

Akshay Swaminathan

akshayswaminathan.com

akshay325@gmail.com

EDUCATION

Stanford University 2021 – 2025
MD, Knight-Hennessy Scholar
PhD, Biomedical Data Science, PI: Nigam Shah, PD Soros Fellow

Harvard University 2015 – 2019
BA, Statistics (High Honors) with Secondary Degree in Global Health and Health Policy

APPLIED STATISTICS EXPERIENCE

Cerebral Inc. 2021 – 2024
Head of Data Science

- Deployed a patent-pending, NLP-enabled crisis message detection system that identifies patients with acute suicidal/homicidal ideation. Validated on 150k chat messages and had 97% sensitivity, 97% specificity. Press coverage: [Yahoo News](#), [Behavioral Health Business](#), [Rama On Healthcare](#), [Cerebral](#), [Stanford HAI](#)
- Developed personalized data reports for mental health clinicians that highlights key metrics to improve clinical quality. Reports are sent out to 2,000+ clinicians per month.
- Executed a pragmatic randomized trial to decrease follow-up time for patients with active suicidal ideation. Abstract accepted for oral presentation at American Telemedicine Association Conference.
- Lead a team of 15 data scientists to design and deploy data-driven clinical quality interventions that impact >700,000 patients.

Stanford University 2021 – 2023
Founder, Unstructured Data Project, PI: Nigam Shah | udp.stanford.edu

- Started a student-led organization with 60+ members to extract data from unstructured clinical documents using machine learning and NLP.
- Developed a desktop application that speeds up the process of clinical chart review by up to 10x compared to traditional methods.
- Developed a machine learning workflow that leverages selective prediction to extract variables typically inaccessible to researchers. Collaborated with clinical faculty in oncology, general surgery, neurosurgery, neonatology, primary care to extract important variables.

University of Pennsylvania 2021 – 2023
Lead Data Scientist, PI: David Fajgenbaum

- Used evidence synthesis methods to grade the real-world effectiveness of repurposed drugs to treat COVID-19. Built a [data dashboard](#) that was used by policy-makers and academics. Press coverage: [CNN](#), [STAT](#)

Flatiron Health 2019 – 2021
Data Scientist, Clinico-Genomics Team

- Conduct outcomes research using real-world clinical and genomic data to improve oncology care.
- Research on real-world outcomes of MSI-H patients treated with pembrolizumab presented at ASCO 2020.
- Developed methods to quantify and mitigate left-truncation bias in real-world data.
- Led study of real-world outcomes and treatment patterns of Castleman Disease patients.

Harvard TH Chan School of Public Health

2017 – 2021

Researcher at Harvard Center for Population and Development Studies, PI: Dr. SV Subramanian

- Developed four novel geographic information system methodologies to estimate child malnutrition for 543 Indian parliamentary constituencies.
- Applied multilevel modeling and predictive algorithms to analyze maternal/child health outcomes in India.
- Built a [policy tracker](#) that has been featured by UNICEF.

Harvard Medical School

2017 – 2021

Researcher in Department of Global Health and Social Medicine, PI: Dr. Vikram Patel

- Developed a predictive algorithm for child cognitive impairment with Public Health Foundation of India. Outperformed proprietary algorithm designed by third-party consultants.
- Used causal random forests to assess effectiveness of randomized control trial of 18,000+ subjects in Bihar, India. Outcomes were depression, gender equity, and school climate.
- Serving as principal biostatistician on the HUHS iWellness Study.

Ottawa Hospital Research Institute, OMNI Research Group

2018 – 2021

Visiting Researcher at Clinical Epidemiology Program, PI: Dr. Daniel Corsi

- Used causal inference techniques to study stillbirths of 1 million women from 52 low and middle-income countries
- Analyzing clinical longitudinal data to develop a predictive model for pre-eclampsia

St. John's National Academy of Health Sciences, Bangalore, India

2017 – 2018

Intern in Division of Epidemiology and Biostatistics, St. John's Research Institute, PI: Dr. Tinku Thomas

- Studied diabetes, anemia, and maternal/child health outcomes across India.

Radcliffe Research Partnership Program, Harvard University

2016 – 2017

Applied Statistics Research Assistant, PI: Dr. Gidon Eshel

- Applied statistical techniques to find plant-based substitutes for beef, pork, and chicken that minimize environmental costs.

GLOBAL HEALTH EXPERIENCE

Refresh Bolivia Inc. (501(c)(3) nonprofit)

2015 – 2019

Executive Director | www.refreshbolivia.com | www.refreshbolivia.wixsite.com/rbexecblog

- Constructed a community health center in peri-urban Cochabamba that serves an area of 10,000+ residents.
- Led 7 trips to Cochabamba, Bolivia, and mobilized 400+ students, professionals, and volunteers from 12+ universities in the US and Bolivia.
- Public health interventions: built 20+ one-cup-flush ecological toilets, produced and distributed 100+ reusable handmade menstrual pads, educated 2,500+ women and children

on personal health and hygiene, helped deworm 1,000+ people, gave workshops on conflict resolution to 80+ incarcerated youth

- Raised \$120,000+, including \$85,000 in grants from Ford Motor Company.

United Nations Population Fund

2018

Intern in Sexual and Reproductive Health Branch

- Designed a data collection tool for 40+ target countries in Sub-Saharan Africa and Asia, analyzing data from regional and country offices, and preparing materials for resource mobilization.

FOREIGN LANGUAGE EDUCATION EXPERIENCE

Autodidactic Polyglot | www.startspeaking.org

2011 – Present

- Experience in **11** languages: [*Advanced*] English, Spanish, French, Mandarin, Tamil, [*Intermediate*] Portuguese, Italian, Hindi, [*Conversational*] Russian, Cantonese, Quechua
- Released 7 books on language acquisition which have sold 2,000+ copies (not including free downloadable materials); used by universities, high schools, companies, other institutions and individuals around the world.
- Run a YouTube channel with 250,000+ views and 3,900+ subscribers.

Harvard Chinatown ESL

2016 – 2018

Director

- Wrote textbooks for 5 levels of English instruction. Sold 1,500+ copies of newly written textbooks. Materials and teaching entirely in Mandarin Chinese.
- Established YouTube and Youku channels with pedagogical resources that have received 20,000 views.
- Designed an ESL teacher training program for all of Harvard's ESL classes. Trained 200+ ESL tutors.

SELECTED PEER-REVIEWED PUBLICATIONS

For complete list of publications see [Google Scholar](#)

- [Book] *Winning with Data Science*. Columbia University Press. 2024. Co-authored with Howard Friedman. Covers fundamental concepts in statistics, machine learning, causal inference. Target audience is business professionals who collaborate with data scientists.
- Natural language processing system for rapid detection and intervention of mental health crisis chat messages. *npj Digital Medicine*. Nov 2023. First and corresponding author.
- Selective prediction for extracting unstructured clinical data. *Journal of the American Medical Informatics Association*. September 2023. First and corresponding author.
- Association between interpregnancy interval and subsequent stillbirth: a retrospective analysis of 716,478 births from 58 low and middle-income countries. *Lancet Global Health*. Jan 2020. First author.
- The State of School Infrastructure in the Assembly Constituencies of Rural India: Analysis of 11 Census Indicators from Pre-Primary to Higher Education. *International Journal of Environmental Research and Public Health*. Jan 2020. First author. **Featured press:** [Harvard School of Public Health](#).
- Burden of Child Malnutrition in India: A View from Parliamentary Constituencies. *Economic and Political Weekly*. January 12, 2019. **Featured press:** [Times of India](#), [Harvard Gazette](#).

[Indian Express](#), [DownToEarth](#), [New Indian Express](#), [LiveMint](#), [IndiaOutbound](#), [IndiaSpend](#), [Outlook](#), [Business Standard](#), [Governance Now](#), [Financial Express](#), [The Week](#), [Scroll.In](#), [The Telegraph](#). First author.

- Estimating Water, Sanitation, and Hygiene Indicators for Parliamentary Constituencies in India. *Journal of Development Policy and Practice*. September 2019. Third author. **Featured press:** [DownToEarth](#), [The CSR Journal](#).
- Environmentally Optimal, Nutritionally Aware Meat Replacement Plant-Based Diets. *Nature Scientific Reports*. August 2019. **Top 100 downloaded papers in 2019. Featured press:** [New York Times](#), [Scientific American](#), [Business Insider](#), [WebMD](#). Altimetric: 367, Downloaded 16,000+ times.
- Precision mapping child undernutrition for nearly 600,000 inhabited census villages in India. *PNAS*. April 2021.
- Proof of Concept of a Gamified DEvelopmental Assessment on an E-Platform (DEEP) Tool to Measure Cognitive Development in Rural Indian Preschool Children. *Frontiers in Psychology*. June 2020. Third author.
- Association does not imply prediction: The discriminatory accuracy of low birthweight in predicting child mortality and anthropometric failure. *Annals of Epidemiology*. October 2020. First author.
- Forecasting efforts from prior epidemics and COVID-19 predictions. *European Journal of Epidemiology*. July 2020. Second author.
- Reflections on designing population surveys for COVID-19 infection and prevalence. *GeroScience*. August 2020. First author.
- Validity of Maternal Report of Birthweight in a Cohort Study and Its Implication on Low Birthweight Rate Using Simulations. *Indian Pediatrics*. November 2019. First author.
- Robust Data Science Approaches to Estimating Parliamentary Constituency Indicators. *Economic and Political Weekly*. May 2019. Second author.
- Estimating the Burden of Child Malnutrition across Parliamentary Constituencies in India: A Methodological Comparison. *Social Science & Medicine: Population Health*. April 2019. Second author.
- Prevalence and Correlates of Gestational Diabetes in India: Cross-Sectional Analysis of a Nationally Representative Survey. *JAMA Network Open*. December 2020. Second author.
- Real-world association of HER2/ERBB2 concordance with trastuzumab clinical benefit in advanced esophagogastric cancer. *Future Oncology*. July 2021.

OTHER PUBLICATIONS

- 流利英语1000句:一年级 (1000 Sentences to Start Speaking English: Level 1). Amazon Publishing. 2016.
- 流利英语1000句:二年级 (1000 Sentences to Start Speaking English: Level 2). Amazon Publishing. 2017.
- 流利英语1000句:三年级 (1000 Sentences to Start Speaking English: Level 3). Amazon Publishing. 2017.
- 流利英语1000句:四年级 (1000 Sentences to Start Speaking English: Level 4). Amazon Publishing. 2018.
- 流利英语1000句:四年级 (1000 Sentences to Start Speaking English: Level 5). Self-Published. 2018.
- 1000 Sentences to Improve Your Spanish: Advanced Level. Amazon Publishing. 2018.
- 1000 Sentences to Start Speaking Quechua: Level 1. Self-published. 2018
- "Audio Dissection". *Language Magazine*. 2015.

- Parliamentary Constituency Factsheet for Indicators of Nutrition, Health and Development in India. HCPDS Working Paper Vol. 18, No. 4. Second author.
- State of Nutrition among Children in Parliamentary Constituencies of India. Tata Trusts: Mumbai, India. 2018. Collaborator.

ACCEPTED ABSTRACTS

- *Randomized trial of patient and clinician alerts to decrease follow-up time following suicidal crises.* American Telemedicine Association 2023.
- *Natural language processing (NLP) for faster detection and intervention of mental health crisis messages.* American Telemedicine Association 2023.
- *TeleHelp Ukraine: A Novel Telemedicine Program Serving War-Impacted Patients.* American Telemedicine Association 2023.
- *Changes over time in real-world next-generation sequencing test use in patients with advanced non-small cell lung cancer.* American Association for Cancer Research 2021.
- *Real-world overall survival and time to therapy discontinuation of patients (pts) with mCRPC treated with second-generation novel hormonal therapies associated with tissue-based comprehensive genomic profiling.* American Society of Clinical Oncology Genitourinary Cancers Symposium 2021.
- *Enhanced Cost-Effectiveness Analysis Using HER Data for Real World Value.* International Society for Pharmacoeconomics and Outcomes Research 2021.
- *Statistical Methods for Pan-tumor Analysis: Models to Account for Tumor-Level Heterogeneity.* International Society for Pharmacoeconomics and Outcomes Research 2021.
- *Characteristics and outcomes of real-world patients with microsatellite instability-high solid tumors treated with pembrolizumab monotherapy after FDA approval.* American Society of Clinical Oncology 2020.
- *The effect of interpregnancy interval on subsequent stillbirth.* Society for Pediatric and Perinatal Epidemiologic Research 2019.
- *Interpregnancy Interval and Stillbirth: A Causal Analysis of the Demographic and Health Surveys.* Canadian National Perinatal Research Meeting (CNPRM) 2019.
- *Global Epidemiology of Stillbirth: Analysis of the Demographic and Health Surveys.* CNPRM 2019.
- *Bolivian water wars: A fight for the right to water.* American Public Health Association Conference (APHA) 2018.
- *Can serious games measure cognitive development?* Public Health Foundation of India Research Symposium 2018.
- *Cost-effectiveness and cost-utility analysis of DIGNITY multidisciplinary rehabilitation model for traumatized refugees during 2012-2015.* APHA 2018.
- *The Harvard Model for Student-led Adult Education Programming.* Massachusetts Coalition for Adult Education Network Conference 2018.
- *Health Behaviors and Perceptions in La Zona Sur, Cochabamba.* APHA 2017, Community-based primary health care and community health workers workshop.

GUEST LECTURES/TALKS

- *Straying from the path.* TEDx Bergen County Academies. October 9, 2019.
- *A Model for Youth-Led Global Health Initiatives.* United Nations Population Fund, August 9, 2018; January 2019; Bergen County Academies.

- *Fundamentals of Language acquisition.* Harvard Extension School. September 2016, November 2017, November 2018.
- *La adquisición de los idiomas extranjeros.* San Simon University Linguistics Department. May 2016.
- *Strategies for Learning Chinese.* Point Pleasant Beach High School. June 2017, June 2018.

TV INTERVIEWS

- *Que no me pierda,* Bolivia's largest evening news program. June 1, 2016.
- TVC, news station of Cochabamba. May 31, 2016.
- Univalle, news station of Cochabamba. May 2016, January 2017, January 2018.
- Red Uno, national Bolivian news network. June 2, 2016 and January 14, 2017.

HONORS AND AWARDS

- PD Soros Fellowship 2024
- Knight-Hennessy Fellowship 2021
- Rhodes Scholarship Finalist 2018
- Hertz Fellowship Semi-finalist 2024
- Benjamin Franklin Legacy Prize 2019
(for creativity and leadership in public service, 5 winners nationally)
- Pforzheimer Prize for Public Service 2019
- 2x Harvard Global Health Institute Summer Internship Grant Winner (\$11,000) 2017, 2018
- 4x Ford College Community Challenge (\$85,000) 2016, 2017, 2018, 2019
- 2x Harvard College Research Program Grant Winner (\$2,000) 2016, 2017
- Arist Course Competition 1st place (\$2,000). 2020
 - Try our [free course](#) called "Spotting Fake Statistics in Real Life"