Impact of Area Deprivation Index on rate of hospitalization and quality of treatment for food allergy and food anaphylaxis events

Ayushman Choudhury; Linying Zhang, PhD, Washington University in St. Louis

1. Introduction

Food Allergy (abbreviated as FA) is a type of immune reaction to a food. It is triggered by food protein antigens, and is typically mediated by Immunoglobulin E (abbreviated as IgE). Food Anaphylaxis is a particularly severe and life-threatening form of FA, characterized by its rapid evolution and impacts on multiple systems of the body [1,2].

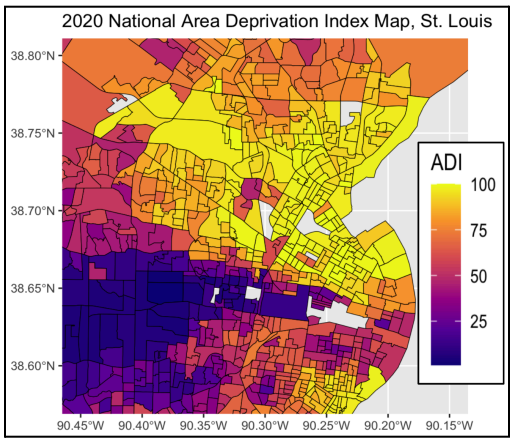
In recent years, observed prevalence rates of IgE-mediated FA have increased worldwide. This increase within a single generation implies that genetic mutations, widely believed to be one among several drivers for food allergies, likely cannot be a significant factor [1,3,4]. Modern food allergy research has focused on how environmental factors (Vitamin D, local biodiversity, air pollution, etc.), social factors (infant exposure to allergens, breastfeeding), and economic factors (access to safe childcare, ability to purchase allergen-free foods, insurance coverage for epinephrine auto-injectors) impact both the diagnosis of FA conditions, and the outcome of severe FA events [3,5,6].

Investigating socio-economic impacts on health outcomes has not been limited to food allergies. Social Determinants of Health (abbreviated as SDOH) are a key focus of the U.S. Department of Health and Human Services's "Healthy People 2030" set of objectives [7], and academic SDOH research has been widespread. In the context of FA, others have investigated, particularly in the United States, whether racial disparities, both institutional and social, may be associated with the aforementioned risk factors [5,6,8].

2. Methods

2.1. Area Deprivation Index

In this paper, we use the neighborhood-level Area Deprivation Index (abbreviated as ADI) as a proxy for socio-economic factors. ADI assigns each US Census Block Group (abbreviated as CBG) a percentile ranking from 1 to 100, where a higher value indicates a more-deprived neighborhood [9,10]. For each patient, we geocode their address — convert the plaintext address to geographic coordinates — using our institutional ArcGIS Server, identify the CBG that contains this coordinate pair using US Census Shapefiles, and assign this CBG's ADI value to the patient. While the US Census claims that Census Tracts are designed to be economically homogeneous, and CBG's are a subdivision of Census Tracts [11], it is important to note that individual disparities are possible within a CBG. However, in the absence of individual socio-economic data, we accept ADI as a proxy. The figure to the left shows the map of CBGs around St. Louis, colored by their ADI.

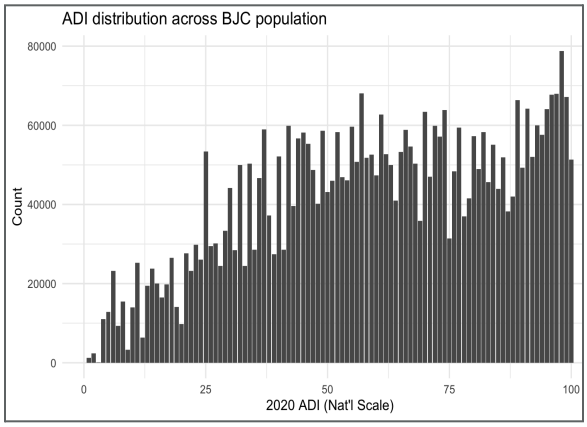


2.2. Data Definitions

Our data is structured following the OHDSI OMOP Common Data Model. This data model is an international standard for organizing observational health data.

The first dataset is person-oriented: a BJC Patient Population, which includes any patient whose location was in a valid CBG in Missouri, and who had a condition_occurrence, visit_occurrence, or observation with a start_date on or after January 1st, 2018. This is due to the limitation that we currently only have access to the records of those who visited BJC Hospital, despite there being several different hospital systems in and around St. Louis. So, when calculating prevalence, we can evaluate it in relation to people within BJC's network of patients who would choose BJC over other hospitals in the region, whether due to proximity, insurance coverage, or other reasons.

One major limitation is that we do not have access to records for people who experience a medical emergency, but do not go to the hospital. There is a concern that this latter group may be skewed socio-economically (one can imagine that a patient who is poorer may be less likely to go to the hospital for a serious FA event). The figure to the right indicates that the BJC Patient Population includes patients from across the spectrum of ADI values, so we feel comfortable continuing this analysis, keeping this caveat in mind.



For each patient, we include their race, age, and sex, and compress them to binary columns: isYoung (born in or after 2000), isMinority (race falls under BIPOC categories), and isFemale. We also include binary columns for whether they've experienced an FA or food anaphylaxis event respectively. These are defined by specific condition_concept_id values in the condition_occurrence table. The food anaphylaxis concept codes are a subset of the FA concept codes, as food anaphylaxis is a type of FA event. As previously described, each patient has an assigned ADI value. We also compress this to a binary column isPoorADI, if their ADI is greater than 50th percentile in the BJC Patient Population.

The second dataset is visit-oriented: an Food Anaphylaxis Visit Occurrence table. This dataset includes all unique visits to BJC for a food anaphylaxis concept code since January 1st, 2018. For each visit, we include the visit length, their admit and discharge locations (converted to binary columns), binary columns for whether they took epinephrine, sodium chloride, prednisone, cetirizine, famotidine, diphenhydramine, or albuterol, and the patient's previously-calculated isYoung, isMinority, isFemale, and isPoorADI.

2.3. Statistical Questions and Approach

We take a Causal Inference approach to investigating this data. Traditionally, observational data only allows for associations between data; to assert causation between a treatment and an outcome requires a randomized controlled trial. RCTs limit the effect of confounding variables by balancing them between the treatment and control groups, whereas observational data can be influenced by these confounding variables. However, in many situations, for example SDOH, RCTs for a treatment like socio-economic status is impossible. Causal Inference provides tools to emulate a RCT from observational data, by controlling for covariate variables and allowing us to extract causal relationships [12].

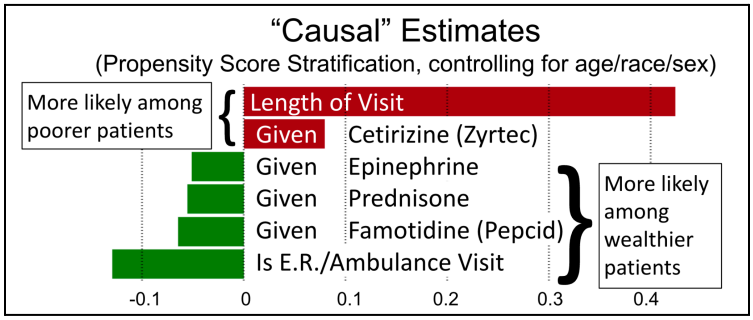
The specific technique that we use is Propensity Score Stratification. For each row in either data table, we define the Propensity Score as the probability that an individual will have a poor ADI, given the binary compression of their race, sex, and age, estimated using a logistic regression. We then divide the rows into K buckets by this score, assert that the individuals in each bucket are comparable, and treat the data as K different RCTs, also known as a stratified randomized experiment. From here, we can apply Neymanian techniques using increasing values of K (for our study: 3, 5, 8, 10, 20, 50, 80) to calculate point and variance estimates for the average causal effect of having a poor ADI on a particular outcome [12]. The limitations of this covariate set are discussed later.

3. Results

	Estimate of "Average Causal Effect"	Variance of estimate
Among the BJC Patient Population, does living in a socio-economically deprived neighborhood (high ADI value) cause a higher likelihood of a FA event occurrence?		
	$-3.4 \cdot 10^{-3}$	$1.3 \cdot 10^{-8}$
Does a CBG being socio-economically deprived (having a high ADI value) cause more patients within the BJC Patient Population to have a FA event occurrence?		
	$-1.9 \cdot 10^{-4}$	$3.0 \cdot 10^{-9}$
For a particular Food Anaphylaxis hospitalization, does the patient living in a socio-economically deprived neighborhood (high ADI value) cause a higher:		
Length of visit?	0.4245680	0.1252157
Likelihood of being admitted via the E.R. or an Ambulance?	-0.129438832	0.001078195
Likelihood of receiving Epinephrine (Epipen/Auvi-Q)?	-0.0514524799	0.0007428435
" " Saline?	0.037618502	0.001071244
" " Prednisone?	-0.055048408	0.001164834
" " Cetirizine (Zyrtec)?	0.079820786	0.001089863
" " Famotidine (Pepcid)?	-0.064628008	0.001234716

“ ” Diphenhydramine?	-0.020574430	0.001168205
“ ” Albuterol (ProAir)?	0.029695707	0.000981889

Part 1 of our analysis showed that a high ADI caused a lower likelihood of FA hospitalization; however, the estimated average causal effects are near 0. Part 2 of our analysis had more significant results — high ADI caused the visit to be 0.42 days longer, increased the likelihood of receiving Cetirizine (Zyrtec), while decreasing the likelihood of receiving “stronger” medications like Epinephrine and Prednisone. High ADI was also associated with patients coming from other healthcare centers, whether different sectors within BJC or other locations altogether, while low ADI was associated with patients being admitted to the E.R. directly from home or by ambulance. The figure to the left is a visual summary of this.



4. Discussion

One of the main assumptions of Causal Inference is ignorability (also called unconfoundedness) — given the covariate set, the potential outcomes of either the treatment or the control are wholly independent of whether the individual unit actually received the treatment or not. This is understood to be true if we have a rich set of covariates, such that there are no unmeasured covariates that affect both the treatment assignment and the outcome [12].

Unfortunately, as stated previously, we currently have access only to hospital records. Ideally, when studying a FA-related hospitalization, it would be important to control on a previous diagnosis of FA, most likely done through a primary care physician, and a record of family history of FA, most likely self-reported. As it stands, our analysis is affected by these unmeasured confounders.

This is in addition to our prior concern on lack of individual economic data. While neighborhood-level ADI is a valuable proxy when this information is not available, it is not a perfect substitute for all the aspects of an individual’s socio-economic condition: their household income, their insurance coverage and access to healthcare resources, their ability to purchase allergen-free meals or access allergen-safe childcare facilities, and other factors. A longer and more in-depth study would benefit from measuring these individual factors.

5. Conclusion

We do not find evidence of a causal effect of ADI on the rate of hospitalization due to a FA event. We find some evidence of a causal effect on the treatment of a food anaphylaxis hospitalization. Patients from more socio-economically deprived areas are more likely to stay in the hospital for longer and receive “weaker” medicines like Zyrtec, but are less likely to receive “stronger” medicines like Epinephrine and Prednisone.

We also acknowledge that fully investigating a causal relationship between socio-economic conditions and FA prevalence would require data beyond the hospital setting. Data on a patients’ past diagnoses through blood or skin testing, family history, or even genetic makeup would improve the covariate set, and thus strengthen the validity of any potential causal relationships.

The ultimate goal is to ascertain what impacts an individual’s socio-economic deprivation might have on the whole range of their experiences with FA: diagnosis, management, hospitalization, and treatment of severe events. By investigating these causal links, we arm medical and governmental institutions with the knowledge to address these socio-economic disparities in FA care, and work towards more equitable FA outcomes.

6. References

1. Yu W, Freeland DMH, Nadeau KC. Food allergy: immune mechanisms, diagnosis and immunotherapy. Nat Rev Immunol. 2016 Dec;16(12):751-765. doi: 10.1038/nri.2016.111. Epub 2016 Oct 31. PMID: 27795547; PMCID: PMC5123910.
2. McLendon K, Sternard BT. Anaphylaxis. [Updated 2023 Jan 26]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482124/
3. Peters RL, Mavoa S, Koplin JJ. An Overview of Environmental Risk Factors for Food Allergy. Int J Environ Res Public Health. 2022 Jan 10;19(2):722. doi: 10.3390/ijerph19020722. PMID: 35055544; PMCID: PMC8776075.
4. Hong X, Tsai HJ, Wang X. Genetics of food allergy. Curr Opin Pediatr. 2009 Dec;21(6):770-6. doi: 10.1097/MOP.0b013e32833252dc. PMID: 19851108; PMCID: PMC2892276.
5. Dupuis R, Phipatanakul W, Bartnikas LM. Social disparities in early childhood prevention and management of food allergy. J Allergy Clin Immunol. 2023 Jan;151(1):37-41. doi: 10.1016/j.jaci.2022.10.015. PMID: 36608981; PMCID: PMC9830563.
6. Trogen, B., Jacobs, S. & Wang, J. Disparities in the Diagnosis and Management of Anaphylaxis. Curr Allergy Asthma Rep 23, 13–19 (2023). https://doi.org/10.1007/s11882-022-01053-y.
7. HHS Office of Disease Prevention and Health Promotion. Social Determinants of Health [Internet]. Rockville, MD: Office of Disease Prevention and Health Promotion; 2024 [cited 2024 August 6]. Available from: https://health.gov/healthypeople/priority-areas/social-determinants-health.
8. Coulson E, Rifas-Shiman SL, Sordillo J, Bunyavanich S, Camargo CA Jr, Platts-Mills TAE, Coull BA, Luttmann-Gibson H, Oken E, Gold DR, Rice MB. Racial, ethnic, and socioeconomic differences in adolescent food allergy. J Allergy Clin Immunol Pract. 2020 Jan;8(1):336-338.e3. doi: 10.1016/j.jaip.2019.06.006. Epub 2019 Jun 22. PMID: 31233938; PMCID: PMC7083174.
9. Kind AJH, Buckingham W. Making Neighborhood Disadvantage Metrics Accessible: The Neighborhood Atlas. New England Journal of Medicine, 2018. 378: 2456-2458. DOI: 10.1056/NEJMp1802313. PMCID: PMC6051533.
10. University of Wisconsin School of Medicine Public Health. 2020 Area Deprivation Index v4. Downloaded from https://www.neighborhoodatlas.medicine.wisc.edu/ July 1, 2024.
11. United States Census Bureau. Census Survey Explorer | Geographies [Internet]. Suitland, MD: United States Census Bureau; 2024 [cited 2024 August 6]. Available from: https://www.census.gov/data/data-tools/survey-explorer/geo.html.
12. Ding P. A first course in Causal Inference. [1st ed.] Chapman & Hall; 2024 July 31. ISBN: 9781032758626.