

NIH BIOGRAPHICAL SKETCH COMMON FORM

Name: Butzin-Dozier, Zachary

Persistent Identifier (PID) of the Senior/Key Person: <https://orcid.org/0000-0001-6419-0008>

Position Title: Assistant Professor

Organization and Location: Stanford University School of Medicine, Department of Pediatrics, Division of Clinical Informatics, Stanford, California, United States

PROFESSIONAL PREPARATION

INSTITUTION AND LOCATION	DEGREE	Start Date	Completion Date	FIELD OF STUDY
UC Berkeley, Berkeley, California, United States	Postdoctoral Fellow	09/2022	05/2026	Biostatistics
UC Berkeley, Berkeley, California, United States	DOCTOR OF PHILOSOPHY	09/2020	08/2022	Epidemiology
UC Berkeley, Berkeley, California, United States	MASTER OF PUBLIC HEALTH	09/2018	05/2020	Epidemiology and Biostatistics
Tufts University, Medford, Massachusetts, United States	BACHELOR OF SCIENCE	09/2011	05/2015	Psychology

Appointments and Positions

2026 - present	Assistant Professor, Stanford University School of Medicine, Department of Pediatrics, Division of Clinical Informatics, Stanford, California, United States
2026 - present	Assistant Professor, Stanford University School of Medicine, Department of Medicine, Division of Computational Medicine, Stanford, California, United States
2022 - present	Instructor, UC Berkeley School of Public Health, Department of Epidemiology, Berkeley, CA, United States
2019 - 2022	Graduate Student Researcher, UC Berkeley School of Public Health, Berkeley, CA, United States
2019 - 2021	Graduate Student Instructor, UC Berkeley School of Public Health, Berkeley, CA, United States

Products**Products Closely Related to the Proposed Project**

- Butzin-Dozier Z, Ji Y, Wang LC, Anzalone AJ, Hurwitz E, Patel RC, van der Laan MJ, Colford JM Jr, Hubbard AE. Causal Inference via Electronic Health Records in the National Clinical Cohort Collaborative: Challenges and Solutions in Long COVID Research. medRxiv. 2025 Jun 11; PubMed Central PMCID: [PMC12155030](https://pubmed.ncbi.nlm.nih.gov/PMC12155030/).
- Butzin-Dozier Z, Ji Y, Wang LC, Anzalone AJ, Coyle J, Phillips RV, Patel RC, Sun J, Hurwitz E, Deshpande S, Shi JS, Mertens A, van der Laan MJ, Colford JM Jr, Hubbard AE. COVID-19 Vaccination Timing, Relative to Acute COVID-19, and Subsequent Risk of Long COVID. medRxiv. 2025 Apr 23; PubMed Central PMCID: [PMC12045423](https://pubmed.ncbi.nlm.nih.gov/PMC12045423/).
- Butzin-Dozier Z, Ji Y, Coyle J, Malenica I, Rogawski McQuade ET, Grembi JA, Platts-Mills JA, Houghton ER, Graham JP, Ali S, Rahman MZ, Alauddin M, Famida SL, Akther S, Hossen MS, Mutsuddi P, Shoab AK, Rahman M, Islam MO, Miah R, Taniuchi M, Liu J, Alauddin ST, Stewart CP, Luby SP, Colford JM Jr, Hubbard AE, Mertens AN, Lin A. 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 Negl Trop Dis. 2025 Feb;19(2):e0012881. PubMed Central PMCID: [PMC11882089](https://pubmed.ncbi.nlm.nih.gov/PMC11882089/).
- Butzin-Dozier Z, Qiu S, Hubbard AE, Shi JS, van der Laan MJ. HIGHLY ADAPTIVE LASSO: MACHINE LEARNING THAT PROVIDES VALID NONPARAMETRIC INFERENCE IN REALISTIC MODELS. medRxiv. 2024 Oct 19; PubMed Central PMCID: [PMC11527044](https://pubmed.ncbi.nlm.nih.gov/PMC11527044/).
- Butzin-Dozier Z, Ji Y, Deshpande S, Hurwitz E, Anzalone AJ, Coyle J, Shi J, Mertens A, van der Laan MJ, Colford JM Jr, Patel RC, Hubbard AE. SSRI use during acute COVID-19 and risk of long COVID among patients with depression. BMC

Med. 2024 Oct 8;22(1):445. PubMed Central PMCID: [PMC11462648](#).

Other Significant Products Highlighting Contributions to Science

1. Ercumen A, Mertens AN, Butzin-Dozier Z, Jung DK, Ali S, Achando BS, Rao G, Hemlock C, Pickering AJ, Stewart CP, Tan ST, Grembi JA, Benjamin-Chung J, Wolfe M, Ho GG, Rahman MZ, Arnold CD, Dentz HN, Njenga SM, Meerkerk T, Chen B, Nadimpalli M, Islam MA, Hubbard AE, Null C, Unicomb L, Rahman M, Colford JM Jr, Luby SP, Arnold BF, Lin A. Water, sanitation, handwashing, and nutritional interventions can reduce child antibiotic use: evidence from Bangladesh and Kenya. *Nat Commun.* 2025 Jan 9;16(1):556. PubMed Central PMCID: [PMC11718192](#).
2. Bernard K, Butzin-Dozier Z, Rittenhouse J, Dozier M. Cortisol production patterns in young children living with birth parents vs children placed in foster care following involvement of Child Protective Services. *Arch Pediatr Adolesc Med.* 2010 May;164(5):438-43. PubMed Central PMCID: [PMC3213033](#).
3. Butzin-Dozier Z, Waters WF, Baca M, Vinueza RL, Saraiva-Garcia C, Graham J. Assessing Upstream Determinants of Antibiotic Use in Small-Scale Food Animal Production through a Simulated Client Method. *Antibiotics (Basel).* 2020 Dec 23;10(1) PubMed Central PMCID: [PMC7822171](#).
4. Butzin-Dozier Z, Athni TS, Benjamin-Chung J. A Review of the Ring Trial Design for Evaluating Ring Interventions for Infectious Diseases. *Epidemiol Rev.* 2022 Dec 21;44(1):29-54. PubMed Central PMCID: [PMC10362935](#).
5. Butzin-Dozier Z, Ji Y, Li H, Coyle J, Shi J, Philips R, Mertens A, Pirracchio R, van der Laan M, Patel R, Colford J, Hubbard A. Predicting Long COVID in the National COVID Cohort Collaborative Using Super Learner. [Preprint]. 2023 August 04. DOI: 10.1101/2023.07.27.23293272

Certification:

I certify that the information provided is current, accurate, and complete. This includes, but is not limited to, information related to current, pending, and other support (both foreign and domestic) as defined in 42 U.S.C. § 6605.

In accordance with Section 10632 of the CHIPS and Science Act of 2022 (42 U.S.C. § 19232), each individual identified as a senior/key person must certify that they are not a party to a malign foreign talent recruitment program.

Research Security Training Requirement for Federal Award Personnel: In accordance with Section 10634 of the CHIPS and Science Act of 2022 (42 U.S.C. § 19234), each individual identified as a senior/key person must certify that they have completed the requisite research security training that meets the requirements specified in Item 2 of Important Notice No. 149 within 12 months prior to proposal submission.

Misrepresentations and/or omissions may be subject to prosecution and liability pursuant to, but not limited to, 18 U.S.C. §§287, 1001, 1031 and 31 U.S.C. §§3729-3733 and 3802.

Certified by Butzin-Dozier, Zachary in SciENcv on 2026-05-14 20:28:06

NIH BIOGRAPHICAL SKETCH SUPPLEMENT

Name: Butzin-Dozier, Zachary

Persistent Identifier (PID) of the Senior/Key Person: <https://orcid.org/0000-0001-6419-0008>

Position Title: Assistant Professor

Organization and Location: Stanford University School of Medicine, Department of Pediatrics, Division of Clinical Informatics, Stanford, California, United States

Personal Statement

I am an Assistant Professor at Stanford University School of Medicine with a joint appointment in the Department of Pediatrics, Division of Clinical Informatics, and the Department of Medicine, Division of Computational Medicine. I am an expert in applying causal inference and machine learning methods to observational electronic health record (EHR) data.

As the principal investigator of this K01 award, I am responsible for ensuring that all training goals and research aims are met in a timely manner. I have made substantial progress during the first year of this award: I published two first-author papers in BMC Medicine and JMIR Public Health and Surveillance, submitted three additional first-author manuscripts currently under peer review, and delivered four invited presentations to national audiences. My work was recognized as the NCATS Clinical and Translational Sciences Awards spotlight article of the month in September 2024.

I completed my postdoctoral training in the Division of Biostatistics at the UC Berkeley School of Public Health under Professors Alan Hubbard and Jack Colford. My experience with advanced biostatistical and causal inference methods, infectious disease epidemiology, and the National Clinical Cohort Collaborative (N3C) makes me uniquely qualified to conduct this research. I received my Ph.D. in Epidemiology and MPH in Epidemiology and Biostatistics from UC Berkeley. My work applies Targeted Machine Learning methodology, which is a method of applied machine learning for causal inference, to data from N3C in order to evaluate determinants of long-term sequelae of viral infection.

In addition, my research evaluates drivers of WASH intervention effectiveness and developmental outcomes for children in Bangladesh. I have developed expertise in the application of advanced study design and analysis methods to facilitate robust causal inferences in observational settings. I have situated myself between cutting-edge developments in analytic methods and urgent infectious disease research that seeks to answer pressing clinical and epidemiologic questions. My previous research and training have prepared me to connect innovations in biostatistical methodology with these high-impact applications, which will apply to infectious disease epidemiology more broadly.

Honors

2024	UC Berkeley School of Public Health Committee on Teaching Excellence Honoree, UC Berkeley School of Public Health
2023	Placed third in the NIH Long COVID Computational Challenge for building an ensemble machine learning model to predict individual risk of Long COVID diagnosis, Long COVID Computational Challenge (sponsored by the National Institutes of Health)
2023	Mentor's Choice and Community Choice Award for Outstanding Research Proposal, Biomedical Data Science Lab
2020 - 2022	Marcus Foster Education Institute Graduate Fellowship for Research on Equity in Education, Marcus Foster Education Institute
2019	Center for Latin American Studies Field Research Grant, UC Berkeley Center for Latin American Studies
2019 - 2020	Foreign Language and Area Studies (FLAS) Academic Year Fellowship for Portuguese language and Latin American area studies, US Department of Education

Contributions to Science

1. My primary research contribution is the development and application of Targeted Machine Learning methods to large-scale electronic health record data in N3C to evaluate long-term sequelae of viral infection. I led a study evaluating the relationship between SSRI use during acute viral infection and subsequent post-viral conditions among patients with depression, finding that SSRI use was associated with a lower risk of post-viral conditions compared to non-use — providing support for SSRIs as a preventive intervention and for the hypothesis that serotonin is a mechanistic biomarker of post-viral pathophysiology (BMC Medicine, 2024). I led a 10-member team in an NIH-sponsored machine learning competition, where we built an ensemble machine learning model using Super Learner to predict individual risk of post-viral conditions, achieving an AUC of 0.95 and placing third; I published this work in JMIR Public Health and Surveillance (2024), which was recognized as the NCATS CTSA spotlight article of the month. I am additionally developing methods for causal inference in EHR data, including a manuscript under review at BMC Medical Research Methodology on analytic challenges in N3C, and contributing to statistical theory through work on the Highly Adaptive LASSO under review at the Journal of Machine Learning Research.
2. I have led and supported research through the WASH Benefits Bangladesh study, which randomized pregnant mothers in rural Bangladesh to receive water, sanitation, hygiene, and nutrition interventions and prospectively monitored child health and development. I led a study evaluating the relationship between a broad panel of child stress biomarkers and child development, where we found that stress biomarkers from the hypothalamic-pituitary-adrenal axis were most associated with child developmental status (accepted for publication by Psychoneuroendocrinology).
3. Ring trial design: My previous work evaluated an underutilized trial design, the ring trial design, which may be useful in disease settings with high spatiotemporal clustering. I conducted a systematic review to evaluate the merits and drawbacks of this design, previous uses of rings in trials and interventions, and potential applications of this design in the future.
4. My early work developed a novel method to assess veterinary antibiotic use, a simulated client method, which may be generalized to other research that aims to assess authentic veterinary salesperson behavior. Furthermore, this research evaluated antibiotic recommendation practices by veterinary salespeople in rural Ecuador.
5. I supported a project evaluating the relationship between a caregiver intervention (Attachment and Biobehavioral Catchup) and child stress in a sample of children involved with child protective services. We found that children living with caregivers who received the Attachment and Biobehavioral Catchup demonstrated a less blunted cortisol response than children living with caregivers who did not receive this intervention, providing support for the effectiveness of this intervention with respect to child stress.

Certification:

I certify that the information provided is current, accurate, and complete. This includes, but is not limited to, information related to current, pending, and other support (both foreign and domestic) as defined in 42 U.S.C. § 6605.

In accordance with Section 10632 of the CHIPS and Science Act of 2022 (42 U.S.C. § 19232), each individual identified as a senior/key person must certify that they are not a party to a malign foreign talent recruitment program.

Research Security Training Requirement for Federal Award Personnel: In accordance with Section 10634 of the CHIPS and Science Act of 2022 (42 U.S.C. § 19234), each individual identified as a senior/key person must certify that they have completed the requisite research security training that meets the requirements specified in Item 2 of Important Notice No. 149 within 12 months prior to proposal submission.

Misrepresentations and/or omissions may be subject to prosecution and liability pursuant to, but not limited to, 18 U.S.C. §§287, 1001, 1031 and 31 U.S.C. §§3729-3733 and 3802.

Certified by Butzin-Dozier, Zachary in SciENcv on 2026-05-1420:28:06