveries from multiple testing. Furthermore, I evaluated potential sources of bias, including unmeasured confounding, and how these biases may influence the observed results. I supported the interpretation of results, manuscript review and editing, and the writing of the discussion.
4. Yang, H., Han, D., Hurwitz, E., Anzalone, J., Varley, C. D., Safo, S. E., Sun, J., Li, D., Camacho-Rivera, M., Fadul, N., **Butzin-Dozier, Z.***, Nosyk, B., Patel, R., & Zang, X. (2025). Racial/ethnic and regional disparities in HIV testing amidst the COVID-19 pandemic: A nationwide retrospective analysis using N3C data. *Lancet HIV*. <https://doi.org/10.2139/ssrn.5273921>
 - a. *Supported data analysis and troubleshooting of analytic pipeline in N3C. Proposed alternative interpretations and potential limitations of study design. Reviewed and revised manuscript draft.
5. Hurwitz, E., Varley, C. D., Anzalone, A. J., Madhira, V., Olex, A. L., Sun, J., Vaidya, D., Fadul, N., Islam, J. Y., Jackson, L. E., Wilkins, K. J., **Butzin-Dozier, Z.***, Li, D., Safo, S. E., McMurry, J. A., Maheria, P., Williams, T., Hassan, S. A., Haendel, M., & Patel, R. C., ... National COVID Cohort Collaborative (N3C) Consortium. (2025). Identifying people living with or those at risk for HIV in a nationally-sampled electronic health record repository called the National Clinical Cohort Collaborative: Computational phenotyping study. *JMIR Medical Informatics*, *13*, e68143.
 - a. *Supported methodologic troubleshooting and data processing. Supported writing and revision of manuscript.
6. Hurwitz, E., Meltzer-Brody, S., **Butzin-Dozier, Z.***, Patel, R. C., Elhadad, N., & Haendel, M. A. (2025). Unlocking the potential of wear time of a wearable device to enhance postpartum depression screening and detection: Cross-sectional study. *JMIR Formative Research*, *9*, e67585. <https://doi.org/10.2196/67585>

- a. *Contributed to study design and causal inference considerations. Reviewed, edited, and provided feedback for the manuscript.
7. **Butzin-Dozier, Z.**, Ji, Y., Coyle, J., Malenica, I., Rogawski McQuade, E. T., Grembi, J. A., Platts-Mills, J. A., Houpt, E. R., Graham, J. P., Ali, S., Rahman, M. Z., Alauddin, M., Famida, S. L., Akther, S., Hossen, Md. S., Mutsuddi, P., Shoab, A. K., Rahman, M., Islam, Md. O., Miah, R., ... Lin, A. (2025). Treatment heterogeneity of water, sanitation, hygiene, and nutrition interventions on child growth by environmental enteric dysfunction and pathogen status for young children in Bangladesh. *PLOS Neglected Tropical Diseases*, *19*(2), e0012881. <https://doi.org/10.1371/journal.pntd.0012881>
8. Figueroa, D., Al Mamun, M. M., Jung, D. K., Li, A., Tan, S. T., Jamshed, F., **Butzin-Dozier, Z.***, Mertens, A. N., Lin, J., Pitchik, H. O., Parvin, K., Silvera, A., Fernald, L. C. H., Arnold, B. F., Ali, S., Shoab, A. K., Famida, S. L., Akther, S., Rahman, M. Z., Hossen, M. S., ... Lin, A. (2025). Maternal experience of intimate partner violence, maternal depression, and parental stress are not associated with child telomere length in Bangladesh. *Scientific Reports*, *15*(1), 8499.
 - a. *Supported data analysis, literature review, and interpretation of findings. Reviewed and revised manuscript draft.
9. Ercumen, A., Mertens, A. N., **Butzin-Dozier, Z.***, Jung, D. K., Ali, S., Achando, B. S., Rao, G., Hemlock, C., Pickering, A. J., Stewart, C. P., Tan, S. T., Grembi, J. A., Benjamin-Chung, J., Wolfe, M., Ho, G. G., Rahman, Md. Z., Arnold, C. D., Dentz, H. N., Njenga, S. M., Meerkerk, T., ... Lin, A. (2025). Water, sanitation, handwashing, and nutritional interventions can reduce child antibiotic use: Evidence from Bangladesh and Kenya. *Nature Communications*, *16*(1), 556. <https://doi.org/10.1038/s41467-024-55801-x>
 - a. *Generated all manuscript figures and tables. Reviewed and revised manuscript.
10. Soff, S., Yoo, Y. J., Bramante, C., Reusch, J. E. B., Huling, J. D., Hall, M. A., Brannock, D., Sturmer, T., **Butzin-Dozier, Z.***, Wong, R., & Moffitt, R. & N3C Consortium (2025). Association of glycemic control with Long COVID in patients with type 2 diabetes: Findings from the National COVID Cohort Collaborative (N3C). *BMJ Open Diabetes Research & Care*, *13*(1), e004536. <https://doi.org/10.1136/bmjdr-2024-004536>
 - a. *Reviewed, edited, and provided feedback for manuscript draft. Discussed interpretation with first author and shared additional limitations.
11. **Butzin-Dozier, Z.**, Ji, Y., Deshpande, S., Hurwitz, E., Anzalone, A. J., Coyle, J., Shi, J., Mertens, A., van der Laan, M. J., Colford, J. M., Jr, Patel, R. C., Hubbard, A. E., & the National COVID Cohort Collaborative (N3C) Consortium. (2024). SSRI use during acute COVID-19 and risk of Long COVID among patients with depression. *BMC Medicine*, *22*(1), 445. <https://doi.org/10.1186/s12916-024-03655-x>
12. Bergquist, T., Loomba, J., Pfaff, E., Xia, F., Zhao, Z., Zhu, Y., Mitchell, E., Bhattacharya, B., Shetty, G., Munia, T., Delong, G., Tariq, A., **Butzin-Dozier, Z.***, Ji, Y., Li, H., Coyle, J., Shi, S., Philips, R. V., Mertens, A., Pirracchio, R., ... N3C Consortium. (2024). Crowd-sourced machine learning prediction of long COVID using data from the National COVID Cohort Collaborative. *eBioMedicine*, *108*, 105333. <https://doi.org/10.1016/j.ebiom.2024.105333>
 - a. *Reviewed, edited, and provided feedback for manuscript draft. Led award-winning team entry, which was featured in this manuscript.

13. **Butzin-Dozier, Z.**, Mertens, A. N., Tan, S. T., Granger, D. A., Pitchik, H. O., Il'yasova, D., Tofail, F., Rahman, M. Z., Spasojevic, I., Shalev, I., Ali, S., Karim, M. R., Shahriar, S., Famida, S. L., Shuman, G., Shoab, A. K., Akther, S., Hossen, M. S., Mutsuddi, P., Rahman, M.,... Lin, A. (2024). Stress biomarkers and child development in young children in Bangladesh. *Psychoneuroendocrinology*, *164*, 107023. <https://doi.org/10.1016/j.psyneuen.2024.107023>
14. Hurwitz, E., **Butzin-Dozier, Z.***, Master, H., O'Neil, S. T., Walden, A., Holko, M., Patel, R. C., & Haendel, M. A. (2024). Harnessing consumer wearable digital biomarkers for individualized recognition of postpartum depression using the All of Us Research Program data set: Cross-sectional Study. *JMIR Mhealth Uhealth*, *12*, e54622. <https://doi.org/10.2196/54622>
 - a. Designed analysis and supported troubleshooting of analytic pipeline. Supported writing of first-draft manuscript and reviewed and edited final manuscript version.
15. **Butzin-Dozier, Z.**, Ji, Y., Li, H., Coyle, J., Shi, J. (Seraphina), Phillips, R. V., Mertens, A., Pirracchio, R., Van Der Laan, M. J., Patel, R. C., Colford, J. M., & Hubbard, A. E; National COVID Cohort Collaborative (N3C) Consortium (2024). Predicting Long COVID in the National COVID Cohort Collaborative using Super Learner: Cohort study. *JMIR Public Health and Surveillance*, *10*, e53322. <https://doi.org/10.2196/53322>
16. Waters, W. F., Baca, M., Graham, J. P., **Butzin-Dozier, Z.***, & Vinueza, L. (2022). Antibiotic use by backyard food animal producers in Ecuador: A qualitative study. *BMC Public Health*, *22*(1), 685. <https://doi.org/10.1186/s12889-022-13073-4>
 - a. Conducted study design, data collection, and portions of data analysis. Wrote first draft of manuscript sections. Reviewed, edited, and provided feedback on final manuscript draft.
17. **Butzin-Dozier, Z.**, Waters, W. F., Baca, M., Vinueza, R. L., Saraiva-Garcia, C., & Graham, J. (2020). Assessing upstream determinants of antibiotic use in small-scale food animal production through a simulated client method. *Antibiotics (Basel, Switzerland)*, *10*(1), 2. <https://doi.org/10.3390/antibiotics10010002>
18. Guhan, M., Kumar, M., **Butzin-Dozier, Z.***, & Graham, J. (2021). Challenges and opportunities for scaling up infection prevention and control programmes in rural district hospitals of Tamil Nadu, India. *International Journal of Infection Control*, *17*(1). <https://doi.org/10.3396/ijic.v17.20628>
 - a. Mentored undergraduate student in manuscript preparation. Reviewed, edited, and provided feedback on manuscript draft. Iterated on overall manuscript format and focus.
19. Bernard, K., **Butzin-Dozier, Z.***, Rittenhouse, J., & Dozier, M. (2010). Cortisol production patterns in young children living with birth parents vs children placed in foster care following involvement of Child Protective Services. *Archives of Pediatrics & Adolescent Medicine*, *164*(5), 438–443. <https://doi.org/10.1001/archpediatrics.2010.54>
 - a. Designed study and conducted data analysis. Reviewed, edited, and provided feedback for manuscript draft.

Peer-Reviewed Journal Article – Other (Systematic Review)

1. **Butzin-Dozier, Z.**, Athni, T. S., & Benjamin-Chung, J. (2022). A review of the ring trial design for evaluating ring interventions for infectious diseases. *Epidemiologic Reviews*, 44(1), 29–54. <https://doi.org/10.1093/epirev/mxac003>

Peer-Reviewed Conference Paper

1. McCoy D, Wu Y, **Butzin-Dozier Z.** AI Progress Should Be Measured by Capability-Per-Resource, Not Scale Alone: A Framework for Gradient-Guided Resource Allocation in LLMs. In: *The Thirty-Ninth Annual Conference on Neural Information Processing Systems (NEURIPS) Position Paper Track*. 2025. <https://openreview.net/forum?id=6plSmhBI33>

Manuscripts Currently Under Peer Review

1. **Butzin-Dozier, Z.**, Ji, Y., Wang, L., Anzalone, A. J., Coyle, J., Phillips, R. V., Patel, R. C., Sun, J., Hurwitz, E., Deshpande, S., Shi, J. (Seraphina), Mertens, A. N., van der Laan, M., Colford, J. M., Hubbard, A. E., & on behalf of the National Clinical Cohort Collaborative. (2025). COVID-19 Vaccination Timing, Relative to Acute COVID-19, and Subsequent Risk of Long COVID. Under review at *eBioMedicine*.
2. **Butzin-Dozier, Z.**, Ji, Y., Wang, L.-C., Anzalone, A. J., Hurwitz, E., Patel, R. C., van der Laan, M. J., Colford, J. M., Hubbard, A. E., & the N3C Consortium. (2025). Causal Inference via Electronic Health Records in the National Clinical Cohort Collaborative: Challenges and Solutions in Long COVID Research. Under review at *BMC Medical Research Methodology*.
3. **Butzin-Dozier, Z.**, Qiu, S., Hubbard, A. E., Shi, J. (Seraphina), & van der Laan, M. J. (2024). Highly adaptive LASSO: Machine learning that provides valid nonparametric inference in realistic models. Under review at the *Journal of Machine Learning Research*.
4. Nkodo, E. N. E., Maheria, P., Hurwitz, E., Anzalone, A., Li, D., Islam, J., Sun, J., Varley, C., **Butzin-Dozier, Z.***, Safo, S., Kirksey, K., Hassan, S., Camacho-Rivera, M., Patel, R., & Fadul, N. (2025). HIV status and COVID-19 treatment disparities in the US National Clinical Cohort Collaborative. Under review at *Open Forum Infectious Diseases*.
 - a. *Supported study design, causal inference considerations, incorporation of heterogeneous healthcare utilization, interpretation of results, and manuscript editing.

Non Peer-Reviewed Publications

1. Auerswald, C., **Butzin-Dozier, Z.**, & et al. (2020). For the good of us all: Addressing the needs of our unhoused neighbors during the COVID-19 pandemic. *UC Berkeley School of Public Health*. <https://bph-storage.s3.us-west-1.amazonaws.com/wp-content/uploads/2020/04/For-the-Good-of-Us-All-Report.pdf>
 - a. Provided epidemiologic and methodologic expertise. Wrote initial drafts of sections that described potential epidemiologic consequences of inadequate policy during the COVID-19 pandemic.

Trainees

Current

Yunwen Ji (Biostatistics PhD student)
Eric Hurwitz (Health Informatics postdoc)
Manav Kumar (undergraduate)
Sam (Lin-Chiun) Wang (undergraduate)
Ariana Budhihartanto (undergraduate)

Former

Seraphina Shi (Biostatistics PhD student)
Haodong Li (Biostatistics PhD student)
Sky Qiu (Biostatistics PhD student)
Sarang Deshpande (undergraduate)
Alexis Silvera (undergraduate)
Farheen Jamshed (undergraduate)
Da Kyung Jung (undergraduate)
Ewen Huang (undergraduate)

Presentations

Invited presentations

1. “Causal Inference via Electronic Health Records in N3C,” Center for Targeted Machine Learning seminar series, September 17, 2025 (Berkeley, CA).
2. “GLP-1 Receptor Agonists and Mortality/Long COVID in Type 2 Diabetes Patients,” Diabetes and Obesity Domain Team, National Clinical Cohort Collaborative, July 18, 2025 (remote via Zoom).
3. “Targeted Learning in Health Services Research,” International Society of Pharmacoepidemiology (ISPE) University of Maryland, Baltimore (UMB), March 14, 2025 (remote via Zoom).
4. “Racial bias in Long COVID diagnosis,” National Clinical Cohort Collaborative Immunosuppression Domain Team, October 7, 2024 (remote via Zoom).
5. “Evaluating preventive interventions against Long COVID through Targeted Learning,” UC Santa Cruz Microbiology and Environmental Toxicology Seminar Series, February 8, 2024 (Santa Cruz, CA).
6. “Evaluating preventive interventions against Long COVID,” UC Berkeley Center for Targeted Machine Learning Seminar Series, April 9, 2024 (Berkeley, CA).

Teaching presentations

1. “Power and Sample Size Guest Lecture,” UC Berkeley Public Health, Epidemiologic Methods (graduate-level course), February 28, 2024 (Berkeley, CA).
2. “Adaptive Trials Guest Lecture, UC Berkeley Public Health,” Intervention Trial Design (graduate-level course), February 28, 2024 (Berkeley, CA).
3. “WASH Benefits Guest Lecture,” UC Berkeley Public Health, Introduction to Public Health guest lecture (undergraduate-level course), February 14, 2024 (Berkeley, CA).

Media Coverage

1. Study shows use of SSRIs may protect against long COVID. Sheila Kaplan, December 6, 2024. University of California, Berkeley, School of Public Health. Retrieved February 13, 2025,

from <https://publichealth.berkeley.edu/news-media/research-highlights/ssris-may-protect-against-long-covid>

2. [NIH NCATS Clinical and Translational Sciences Awards spotlight article of the month](#), September 10, 2024. Featured article: "[Predicting Long COVID...](#)" (JMIR *Public Health and Surveillance*, 2024)
3. Postdoc Spotlight: Zachary Butzin-Dozier Sheds Light on Long COVID with Advanced Machine Learning. Simmerdeep Kaur, June 24, 2024. *Berkeley Visiting Scholars and Postdoc Affairs*. <https://vspa.berkeley.edu/news/postdoc-spotlight-zachary-butzin-dozier-sheds-light-long-covid-advanced-machine-learning>
4. UC Berkeley research teams wins [redacted] prize for long COVID-19 prediction model. Sheila Kaplan, April 11, 2023. *University of California, Berkeley, School of Public Health*. <https://publichealth.berkeley.edu/covid-19/berkeley-research-teams-wins-prize-for-long-covid-prediction-model>

Peer Review Activities

1. Social Sciences & Medicine, January 25, 2025
2. Rapid Reviews for Infectious Diseases, October 1, 2024
3. Cochrane Review Infectious Diseases Group, September 30, 2021