Machine Learning in Population and Public Health: Challenges and Opportunities

Rumi Chunara, Yuan Zhao, Vishwali Mhasawade

ACM CHIL 2020 Tutorial

NYU SCHOOL OF GLOBAL PUBLIC HEALTH

NYU TANDON SCHOOL OF ENGINEERING

Tutorial Goals

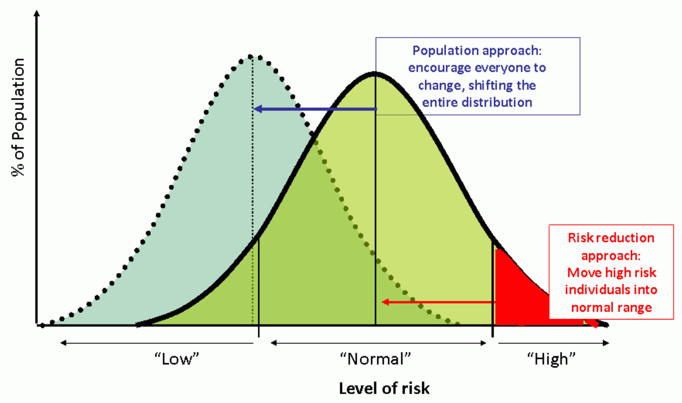
Familiarize with research in public and population health
Identify open areas related to health equity
Activate the machine learning community

- Introduce Public and Population Health

- Theory and framework of social determinants of health (SoDH)
- Measurement of SoDH
- SoDH interventions
- Integration of SoDH in machine learning models

- Taxonomy of health tasks
- Causal inference in public health
- Challenges with using proxies
- Algorithmic fairness and health disparities

Focus is on what makes and keeps people healthy (holistically). Aim is to improve the health of the entire population and to reduce health inequities among population groups [Rose, 1985].



What is Health Equity?

Health equity

"Minimizing avoidable disparities in health and its determinants–including but not limited to health care–between groups of people who have different levels of underlying social advantage or privilege, i.e., different levels of power, wealth, or prestige due to their positions in society relative to other groups."
[Braveman, 2006]

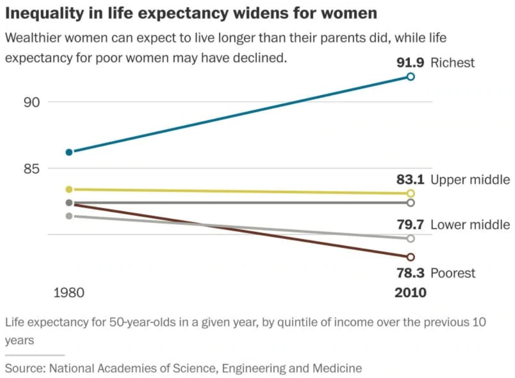


Figure 1: Growing inequalities for women in the United States
[National Academies, 2015].

Figure 2: Socio-ecological model of health [Bronfenbrenner, 1977].

Impact

Figure 3: [CDC, 2014]

United States Social factors account for 25–60 percent of deaths in any given year according to results from various meta-analyses. [Heiman and Artiga, 2015]

Worldwide: Eighty per cent of noncommunicable diseases could be prevented through primary prevention – through modifying behaviours such as reducing tobacco consumption and fat, alcohol and salt intake, preventing obesity, and promoting physical activity, and improving environmental conditions such as air quality and urban planning [WHO, 2009].

Importance of considering all determinants of health - COVID-19 case study

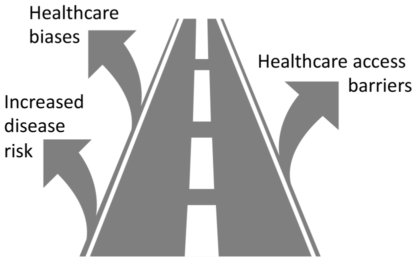


Figure 4: “Highway to health”.

Determinant (COVID relevance [Rollston and Galea, 2020])

- Housing conditions (crowding, poor sanitation)
- Healthy food access (increased comorbidities)
- Education access and quality (health literacy, future socioeconomic status)
- Socio-economic status (occupation/essential worker, ability to self-isolate)
- Healthcare trust (access treatment)
- Racism/discrimination (socioeconomic status, segregation, housing quality, health care access, quality, etc..)

For a complete list of related articles, more information and to give feedback:
https://ChunaraLab.github.io/MLPH/

Ahsan, G. T., Addo, I. D., Ahamed, S. I., Petereit, D., Kanekar, S., Burhansstipanov, L., and Krebs, L. U. (2013). Toward an mhealth intervention for smoking cessation. In *2013 IEEE 37th Annual Computer Software and Applications Conference Workshops*, pages 345–350. IEEE.

Akbari, M. and Chunara, R. (2019). Using contextual information to improve blood glucose prediction. *Machine Learning for Healthcare*, arXiv preprint arXiv:1909.01735.

Bauer, G. R. (2014). Incorporating intersectionality theory into population health research methodology: challenges and the potential to advance health equity. *Social science & medicine*, 110:10–17.

Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., Drake, J. M., Brownstein, J. S., Hoen, A. G., Sankoh, O., et al. (2013). The global distribution and burden of dengue. *Nature*, 496(7446):504–507.

Bonham, V. L., Green, E. D., and Pérez-Stable, E. J. (2018). Examining how race, ethnicity, and ancestry data are used in biomedical research. *Jama*, 320(15):1533–1534.

Braveman, P. (2006). Health disparities and health equity: concepts and measurement. *Annu. Rev. Public Health*, 27:167–194.

Braveman, P. A., Egerter, S. A., Cubbin, C., and Marchi, K. S. (2004). An approach to studying social disparities in health and health care. *American Journal of Public Health*, 94(12):2139–2148.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American psychologist*, 32(7):513.

Cable, N. (2014). Life course approach in social epidemiology: an overview, application and future implications. *Journal of epidemiology*, page JE20140045.

CDC (2014). Up to 40 percent of annual deaths from each of five leading us causes are preventable. *Atlanta, GA: Centers for Disease Control and Prevention*.

Chunara, R., Plymoth, A., and Martin, L. (2020). Diversity in surveillance data: implications for infectious disease forecasting models.

Chunara, R., Wisk, L. E., and Weitzman, E. R. (2017). Denominator issues for personally generated data in population health monitoring. *American journal of preventive medicine*, 52(4):549–553.

Collins, P. H. (2002). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. routledge.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *u. Chi. Legal f.*, page 139.

da Silveira Grübler, M., da Costa, C. A., da Rosa Righi, R., Rigo, S. J., and Chiwiacowsky, L. D. (2018). A hospital bed allocation hybrid model based on situation awareness. *CIN: Computers, Informatics, Nursing*, 36(5):249–255.

Dugan, T. M., Mukhopadhyay, S., Carroll, A., and Downs, S. (2015). Machine learning techniques for prediction of early childhood obesity. *Applied clinical informatics*, 6(03):506–520.

Evans, C. R., Williams, D. R., Onnela, J.-P., and Subramanian, S. (2018). A multilevel approach to modeling health inequalities at the intersection of multiple social identities. *Social Science & Medicine*, 203:64–73.

Franklin, J. M., Shrank, W. H., Lii, J., Krumme, A. K., Matlin, O. S., Brennan, T. A., and Choudhry, N. K. (2016). Observing versus predicting: initial patterns of filling predict long-term adherence more accurately than high-dimensional modeling techniques. *Health services research*, 51(1):220–239.

Friedler, S. A., Scheidegger, C., and Venkatasubramanian, S. (2016). On the (im) possibility of fairness. *arXiv preprint arXiv:1609.07236*.

Galiatsatos, P., Follin, A., Uradu, N., Alghanim, F., Daniel, Y., Saria, S., Townsend, J., Sylvester, C., Chanmugam, A., and Chen, E. (2019). The association between neighborhood socioeconomic disadvantage and readmissions for patients hospitalized with sepsis. In *C94. The Impact of Social Determinants in Pulmonary and Critical Care*, pages A5569–A5569. American Thoracic Society.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 316(22):2402–2410.

Hamad, R., Nguyen, T. T., Bhattacharya, J., Glymour, M. M., and Rehkopf, D. H. (2019). Educational attainment and cardiovascular disease in the united states: A quasi-experimental instrumental variables analysis. *PLoS medicine*, 16(6):e1002834.

Heiman, H. J. and Artiga, S. (2015). Beyond health care: the role of social determinants in promoting health and health equity. *Health*, 20(10):1–10.

Hermanspann, T., Schoberer, M., Robel-Tillig, E., Härtel, C., Goelz, R., Orlikowsky, T., and Eisert, A. (2017). Incidence and severity of prescribing errors in parenteral nutrition for pediatric inpatients at a neonatal and pediatric intensive care unit. *Frontiers in pediatrics*, 5:149.

Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudik, M., and Wallach, H. (2019). Improving fairness in machine learning systems: What do industry practitioners need? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–16.

Kasy, M. and Abebe, R. (2020). Fairness, equality, and power in algorithmic decision making. Technical report, Working paper.

Kreif, N., Grieve, R., Díaz, I., and Harrison, D. (2015). Evaluation of the effect of a continuous treatment: a machine learning approach with an application to treatment for traumatic brain injury. *Health economics*, 24(9):1213–1228.

Krieger, N. and Davey Smith, G. (2016). The tale wagged by the dag: broadening the scope of causal inference and explanation for epidemiology. *International journal of epidemiology*, 45(6):1787–1808.

Kusnoor, S. V., Koonce, T. Y., Hurley, S. T., McClellan, K. M., Blasingame, M. N., Frakes, E. T., Huang, L.-C., Epelbaum, M. I., and Giuse, N. B. (2018). Collection of social determinants of health in the community clinic setting: A cross-sectional study. *BMC Public Health*, 18(1):550.

Lê-Scherban, F., Wang, X., Boyle-Steed, K. H., and Pachter, L. M. (2018). Intergenerational associations of parent adverse childhood experiences and child health outcomes. *Pediatrics*, 141(6):e20174274.

Lehman, C. (2019). Addressing social determinants of health. *Physical Therapy in Motion*.

Lodi, S., Phillips, A., Lundgren, J., Logan, R., Sharma, S., Cole, S. R., Babiker, A., Law, M., Chu, H., Byrne, D., et al. (2019). Effect estimates in randomized trials and observational studies: comparing apples with apples. *American journal of epidemiology*, 188(8):1569–1577.

Lord, A. S., Carman, H. M., Roberts, E. T., Torrico, V., Goldmann, E., Ishida, K., Tuhrim, S., Stillman, J., Quarles, L. W., and Boden-Albala, B. (2015). Discharge educational strategies for reduction of vascular events (deserve): design and methods. *International journal of stroke*, 10(SA100):151–154.

Mahamoud, A., Roche, B., and Homer, J. (2013). Modelling the social determinants of health and simulating short-term and long-term intervention impacts for the city of toronto, canada. *Social science & medicine*, 93:247–255.

Marmot, M. (2002). The influence of income on health: views of an epidemiologist. *Health affairs*, 21(2):31–46.

McCall, L. (2005). The complexity of intersectionality. *Signs: Journal of women in culture and society*, 30(3):1771–1800.

Mcclure, L. (2018). *Controlling for smoking?*

McCradden, M. D., Joshi, S., Mazwi, M., and Anderson, J. A. (2020). Ethical limitations of algorithmic fairness solutions in health care machine learning. *The Lancet Digital Health*, 2(5):e221–e223.

McGibbon, E. and McPherson, C. (2011). Applying intersectionality & complexity theory to address the social determinants of women's health.

McGuire, T. G. (2016). Achieving mental health care parity might require changes in payments and competition. *Health Affairs*, 35(6):1029–1035.

Merlo, J., Ohlsson, H., Lynch, K. F., Chaix, B., and Subramanian, S. (2009). Individual and collective bodies: using measures of variance and association in contextual epidemiology. *Journal of Epidemiology & Community Health*, 63(12):1043–1048.

Merlo, J. and Wagner, P. (2013). The tyranny of the averages and the indiscriminate use of risk factors in public health: a call for revolution. *Eur J Epidemiol*, 28(Suppl 1):148.

Mhasawade, V., Elghafari, A., Duncan, D. T., and Chunara, R. (2020a). Role of the built and online social environments on expression of dining on instagram. *International journal of environmental research and public health*, 17(3):735.

Mhasawade, V., Rehman, N. A., and Chunara, R. (2020b). Population-aware hierarchical bayesian domain adaptation via multi-component invariant learning. In *Proceedings of the ACM Conference on Health, Inference, and Learning*, pages 182–192.

Mitchell, M. and Jolley, J. (2004). Research design explained 5 th ed. *Victoria: Wadsworth Publisher. Moebert, J. & Tydecks, P.(2007). Power and Ownership Structures among German Companies. A Network Analysis of Financial Linkages.*

Mitchell, S., Potash, E., Barocas, S., D'Amour, A., and Lum, K. (2018). Prediction-based decisions and fairness: A catalogue of choices, assumptions, and definitions. *arXiv preprint arXiv:1811.07867*.

National Academies (2015). *The growing gap in life expectancy by income: Implications for federal programs and policy responses*. National Academies Press.

National Academies of Sciences, E., Medicine, et al. (2019). *Integrating social care into the delivery of health care: moving upstream to improve the nation's health*. National Academies Press.

Nishtala, S., Kamarthi, H., Thakkar, D., Narayanan, D., Grama, A., Padmanabhan, R., Madhiwalla, N., Chaudhary, S., Ravindra, B., and Tambe, M. (2020). Missed calls, automated calls and health support: Using ai to improve maternal health outcomes by increasing program engagement. *arXiv preprint arXiv:2006.07590*.

Obermeyer, Z. and Mullainathan, S. (2019a). Dissecting racial bias in an algorithm that guides health decisions for 70 million people. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 89–89.

Obermeyer, Z. and Mullainathan, S. (2019b). Dissecting racial bias in an algorithm that guides health decisions for 70 million people. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 89, New York, NY, USA. Association for Computing Machinery.

Osborn, S. R., Yu, J., Williams, B., Vasilyadis, M., and Blackmore, C. C. (2017). Changes in provider prescribing patterns after implementation of an emergency department prescription opioid policy. *The Journal of emergency medicine*, 52(4):538–546.

Penman-Aguilar, A., Talih, M., Huang, D., Moonesinghe, R., Bouye, K., and Beckles, G. (2016). Measurement of health disparities, health inequities, and social determinants of health to support the advancement of health equity. *Journal of public health management and practice: JPHMP*, 22(Suppl 1):S33.

Quisel, T., Kale, D. C., and Foschini, L. (2016). Intra-day activity better predicts chronic conditions. *arXiv preprint arXiv:1612.01200*.

Rahmattalabi, A., Adhikari, A. B., Vayanos, P., Tambe, M., Rice, E., and Baker, R. (2018). Influence maximization for social network based substance abuse prevention. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., Liu, P. J., Liu, X., Marcus, J., Sun, M., et al. (2018). Scalable and accurate deep learning with electronic health records. *NPJ Digital Medicine*, 1(1):18.

Rollston, R. and Galea, S. (2020). Covid-19 and the social determinants of health. *American Journal of Health Promotion*, 34(6):687–689. PMID: 32551932.

Rose, G. (1985). Sick individuals and sick populations, int. journ. *Of Epidemiology*, 14(1).

Rose, S., Bergquist, S. L., and Layton, T. J. (2017). Computational health economics for identification of unprofitable health care enrollees. *Biostatistics*, 18(4):682–694.

Schwartz, S., Prins, S. J., Campbell, U. B., and Gatto, N. M. (2016). Is the "well-defined intervention assumption" politically conservative? *Social science & medicine (1982)*, 166:254.

Snyder, J. J., Salkowski, N., Wey, A., Pyke, J., Israni, A. K., and Kasiske, B. L. (2018). Organ distribution without geographic boundaries: a possible framework for organ allocation. *American Journal of Transplantation*, 18(11):2635–2640.

Vyas, D. A., Eisenstein, L. G., and Jones, D. S. (2020). Hidden in plain sight—reconsidering the use of race correction in clinical algorithms.

Ward, J. B., Gartner, D. R., Keyes, K. M., Fliss, M. D., McClure, E. S., and Robinson, W. R. (2019). How do we assess a racial disparity in health? distribution, interaction, and interpretation in epidemiological studies. *Annals of epidemiology*, 29:1–7.

Weichenthal, S., Olaniyan, T., Christidis, T., Lavigne, E., Hatzopoulou, M., Van Ryswyk, K., Tjepkema, M., and Burnett, R. (2020). Within-city spatial variations in ambient ultrafine particle concentrations and incident brain tumors in adults. *Epidemiology (Cambridge, Mass.)*, 31(2):177.

WHO (2008). *Closing the gap in a generation: Health equity through action on the social determinants of health: Commission on Social Determinants of Health final report*. World Health Organization.

WHO (2009). 2008-2013 action plan for the global strategy for the prevention and control of noncommunicable diseases: prevent and control cardiovascular diseases, cancers, chronic respiratory diseases and diabetes.

WHO (2010). A conceptual framework for action on the social determinants of health.

Wu, Y., Zhang, L., Wu, X., and Tong, H. (2019). Pc-fairness: A unified framework for measuring causality-based fairness. In *Advances in Neural Information Processing Systems*, pages 3404–3414.

Yang, K., Loftus, J. R., and Stoyanovich, J. (2020). Causal intersectionality for fair ranking. *arXiv preprint arXiv:2006.08688*.

Zhan, A., Mohan, S., Tarolli, C., Schneider, R. B., Adams, J. L., Sharma, S., Elson, M. J., Spear, K. L., Glidden, A. M., Little, M. A., et al. (2018). Using smartphones and machine learning to quantify parkinson disease severity: the mobile parkinson disease score. *JAMA neurology*, 75(7):876–880.

Zhao, Y., Mirin, N., Wood, E., Dorice, V., Rajesh, V., Cook, S., and Chunara, R. (2020). Machine learning for integrating social determinants in cardiovascular disease prediction models: A systematic review.