

Zachary B. Dozier, PhD, MPH

zdozier.com · zdozier@stanford.edu · 3145 Porter Drive, Palo Alto, CA 94304

Formerly published as Zachary Butzin-Dozier

Updated October 12, 2026. Generated automatically from ORCID, OpenAlex, Crossref, Europe PMC, and NIH RePORTER.

Summary Statement

I am an Assistant Professor at Stanford School of Medicine (Department of Pediatrics, Division of Clinical Informatics; Department of Medicine, Division of Computational Medicine) with a strong ability to communicate complex methods and findings to a broad range of audiences. My research applies machine learning for causal inference via electronic health record (EHR) data. This work is supported by an NIAID K01 award and primarily uses data from Epic Cosmos, PEDSnet, the UC Health Data Warehouse, and the National Clinical Cohort Collaborative.

Academic Appointments

Assistant Professor, Stanford University School of Medicine

June 2026 – Present

Department of Pediatrics (Clinical Informatics) and Department of Medicine (Computational Medicine / Center for Biomedical Informatics Research)

I lead an independent research program applying artificial intelligence and machine learning for causal inference via electronic health record data to improve pediatric health. As an NIAID K01 scholar, I am building a lab with access to Epic COSMOS, PEDSnet, PCORnet, UCHDW, and N3C.

Postdoctoral Scholar, UC Berkeley School of Public Health

September 2022 – May 2026

Division of Biostatistics, Center for Targeted Machine Learning

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.

Education

UC Berkeley School of Public Health, Berkeley, CA

2022

Ph.D. in Epidemiology

GPA: 4.0/4.0

UC Berkeley School of Public Health, Berkeley, CA
Master of Public Health in Epidemiology and Biostatistics
GPA: 4.0/4.0

2020

Tufts University, Boston, MA
Bachelor of Science in Psychology
Summa Cum Laude
GPA: 3.8/4.0

2015

Funding

Applying a Machine Learning and Causal Inference Approach to Analyzing Long-Term Sequelae of Viral Infection

September 2024 – August 2029

National Institute of Allergy and Infectious Diseases (NIAID), NIH, K01 (K01AI182501). Role: PI.

Applies targeted machine learning and causal inference methods to data from the National COVID Cohort Collaborative (N3C) to study the long-term sequelae of COVID-19 infection.

The effects of air pollution and temperature on perinatal outcomes

June 2025 – June 2026

Gates Foundation, Subcontract. Role: Postdoctoral scholar; PI: Rakesh Ghosh.

As a postdoctoral scholar, I support the study design, methodologic considerations, and data analyses of this project.

Global development grant (OPP1165144) to the University of California, Berkeley

2020–2024

Bill & Melinda Gates Foundation, Global development grant. Role: Postdoctoral scholar.

My previous research (2020-2024) was financially supported by this grant. 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

Publications

Peer-Reviewed Journal Articles (Original Research)

1. Wang, Y., Sabit, A., Anzalone, J., Li, D., Hurwitz, E., Fadul, N., Wei, S., Olex, A. L., Safo, S. E., Islam, J. Y., Brown, T. T., Chute, C. G., Haendel, M. A., Varley, C. D., **Butzin-Dozier, Z.***, Camacho-Rivera, M., Nkodo, E. N. E., Patel, R. C., Sun, J., & National Clinical Cohort Collaborative (N3C). (2026). Increasing Multimorbidity Burden After COVID-19 Infection Among People With and Without HIV in the United States: Insights From the National Clinical Cohort Collaborative (N3C). *Open Forum Infectious Diseases*, 13(9), ofag560. <https://doi.org/10.1093/ofid/ofag560>
*Supported causal inference considerations in the study design, analysis, and interpretation. I reviewed the manuscript and provided written feedback. I also provided insights regarding addressing heterogeneous healthcare utilization and COVID-19 sequelae in N3C.
2. **Butzin-Dozier, Z.**, Ji, Y., Wang, L.-C., Anzalone, A. J., Coyle, J., Phillips, R. V., Patel, R. C., Sun, J., Hurwitz, E., Deshpande, S., Shi, J., Mertens, A., van der Laan, M. J., Colford, J. M., Jr., Hubbard, A. E., & National Clinical Cohort Collaborative (N3C) Consortium. (2026). COVID-19 vaccination timing, relative to acute COVID-19, and subsequent risk of long COVID. *eBioMedicine*, 130, 106382. <https://doi.org/10.1016/j.ebiom.2026.106382>
3. **Butzin-Dozier, Z.**, Ji, Y., Deshpande, S., Anzalone, A. J., Wang, L.-C., Hurwitz, E., Kumar, M., Patel, R. C., Sun, J., Robertson, M., Mertens, A., Camacho-Rivera, M., Colford, J. M., Hubbard, A. E., Madlock-Brown, C., & the National COVID Cohort Collaborative (N3C) Consortium. (2026). Black Patients Underdiagnosed With Long COVID In The US Compared With White Patients. *Health Affairs*, 45(7), 775–781. <https://doi.org/10.1377/hlthaff.2025.00179>
4. Ellenbach, J. C. S., Martin-Carrillo, A., Rahman, M. Z., Ali, S., Grembi, J. A., Jannat, K., Arnold, C. D., Byrd, K., **Butzin-Dozier, Z.**, Alauddin, M., Famida, S. L., Hossen, M. S., Akther, S., Mutsuddi, P., Shoab, A. K., Hussain, Z., Rahman, M., Unicomb, L., Alauddin, S. T., ... Lin, A. (2026). Associations Between Diet and Environmental Enteric Dysfunction in Children Living in Rural Bangladesh. *Current Developments in Nutrition*, 10, 108692. <https://doi.org/10.1016/j.cdnut.2026.108692>
5. 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. C., Zang, X., & Chute, C. G. (2026). Racial, ethnic, and regional disparities in HIV testing during the COVID-19 pandemic in the USA: a nationwide, retrospective, observational study using National Clinical Cohort Collaborative data. *The Lancet HIV*, 13(3), e185–e193. [https://doi.org/10.1016/s2352-3018\(25\)00302-9](https://doi.org/10.1016/s2352-3018(25)00302-9)
*Supported data analysis and troubleshooting of analytic pipeline in N3C. Proposed alternative interpretations and potential limitations of study design. Reviewed and revised manuscript draft.
6. Essam Nkodo, E. N., Maheria, P., Hurwitz, E., Anzalone, A. J., Li, D., Islam, J. Y., Sun, J., Varley, C. D., **Butzin-Dozier, Z.***, Safo, S. E., Kirksey, K., Hassan, S. A., Camacho-Rivera, M., Patel, R. C., & Fadul, N. (2025). HIV Status and COVID-19 Treatment Disparities in the US National Clinical Cohort Collaborative. *Open Forum Infectious Diseases*, 13(1), ofaf731. <https://doi.org/10.1093/ofid/ofaf731>
*Supported study design, causal inference considerations, incorporation of heterogeneous healthcare utilization, interpretation of results, and manuscript editing.
7. 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. (2025). Associations Between Micronutrient Status, Hormones, and Immune Status During Pregnancy and Child Growth in Rural Bangladesh: A Prospective Cohort Study. *Current Developments in Nutrition*, 9(12), 107596. <https://doi.org/10.1016/j.cdnut.2025.107596>
*I conducted a literature review, supported study design and hypothesis formation, supported manuscript drafting and editing, and informed interpretation of results.

8. Hurwitz, E., Varley, C. D., Anzolone, 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. A., ... The National Clinical 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. <https://doi.org/10.2196/68143>
 *Supported methodologic troubleshooting and data processing. Supported writing and revision of manuscript.
9. 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>
 *Contributed to study design and causal inference considerations. Reviewed, edited, and provided feedback for the manuscript.
10. Figueroa, D., Al Mamun, M. M., Jung, D. K., Li, G., 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., ... 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. <https://doi.org/10.1038/s41598-025-90505-2>
 *Supported data analysis, literature review, and interpretation of findings. Reviewed and revised manuscript draft.
11. **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, M. S., Mutsuddi, P., Shoab, A. K., Rahman, M., Islam, M. O., ... 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>
12. 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>
 *Reviewed, edited, and provided feedback for manuscript draft. Discussed interpretation with first author and shared additional limitations.
13. 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, M. Z., Arnold, C. D., Dentz, H. N., Njenga, S. M., ... 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>
 *Generated all manuscript figures and tables. Reviewed and revised manuscript.
14. Wong, R., Wiggen, T., Hall, M. A., Johnson, S. G., Huling, J. D., Turner, L. E., Wilkins, K. J., Yeh, H.-C., Stürmer, T., Bramante, C. T., **Butzin-Dozier, Z.***, Buse, J. B., Reusch, J., & N3C Consortium. (2025). The Effect of COVID-19 on Incident Diabetes in Pediatric Patients: Findings From the National COVID-19 Cohort Collaborative (N3C). *Pediatric Diabetes*, 2025(1), 3545727. <https://doi.org/10.1155/pedi/3545727>
 *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.

15. **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>
16. 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., ... 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>
 *Reviewed, edited, and provided feedback for manuscript draft. Led award-winning team entry, which was featured in this manuscript.
17. **Butzin-Dozier, Z.**, Ji, Y., Li, H., Coyle, J., Shi, J., Phillips, R. V., Mertens, A. N., Pirracchio, R., van der Laan, M. J., Patel, R. C., Colford, J. M., Hubbard, A. E., & The 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>
18. 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 and uHealth*, 12, e54622. <https://doi.org/10.2196/54622>
 *Designed analysis and supported troubleshooting of analytic pipeline. Supported writing of first-draft manuscript and reviewed and edited final manuscript version.
19. **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., ... 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>
20. 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>
 *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.
21. 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>
 *Mentored undergraduate student in manuscript preparation. Reviewed, edited, and provided feedback on manuscript draft. Iterated on overall manuscript format and focus.
22. **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*, 10(1), 2. <https://doi.org/10.3390/antibiotics10010002>
23. 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>
 *Designed study and conducted data analysis. Reviewed, edited, and provided feedback for manuscript draft.

Peer-Reviewed Review Articles

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 Papers

1. McCoy, D., Wu, Y., & **Butzin-Dozier, Z.** (2025). AI Progress Should Be Measured by Capability-Per-Resource, Not Scale Alone: A Framework for Gradient-Guided Resource Allocation in LLMs. In *Advances in Neural Information Processing Systems 38* (pp. 154476–154494). <https://doi.org/10.52202/085713-4650>

Accepted and In Press

1. Singh, N., Anzalone, A. J., Stevenson, H., Robinson, L., Jackson, L. E., Safo, S. E., **Butzin-Dozier, Z.***, Michaud, K., Danila, M. I., Singh, J. A., & Patel, R. C. (in press). Higher risk of Long COVID among patients with autoimmune rheumatic diseases and immunomodulatory therapy: A retrospective cohort analysis of national US data. *Arthritis & Rheumatology*.
*I provided subject-area and methodologic expertise regarding Long COVID, causal inference, and N3C. I provided feedback on the study design, and I gave written feedback on the manuscript draft.
2. **Butzin-Dozier, Z.**, Wang, L.-C., Ji, Y., Anzalone, A. J., Olawore, O., Hafen, R., Hurwitz, E., Kumar, M., Patel, R. C., Budhihartanto, A., van der Laan, M., Colford, J. M., Jr., Hubbard, A. E., Buse, J. B., Johnson, S., Reusch, J., Chan, L. E., Moffitt, R., Wong, R., ... the National Clinical Cohort Collaborative (N3C) Consortium. (2026). Metformin and Severe Post-COVID-19 Outcomes Among Individuals with Diabetes Mellitus. Accepted for publication in *Diabetes, Obesity and Metabolism*. Preprint: *medRxiv*. <https://doi.org/10.64898/2026.07.06.26357398>

Preprints and Manuscripts Under Review

1. **Butzin-Dozier, Z.**, Wang, L.-C., Budhihartanto, A., Bramante, C., & Wong, R. (2026). Long COVID: A critical review of U.S. health, science, and policy. Invited review, *Annual Review of Public Health*.
2. Hurwitz, E., Merchant, L., Flatt, R. E., Reed, K. K., **Butzin-Dozier, Z.**, Haendel, M. A., Thornton, L. M., Bulik, C. M., & Presskreischer, R. (2026). Precision psychological markers enable targeted treatment in digital eating disorder interventions. *medRxiv*. <https://doi.org/10.64898/2026.07.21.26358594>
3. **Butzin-Dozier, Z.**, Ji, Y., Wang, L.-C., Kumar, M., Anzalone, A. J., Hurwitz, E., Budhihartanto, A., Patel, R. C., Hubbard, A., Halpern, J., & the National Clinical Cohort Collaborative (N3C) Consortium. (2026). SSRI prescription during acute COVID-19 and risk of Long COVID symptoms and conditions among patients with depression. *medRxiv*. <https://doi.org/10.64898/2026.07.06.26357401>
4. **Butzin-Dozier, Z.**, Wang, L.-C., Ji, Y., Kumar, M., Anzalone, A. J., Hurwitz, E., Patel, R. C., Budhihartanto, A., Buse, J. B., Johnson, S., Bramante, C., Wong, R., & the National Clinical Cohort Collaborative (N3C) Consortium. (2026). GLP Medications and Severe Post-COVID-19 Outcomes Among Individuals with Type 2 Diabetes Mellitus. *medRxiv*. <https://doi.org/10.64898/2026.07.03.26357246>
5. Byer, B. K., **Butzin-Dozier, Z.**, McGrath, B. M., Muenzer, J., Clarke, L., Haendel, M. A., & O'Neil, S. T. (2026). Differential COVID-19 Outcomes Across Lysosomal Disorders. *medRxiv*. <https://doi.org/10.64898/2026.06.22.26356274>
6. Han, G. Y., Kalogeropoulos, A. P., **Butzin-Dozier, Z.**, Wong, R., & Wang, F. (2026). Causal Machine Learning for Comparative Effectiveness of GLP-1 RA versus SGLT2i in Heart Failure Using Real-World EHR Data. *medRxiv*. <https://doi.org/10.64898/2026.04.06.26350259>
7. Hurwitz, E., Connelly, E., Sklerov, M., Master, H., Hochheiser, H., **Butzin-Dozier, Z.**, Dunn, J., & Haendel, M. A. (2026). Population differences in wearable device wear time: Rescuing data to address biases and advance health equity. *medRxiv*. <https://doi.org/10.64898/2026.03.06.26347799>

8. **Butzin-Dozier, Z.**, Kumar, M., Ji, Y., Wang, L.-C., Anzalone, A. J., Hurwitz, E., Patel, R. C., Wong, R., Bramante, C., Sines, B., & the National Clinical Cohort Collaborative (N3C) Consortium. (2026). IL-6 Receptor Antagonists and Severe Post-COVID-19 Outcomes: An Emulated Target Trial. *medRxiv*. <https://doi.org/10.64898/2026.02.27.26347274>
9. **Butzin-Dozier, Z.**, Ji, Y., Wang, L.-C., Anzalone, A. J., Hurwitz, E., Patel, R. C., van der Laan, M. J., Colford, J. M., Jr., 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. *medRxiv*. <https://doi.org/10.1101/2025.06.06.25329168>
10. **Butzin-Dozier, Z.**, Qiu, S., Hubbard, A. E., Shi, J., & van der Laan, M. J. (2024). Highly adaptive LASSO: Machine learning that provides valid nonparametric inference in realistic models. *medRxiv*. <https://doi.org/10.1101/2024.10.18.24315778>
11. Lin, A., Mertens, A. N., Tan, S., Rahman, M. Z., Hester, L., Kim, L., Arnold, B. F., Karim, M. R., Shahriar, S., Ali, S., Shoab, A. K., Hossen, M. S., Mutsuddi, P., Famida, S. L., Akther, S., Rahman, M., Unicomb, L., **Butzin-Dozier, Z.**, Hemlock, C., ... Dhabhar, F. S. (2021). Effects of drinking water, sanitation, handwashing, and nutritional interventions on immune status in young children: a cluster-randomized controlled trial in rural Bangladesh. *medRxiv*. <https://doi.org/10.1101/2021.11.10.21266206>

Book Chapters

1. **Butzin-Dozier, Z.**, & Anzalone, J. A. (2027). Strengthening causal inference of real-world data. In J. Castner (Ed.), *Environmental Health Nursing: Research, Evidence, and Impact*. Wiley. Forthcoming.

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>

*Provided epidemiologic and methodologic expertise. Wrote initial drafts of sections that described potential epidemiologic consequences of inadequate policy during the COVID-19 pandemic.

Teaching

Instructor, UC Berkeley School of Public Health, Department of Epidemiology August 2022 – June 2025
Instructor for graduate-level courses, including Introductory Methods in Epidemiology (Spring 2023–June 2025), 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.

Trainees

Current

Anna Krasnova (postdoc)
Peter Ju (Data analyst)
Eric Hurwitz (Health Informatics postdoc)
Ariana Budhihartanto (undergraduate)
Manav Kumar (undergraduate)
Sam (Lin-Chiun) Wang (undergraduate)

Former

Yunwen Ji (Biostatistics PhD student)
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*. <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. *NIH NCATS*. Featured article: "Predicting Long COVID..." (JMIR Public Health and Surveillance, 2024). <https://ncats.nih.gov/news-events/news/n3c-study-shows-health-factors-covid-19-infection-may-predict-long-covid-risk>

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 team wins 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. Stanford University Kidney Health Initiative Seed Grant program, May 2026
2. Social Science & Medicine, January 25, 2025
3. Rapid Reviews for Infectious Diseases, October 1, 2024
4. Cochrane Review Infectious Diseases Group, September 30, 2021

Skills

Technical: R, Python, Camtasia, Jupyter Notebooks, Stata, SPSS, Microsoft Office

Languages: Fluent in Spanish, proficient in Portuguese