Ecological Factors Associated with Self-Reported Mental Health Status

Alyssa Berger, Andrew Cistola

Project Description
Ecological factors are known predictors of many health outcomes[1-9] and neighborhood level deprivation is a useful measure in assessing these outcomes to create targeted interventions.[10-14] Machine learning and spatial analysis also have potential to improve these methods[15-18] but there is little benefit over traditional approaches.[19-20] This kind of research within mental health is limited[21] and needs methodological improvement.[22] Therefore, there is a need for studies that can inform both data science[23-25] and translational science in mental health.[26-28]

Specific Aims

1. Identify local socioeconomic factors associated with a higher prevalence of poor self-reported mental health status
2. Identify area health resources that are associated with a lower prevalence of poor self-reported mental health status

Research Design

1. Collect estimated zip code level self-reported poor mental health status as percent of population from the 2020 release of the CDC and RWJF PLACES dataset.
2. Collect approximately 400 zip code level socio-economic variables as percent estimates from the 2020 release of the US Census American Community Survey.
3. Connect outcomes with predictors, remove observations with missing values, impute missing data for predictor variables using median values, standard scale all variables.
4. Collect approximately 2000 county level health resource variables from the 2020 release of the HRSA Area Health resource File.

Specific Aim 1:

5. Utilize open-source machine learning algorithms to identify socioeconomic variables with both high variation and high importance. Conduct cross-validated prediction to identify the smallest number of variables that will achieve the best fitting model for zip codes.
6. Calculate spatially adjusted rates and use an artificial neural network with backwards elimination to predict zip codes in the top quartile. Compare the predictive capability of all predictors, a random set of predictors, other known predictors, and the variables obtained in step 4.
7. Test for OLS assumptions and use domain knowledge to develop an appropriate mixed effects regression model for the purpose of identifying parameter estimates for socioeconomic predictors.

Specific Aim 2:

8. Using local Empirical Bayes smoothing and LISA quadrants, identify ‘hot and cold spot’ regions and assign nominal labels to contained counties.Use algorithms capable of multi-nominal prediction to identify health resources associated with each category.
9. Using zip code predictors, conduct geographic weighted regression to identify regions where each predictor has significantly higher or lower coefficients.Assign nominal labels and use algorithms capable of multi-nominal prediction to identify health resources associated with each category.
10. Test for OLS assumptions and use domain knowledge to develop an appropriate mixed effects regression model. The purpose is to identify parameter estimates for each county predictor identified in step 8 and an interaction term for county and zip code predictors in step 9.

Timeline
Our group has been meeting weekly to discuss project ideas and collaborate since January 29th, 2021. Current work at https://github.com/andrewcistola/PHC6194 (private). Preliminary results available for steps 1-5 of specific aim 1 and step 1-4 of specific aim 2

<u>January-March:</u> Finalize project ideas and proposal
<u>March 22-April 9</u>: Conduct analysis
<u>April 9-April 23</u>: Interpret Results and Finalize Report

References

1. Casper M, Kramer MR, Peacock JM, Vaughan AS. Population health, place, and space: spatial perspectives in chronic disease research and practice. Prev Chronic Dis. 2019 Sep 5;16:E123.
2. Bozigar M, Lawson AB, Pearce JL, King K, Svendsen ER. A Bayesian spatio-temporal analysis of neighborhood pediatric asthma emergency department visit disparities. Health Place. 2020 Nov;66:102426.
3. Cromer SJ, Lakhani CM, Wexler DJ, Burnett-Bowie S-AM, Udler M, Patel CJ. Geospatial Analysis of Individual and Community-Level Socioeconomic Factors Impacting SARS-CoV-2 Prevalence and Outcomes. medRxiv. 2020 Sep 30;
4. Richardson AS, Collins RL, Ghosh-Dastidar M, Ye F, Hunter GP, Baird MD, et al. Improvements in neighborhood socioeconomic conditions may improve resident diet. Am J Epidemiol. 2020 Oct 13;
5. Holder AL, Wallace DJ, Martin GS. Hotspotting sepsis: applying analytic tools from other disciplines to eliminate disparities. Ann Transl Med. 2016 Aug;4(15):295.
6. Kolak M, Bhatt J, Park YH, Padrón NA, Molefe A. Quantification of Neighborhood-Level Social Determinants of Health in the Continental United States. JAMA Netw Open. 2020 Jan 3;3(1):e1919928.
7. Beck AF, Anderson KL, Rich K, Taylor SC, Iyer SB, Kotagal UR, et al. Cooling the hot spots where child hospitalization rates are high: A neighborhood approach to population health. Health Aff (Millwood). 2019;38(9):1433–41.
8. Eibich P, Krekel C, Demuth I, Wagner GG. Associations between Neighborhood Characteristics, Well-Being and Health Vary over the Life Course. Gerontology 2016; 62: 362–370
9. Miles JN, Weden MM, Lavery D, Escarce JJ, Cagney KA, Shih RA. Constructing a Time-Invariant Measure of the Socio-economic Status of U.S. Census Tracts. J. Urban Health 2016; 93: 213–232.
10. Diez Roux AV, Mair C. Neighborhoods and health. Ann N Y Acad Sci. 2010 Feb;1186:125–45.
11. Scaria E, Powell WR, Birstler J, Alagoz O, Shirley D, Kind AJH, et al. Neighborhood disadvantage and 30-day readmission risk following Clostridioides difficile infection hospitalization. BMC Infect Dis. 2020 Oct 16;20(1):762.
12. Kind AJH, Buckingham WR. Making Neighborhood-Disadvantage Metrics Accessible - The Neighborhood Atlas. N Engl J Med. 2018 Jun 28;378(26):2456–8.
13. Messer LC, Laraia BA, Kaufman JS, Eyster J, Holzman C, Culhane J, et al. The development of a standardized neighborhood deprivation index. J Urban Health. 2006 Nov;83(6):1041–62.
14. Walker AF, Hu H, Cuttriss N, Anez-Zabala C, Yabut K, Haller MJ, et al. The neighborhood deprivation index and provider geocoding identify critical catchment areas for diabetes outreach. J Clin Endocrinol Metab. 2020 Sep 1;105(9).
15. Krittanawong C, Virk HUH, Bangalore S, Wang Z, Johnson KW, Pinotti R, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. Sci Rep. 2020 Sep 29;10(1):16057.
16. Benedetto U, Dimagli A, Sinha S, Cocomello L, Gibbison B, Caputo M, et al. Machine learning improves mortality risk prediction after cardiac surgery: Systematic review and meta-analysis. J Thorac Cardiovasc Surg. 2020 Aug 10;
17. Hu L, Liu B, Ji J, Li Y. Tree-Based Machine Learning to Identify and Understand Major Determinants for Stroke at the Neighborhood Level. J Am Heart Assoc. 2020 Nov 3;e016745.

18. Ji J, Hu L, Liu B, Li Y. Identifying and assessing the impact of key neighborhood-level determinants on geographic variation in stroke: a machine learning and multilevel modeling approach. BMC Public Health. 2020 Nov 7;20(1):1666.
19. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol. 2019 Feb 11;110:12–22.
20. Angier H, Jacobs EA, Huguet N, Likumahuwa-Ackman S, Robert S, DeVoe JE. Progress towards using community context with clinical data in primary care. Family Med Commun Hlth. 2019;7(1):e000028.
21. Forthman KL, Colaizzi JM, Yeh H-w, Kuplicki R, Paulus MP. Latent Variables Quantifying Neighborhood Characteristics and Their Associations with Poor Mental Health. International Journal of Environmental Research and Public Health. 2021; 18(3):1202. https://doi.org/10.3390/ijerph18031202.
22. Walsh CG, Ribeiro JD, Franklin JC. Predicting risk of suicide attempts over time through machine learning. Clinical Psychological Science. 2017 May;5(3):457–69.
23. Bian J, Buchan I, Guo Y, Prosperi M. Statistical thinking, machine learning. J Clin Epidemiol. 2019 Aug 16;116:136–7..
24. Prosperi M, Min JS, Bian J, Modave F. Big data hurdles in precision medicine and precision public health. BMC Med Inform Decis Mak. 2018 Dec 29;18(1):139.
25. Wang F. Why public health needs GIS: A methodological overview. Ann GIS. 2020;26(1):1–12.
26. Weaver A, Lapidos A. Mental Health Interventions with Community Health Workers in the United States: A Systematic Review. J Health Care Poor Underserved. 2018;29(1):159–80.
27. Bidargaddi N, Schrader G, Klasnja P, Licinio J, Murphy S. Designing m-Health interventions for precision mental health support. Transl Psychiatry. 2020 Jul 7;10(1):222.
28. Campion J, Knapp M. The economic case for improved coverage of public mental health interventions. Lancet Psychiatry. 2018;5(2):103–5.