Supplementary Materials

Table S1. Uni- and multi-variable odds ratios for 30-day readmission from GEE models.

	Univariable GEE (unadjusted)			Univariable GEE (adjusted†)			Multi-variable GEE (adjusted†)		
	OR	**95% CI**	***p***	**OR**	**95% CI**	***p***	**OR**	**95% CI**	***p***
Race-Ethnicity	Reference: Non-Hispanic White								
African American	1.21	1.16-1.26	<0.001	1.18	1.14-1.23	<0.001	1.08	1.04-1.13	<0.001
Hispanic	0.70	0.62-0.78	<0.001	0.86	0.78-0.96	0.005	0.82	0.73-0.91	<0.001
Other	0.77	0.72-0.78	<0.001	0.90	0.85-0.96	0.001	0.90	0.84-0.96	<0.001
ADI (continuous)*	1.04	1.04-1.05	<0.001	1.04	1.03-1.04	<0.001	1.02	1.01-1.02	0.001
Drug Use	1.48	1.37-1.59	<0.001	1.54	1.43-1.65	<0.001	1.35	1.26-1.46	<0.001
Lives Alone	1.42	1.36-1.47	<0.001	1.17	1.13-1.22	<0.001	1.11	1.07-1.16	<0.001
Depression	1.43	1.36-1.50	<0.001	1.25	1.30-1.31	<0.001	1.21	1.16-1.27	<0.001
Dual Eligible	1.82	1.73-1.90	<0.001	1.37	1.31-1.43	<0.001	1.25	1.19-1.31	<0.001
Insurance	Reference: Commercial								
Medicaid	1.42	1.34-1.51	<0.001	1.36	1.28-1.44	<0.001	1.28	1.21-1.36	<0.001
Medicare	2.44	2.34-2.54	<0.001	1.45	1.37-1.53	<0.001	1.26	1.19-1.33	<0.001
Other/Unknown	0.71	0.59-0.86	<0.001	0.77	0.64-0.92	0.004	0.83	0.69-0.99	0.039

Abbreviations: ADI, area deprivation index; GEE, generalized estimating equations; OR, Odds Ratio; 95% CI, 95% confidence interval; *p*, p-value.

*Odds ratio for 10-unit increase in ADI

†GEE models adjusted for age, gender, individual Charlson related comorbidities, and AHRQ CCSR for primary diagnosis categories.

Table S2. Multi-variable odds ratios for 30-day readmission from GEE models by race-ethnicity.

	Multivariable GEE[†] (adjusted & without interactions)											
	African American			**Non-Hispanic White**			**Hispanic**			**Other**		
	OR	**95% CI**	***p***	**OR**	**95% CI**	***p***	**OR**	**95% CI**	***p***	**OR**	**95% CI**	***p***
ADI (continuous)[*]	1.02	1.00-1.03	0.039	1.01	1.01-1.02	0.001	1.06	1.00-1.12	0.039	1.03	1.00-1.05	0.024
Drug Use	1.35	1.21-1.50	<0.001	1.32	1.19-1.46	<0.001	1.62	1.06-2.45	0.024	1.38	0.95-2.02	0.092
Lives Alone	1.10	1.02-1.18	0.009	1.12	1.06-1.17	<0.001	1.20	0.91-1.59	0.193	1.08	0.89-1.29	0.445
Depression	1.22	1.12-1.35	<0.001	1.18	1.11-1.24	<0.001	1.51	1.12-2.02	0.006	1.46	1.18-1.81	<0.001
Dual Eligible	1.25	1.16-1.36	<0.001	1.24	1.15-1.32	<0.001	1.14	0.85-1.52	0.391	1.29	1.07-1.56	0.009
Insurance	Reference: Commercial											
Medicaid	1.29	1.17-1.43	<0.001	1.32	1.21-1.43	<0.001	0.95	0.72-1.26	0.736	1.06	0.86-1.31	0.587
Medicare	1.22	1.10-1.37	<0.001	1.31	1.22-1.41	<0.001	0.92	0.62-1.35	0.665	1.02	0.81-1.28	0.856
Other/Unknown	0.63	0.44-0.89	0.010	0.93	0.74-1.18	0.560	0.58	0.29-1.19	0.136	1.01	0.49-2.09	0.969

Abbreviations: ADI, area deprivation index; GEE, generalized estimating equations; OR, Odds Ratio; 95% CI, 95% confidence interval; p, p-value.

[*]Odds ratio for 10-unit increase in ADI

[†]For each race-ethnicity group, multi-variable GEE model adjusted for age, gender, individual Charlson related comorbidities, and AHRQ CCSR for primary diagnosis categories.

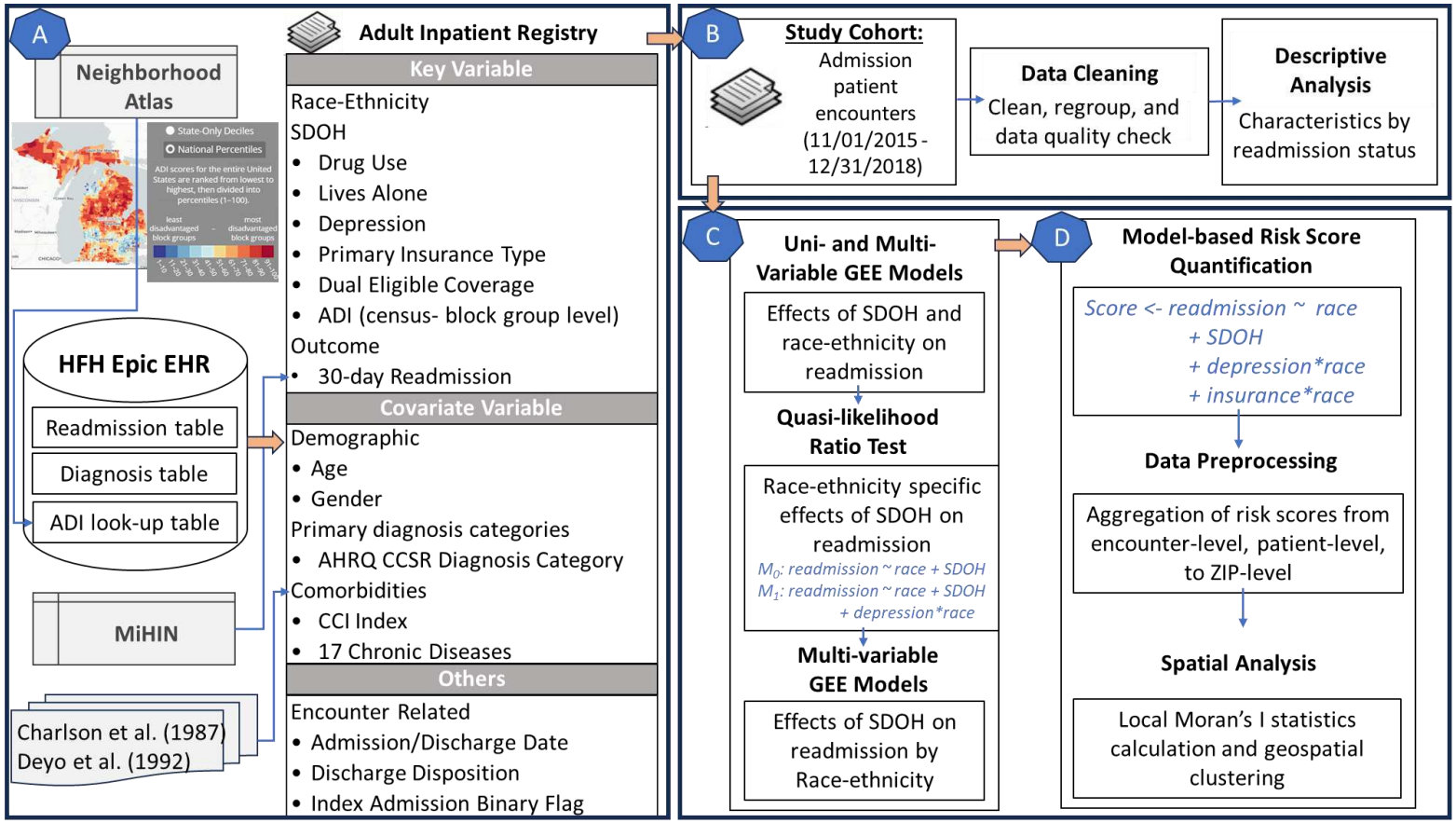

Figure S1. The overview and workflow of the methods

Abbreviations: ADI, area deprivation index; AHRQ CCSR, Agency for Health Care Research and Clinical Classifications Software Refined Categories for primary diagnosis; CCI, Charlson Comorbidity Index; EHR, electronic health record; GEE, generalized

estimating equations; HFH, Henry Ford Health; MiHIN, Michigan Health Information Network; SDOH, Social Determinants of Health.

Phase A – Data Collection Process: The observational retrospective study was conducted using a subset of data elements from the HFH adult inpatient registry which includes demographics, SDOH, comorbidities (CCI and 17 comorbid conditions), and primary diagnosis categories. The primary data source for creating the inpatient registry is the HFH Epic EHR systems which includes the readmission, diagnosis, and ADI lookup tables. The ADI data in the EHR, known as the Neighborhood Atlas, originates from the Center for Health Disparities Research: University of Wisconsin-Madison School of Medicine and Public Health (https://www.neighborhoodatlas.medicine.wisc.edu/). The weighted index, CCI, was calculated using 17 comorbid conditions, which were coded as binary flags based on patients' primary diagnosis records from one year prior to their admission date. The methods for calculating and defining the CCI are detailed in the publications by Charlson et al., 1987 (Reference No.25) and Deyo et al., 1992 (Reference No.26). Furthermore, the method for converting primary diagnosis extracted from the EHR into the AHRQ CCSR diagnosis categories can be found on the AHRQ website (https://www.hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp).

Phase B – Cohort Identification, Data Cleaning, and Data Summary: Patient encounters with admission dates between November 2015 and December 2018 were included. Those identified as non-qualified index admissions were excluded from this study. Following data cleaning and regrouping, as described in the "Key Variable Definitions" section under METHODS, a descriptive analysis was conducted. This analysis reported frequency count and percentages for categorical variables (gender, race-ethnicity, ADI quartiles, drug use, lives alone, depression, dual eligible, insurance type, 17 individual chronic condition indicators, and top 15 most

frequent AHRQ CCSR diagnosis categories). It also reported means and standard deviation for continuous variables (age at admission, ADI national rank, and CCI) overall and by 30-day readmission status.

Phase C – Effect Testing Using GEE Models: Uni- and multi-variable GEE models were applied to assess whether SDOH and race-ethnicity were associated with readmission. In the uni-variable GEE models, both unadjusted and adjusted methods were conducted. The former included only one key variable (e.g. *readmission ~ race* or *readmission ~ drug use*) and the latter included one key variable along with a common set of potential confounders including age, gender, individual Charlson-related comorbidities, and AHRQ CCSR for primary diagnosis categories (e.g. *readmission ~ drug use + potential confounders*). Additionally, the Quasi-likelihood Ratio Test was conducted to examine whether the effects of SDOH were dependent upon race-ethnicity by testing the difference between two multi-variable GEE models: one without (based model, M_0*: readmission ~ race + SDOH + potential confounders*) and one with multiplicative SDOH-by-race-ethnicity interaction terms (M_1*: readmission ~ race + SDOH + one SDOH* [e.g., *depression*]**race + potential confounders*). Furthermore, the data were split into four groups based on race-ethnicity, Non-Hispanic White, African American, Hispanic, and Other. Within each subgroup, a multi-variable GEE model was conducted to assess SDOH effects on readmission.

Phase D – Identification of Geospatial Clustering Using Spatial Analysis: A multi-variable GEE model was constructed utilizing all significantly associated variables including SDOH, race-ethnicity, and their multiplicative interactions (*readmission ~ race + SDOH + depression*race + insurance*race + potential confounders*). The quantified risk score for each patient encounter was calculated based on natural logarithm of their odds of readmission. These encounter-level scores were then linked to the individual

patient's residential ZIP code within the three main counties served by HFH—Wayne, Oakland, and Macomb in Michigan—to calculate ZIP code-level scores. These ZIP code-level scores aggregated based on median patient-level scores mapping to each ZIP were subsequently used to determine geospatial clusters by applying spatial autocorrelation analysis. To identify geospatial clustering, Local Moran's I statistics were calculated using the ZIP code-level scores.