Social Determinants Predict Low Back and Neck Pain Prevalence in Chicago

Ashrith V. Alavilli[1,2], Scott J. Simmons[2,3], John T. Martin[1];

[1]Rush University, Chicago, IL, [2]Case Western Reserve University, Cleveland, OH, [3]Drury University, Springfield MO; Ashrith.Alavilli@gmail.com; Disc: None

INTRODUCTION: The socioeconomic conditions in Chicago paint a vivid picture of racial and health inequity, reflecting affluent White-majority Northern neighborhoods and under-resourced Black and Hispanic-majority Southern and Western neighborhoods [1]. These neighborhoods vary dramatically in local health metrics — where increased rates of violent crime, cancer, and heart disease [2] drastically reduce life expectancy [3]. In recent work, we demonstrated that social determinants are also primary drivers of musculoskeletal health, demonstrating that income and education define the degree of severity and chronicity in low back pain (LBP) phenotypes [4]. Here, we further explore the relationship between spine disease and the social and physical environment in Chicago. We used health record data from Rush University System for Health (RUSH), which spans multiple healthcare centers across Chicago, and mapped LBP and neck pain (NP) prevalence across Chicago zip codes. We then utilized the Chicago Health Atlas (CHA) to determine which of its 140 zip-code-level socioeconomic metrics drive LBP and NP. We hypothesized that 1) LBP and NP would be more concentrated in Chicago's under-resourced South and West regions and that 2) LBP and NP are correlated to factors that affect physical activity and physical health (green space, walkability, manual labor, food access).

METHODS: *Dataset* We collected patient insurance diagnosis codes (ICD10) and zip code data from RUSH health records from 2014–2024 with Institutional Review Board approval. LBP was defined as any ICD10 code related to low back pain with or without sciatica (M54.40, M54.41, M54.42, M54.50, M54.51, N=33,258). NP was defined as the ICD10 code for cervicalgia (M54.2, N=15,068). Then, we calculated the proportion of LBP and NP patients out of all patients with an encounter at RUSH from 2014–2024 (N=848,571) for each of 57 zip codes (removing two zip codes with less than 100 patients). From the CHA, we collected 140 zip-code-level socioeconomic indicators, covering clinical care, physical environment, morbidity, social and economic factors, health behaviors, mortality, and demographic data. US Census data was also collected for each zip code to determine racial demographics. *Socioeconomic analysis of each zip code* A principal component analysis (PCA) was performed on the CHA indicators to identify the primary sources of variability between zip codes. Zip codes were clustered after PCA-reduction (k-means, N=4 clusters). We identified defining CHA variables for each cluster using the Kruskal-Wallis test followed by a Benjamini-Hochberg correction (false discovery rate = 0.1). *LBP and NP analysis by zip code* Zip codes were categorized into three regions based on their geographical location: North, South, and West. A Kruskal-Wallis test followed by Dunn's test was used to detect the impact of city region (North, South, West) on LBP and NP prevalence. *Identifying key drivers of LBP and NP* We calculated Pearson correlations between the prevalence of LBP or NP and the 140 CHA variables, identifying the top 10 most strongly correlated variables. Then, we further refined the best predictors of LBP and NP from this set by performing random forest regressions with k-fold repeated cross-validation (10 folds, 10 repeats) on every combination of two to six variables. The best predictors were chosen based on model importance metrics.

RESULTS: *Chicago zip codes are divided along racial and poverty lines* North (White-majority), South (Black-majority), and West (Black or Hispanic-majority) city regions are primarily drawn across racial lines (Figure 1A). The North has a lower poverty rate than the South and West (Figure 1B). Unsupervised clustering learned these racial differences in zip code, revealing clusters where the White-, Black-, or Hispanic-majority was the defining feature, and demonstrating expected racial disadvantages in those clusters such as education and morbidity. Principal components analysis of CHA metrics revealed two primary modes of variance, deprivation (PC1, crime and poverty indicators) and morbidity (PC2, hospitalization prevalence) (Figure 1C). White-majority zip codes had lower deprivation (PC1) and morbidity (PC2), while Black-majority zip codes had the highest prevalence of both, and Hispanic-majority zip codes

Figure 1. 1A: Majority race in Chicago by zip code (US Census). 1B: Poverty rate by region (Chicago Health Atlas). 1C: PCA plot with race/cluster labels.

were intermediates with the lowest prevalence of morbidity. One Asian-majority zip code (Chinatown) aligned with the intermediate deprivation/morbidity zip codes. *LBP and NP prevalence are drawn across racial and poverty lines* The North region (White-majority, higher SES) had a significantly lower prevalence of LBP than the South region (Black-majority, lower SES) and West region (mixed Black and Hispanic, lower SES), however there was no difference between the South and West (Figures 2A-B). There were also no differences in NP prevalence by city region (Figures 2C-D). *Social drivers of LBP and NP* Random Forest modeling for predicting LBP prevalence by zip code was successful ($R^2 = 0.74$) where importance metrics ranked the following variables as highly predictive (in order): Childhood Opportunity Index (Education subscale), White-majority population, coronary heart disease rate, and opioid emergency department admits. Random forest modeling for NP was successful to a lesser extent ($R^2 = 0.49$), with predictors: Covid-19 Primary sequence vaccination rate, Medicaid behavioral health professionals per capita, leukemia/lymphoma diagnosis rates, and cervical cancer diagnosis rate.

DISCUSSION: Chicago's healthcare outcomes are influenced by its segregated environment. Using unsupervised learning, we found that social and health factors across Chicago's zip codes divide along racial lines, where Black-majority zip codes have the most deprivation and morbidity, White-majority/Asian-majority zip codes have the least deprivation and moderate morbidity, and Hispanic-majority zip codes are intermediates with low reported morbidity. Following these trends, LBP was

Figure 2. Low back (A, B) and neck pain (C, D) in Chicago (RUSH) by zip code and region (*, p<0.05).

higher in the under-resourced South and West regions and lower in the affluent North region. Childhood education has been previously linked to adult health outcomes [5] and plays a key role in LBP prevalence in Chicago as demonstrated here. Given the lower model performance and modest regional differences, we surmise that NP is less sensitive to social determinants than LBP. Furthermore, factors in the CHA with potential relationships to physical activity and health (green space, walkability, air quality, food access) were not related to LBP or NP, contrary to our hypothesis. While Chicago is an extreme example of segregation, similar trends in LBP and other musculoskeletal disease will likely be observed in other cities. Broader datasets that incorporate other cities and hospital systems should be consulted for a clearer picture.

SIGNIFICANCE: The location one resides in can determine their risk of LBP and (to a lesser extent) NP, where education and racial demographics play major roles in disease prevalence. Musculoskeletal healthcare resources may be more effective when allocated to deprived areas. Still, specific physical mechanisms underlying low back pain and social status remain an area of future work.

REFERENCES: [1]Novara+ *Harvard Press* 2017; [2]Bishop-Royse+ *J Racial Ethn Health Disparities* 2013; [3]Kim+ *Am J Prev Med* 2023; [4]Huang+ *JOR Spine* 202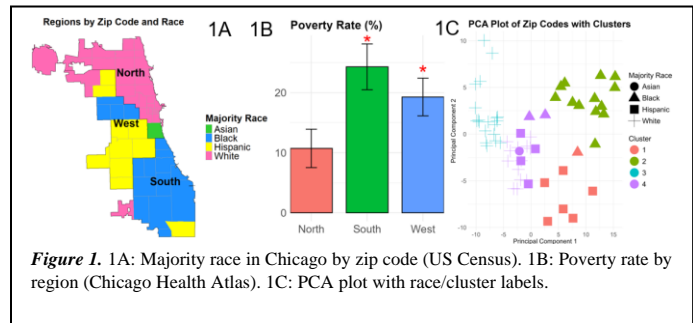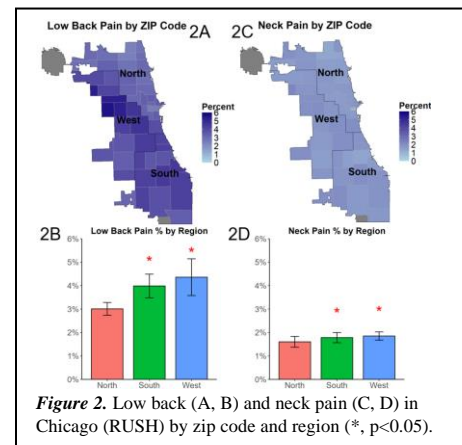3; [5]Campbell+ *Science* 2014.