Research and Applications

Assessing socioeconomic bias in machine learning algorithms in health care: a case study of the HOUSES index

Young J. Juhn[1,2], **Euijung Ryu**[3], **Chung-Il Wi**[1,2], **Katherine S. King** [iD][3],
Momin Malik [iD][4], **Santiago Romero-Brufau** [iD][5], **Chunhua Weng**[6], **Sunghwan Sohn**[7],
Richard R. Sharp[8], and **John D. Halamka** [iD][4,9]

[1]Precision Population Science Lab, Mayo Clinic, Rochester, Minnesota, USA, [2]Artificial Intelligence Program of Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, Minnesota, USA, [3]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota, USA, [4]Center for Digital Health, Mayo Clinic, Rochester, Minnesota, USA, [5]Department of Internal Medicine, Mayo Clinic, Rochester, Minnesota, USA, [6]Department of Biomedical Informatics, Columbia University, New York, New York, USA, [7]Department of Artificial Intelligence and Informatics, Mayo Clinic, Rochester, Minnesota, USA, [8]Biomedical Ethics Program, Mayo Clinic, Rochester, Minnesota, USA, and [9]Mayo Clinic Platform, Rochester, Minnesota, USA

Corresponding Author: Young J. Juhn, MD, MPH, Precision Population Science Lab, Mayo Clinic, Rochester, MN, USA; juhn.young@mayo.edu; Euijung Ryu, PhD, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA; ryu.euijung@mayo.edu

Young J. Juhn and Euijung Ryu contributed equally to this work.

Received 8 October 2021; Revised 24 March 2022; Editorial Decision 25 March 2022; Accepted 5 April 2022

ABSTRACT

Objective: Artificial intelligence (AI) models may propagate harmful biases in performance and hence negatively affect the underserved. We aimed to assess the degree to which data quality of electronic health records (EHRs) affected by inequities related to low socioeconomic status (SES), results in differential performance of AI models across SES.

Materials and Methods: This study utilized existing machine learning models for predicting asthma exacerbation in children with asthma. We compared balanced error rate (BER) against different SES levels measured by HOUsing-based SocioEconomic Status measure (HOUSES) index. As a possible mechanism for differential performance, we also compared incompleteness of EHR information relevant to asthma care by SES.

Results: Asthmatic children with lower SES had larger BER than those with higher SES (eg, ratio = 1.35 for HOUSES Q1 vs Q2–Q4) and had a higher proportion of missing information relevant to asthma care (eg, 41% vs 24% for missing asthma severity and 12% vs 9.8% for undiagnosed asthma despite meeting asthma criteria).

Discussion: Our study suggests that lower SES is associated with worse predictive model performance. It also highlights the potential role of incomplete EHR data in this differential performance and suggests a way to mitigate this bias.

Conclusion: The HOUSES index allows AI researchers to assess bias in predictive model performance by SES. Although our case study was based on a small sample size and a single-site study, the study results highlight a potential strategy for identifying bias by using an innovative SES measure.

Key words: artificial intelligence, HOUSES, algorithmic bias, electronic health records, social determinants of health

INTRODUCTION

Augmented computing power, storage capability, and predictive analytics have accelerated the adoption and deployment of artificial (augmented or machine) intelligence (AI) tools in health care system of the United States.[1,2] As of 2017, 96% of US hospitals adopted certified electronic health records (EHRs) and 98% of US hospitals demonstrated meaningful use of at least one certified health information technology.[1] A recent survey showed that 90% of health care executives in the United States reported they had AI tools and automation strategy in 2020 compared to 53% in 2019.[2] As of 2020, more than 80 imaging-related AI algorithms have been cleared by the US Food and Drug Administration (FDA).[3,4] A research group at Mayo Clinic showed that an electrocardiogram-based, AI-powered clinical decision support tool demonstrated early diagnosis of low ejection fraction, a condition that is underdiagnosed but treatable, via a randomized clinical trial (RCT).[5] The intervention increased the diagnosis of low ejection fraction in the overall cohort (1.6% in the control arm versus 2.1% in the intervention arm, odds ratio [OR] 1.32 (1.01–1.61), $P = .007$).

Health care research continues to try and develop AI tools to improve health and reduce disparities (eg, reducing unexplained self-reported pain disparities by applying image-based AI algorithm in underserved populations).[6] However, AI systems can introduce or reify biases and negatively influence health in under-resourced or racial/ethnic minority populations.[7–10] For example, a recent study of a widely used commercial AI algorithm based on health care costs identified only 18% of Black patients needing additional care for chronic disease management, compared to 47% after controlling for chronic historical under-spending on Black patients.[7] This demonstrated that building predictive models from easily available proxies (health care spending, rather than health) without considering possibly differential performance by marginalization status has the potential to reify bias, but also that appropriate analysis may be able to mitigate this bias. Others reported biases in model performance by race and socioeconomic status (SES) in predicting post-partum depression (logistic regression),[11] intensive care unit (ICU) mortality (logistic regression),[12] and 30-day psychiatric readmission (logistic regression).[12] When models have differential predictive performance by patient characteristics such as SES,[13] the adoption of AI in clinical care on a large scale risks exacerbating inequities.[14–16]

Given the significant associations of SES with health risk and health care access especially driven by upstream social determinants of health (SDH),[17,18] quantifying the degree of bias in differential model performance by SES has important ethical implications for AI research as well as health care delivery research. SES is a key element of SDH in health care delivery and research[19–28] and is considered a major factor accounting for differential health outcomes through broad and fundamental mechanisms, from biological (eg, epigenetics, gene expression or telomere length) and behavioral (eg, nonadherence and stress) factors to environmental factors (eg, indoor [eg, molds and other allergens] or outdoor environment [eg, high traffic volume]).[20,28–31] We investigate a specific mechanism: if EHRs are less comprehensive or complete for those of lower SES due to limited health care access (eg, less access to health care resources), and AI models are developed from EHRs and rely on data completeness and quality for predictive success, then the straightforward adoption of AI has the potential to discriminate against those of lower SES.

Unfortunately, individual-level SES measures that are validated, reliable, and scalable are frequently unavailable in commonly-used data sources for clinical care and research,[32] posing a major barrier to health care delivery and research as acknowledged by National Academy of Medicine and the National Quality Forum.[9,33,34] Additionally, the limited availability of suitable individual-level SES in EHRs is a major challenge in studying possible bias in the adoption of AI systems. Consequently, current AI fairness research is limited to considering readily available demographic factors such as age, sex, and race/ethnicity,[35] leaving the role of SES in AI bias (on its own, or in interactions with other factors) poorly understood.

To address this major roadblock in the equitable implementation of health care AI, we propose using the HOUSES (individual HOUsing-based SES) index as a measure of SES with important features (validity, precision, objectivity [instead of self-report] and scalability) that can be integrated with AI model development. In this work, we (1) assessed differential data availability and validity of EHRs among study subjects according to SES as measured by HOUSES and (2) applied HOUSES index to quantify bias in commonly used metrics of model performance by SES. Thus, the goal of this study is to demonstrate the importance of considering SES in AI research; we plan to do a future study to create recommendations of specific mitigation steps using SES.

MATERIALS AND METHODS

Study population and setting

The study population and setting were described in our recent report.[36] Briefly, Olmsted County is a virtually self-contained health care environment with only 2 health care systems providing clinical care to nearly all residents. About 98% of residents authorize the use of their medical records for research.[37] According to 2010 US Census data, the age, sex, and ethnic characteristics of Olmsted County residents were similar to those of the state of Minnesota and the Upper Midwest.[38] However, Olmsted County has become more diverse as indicated by the racial/ethnic characteristics of children enrolled in public schools (in 2019, 35.2% reported belonging to a racial/ethnic minority group). Mayo Clinic Primary Care Pediatric Practices offers primary care service at 4 locations within Olmsted County. This study was conducted at the Baldwin Primary Care Practice site, the largest of the 4 practice sites (ie, including teaching pediatric faculty, residents, and nurse practitioners). In Olmsted County, Asthma is the most prevalent chronic illness, with the third highest health care expenditures in children and adolescents.[39] The asthma prevalence in the primary care practice (14%) is slightly lower than that of the county overall (17.6%).[40]

Study design and subjects

The design and subjects of the original study as an RCT were described in our recent report.[36] Briefly, the present study was designed as a cross-sectional study. Patients were randomly assigned to 2 independent cohorts (ie, training and testing cohorts) for the development of machine learning (ML) models. A previous study developed a novel AI-assisted clinical decision support system named A-GPS (Asthma-Guidance and Prediction System), based on a single-center pragmatic RCT. An A-GPS-based intervention in the original study provided clinicians with the summary of most relevant information for asthma management, along with a prediction of asthma exacerbation (AE) made using the trained ML models.[36] Despite providing a great deal of actional guidance and intervention information, it overall significantly reduced clinicians' EHR review burden, resulting in more efficient asthma management. The focus

of the original study was to assess the effectiveness of the use of A-GPS on asthma outcomes (eg, AE, asthma control, asthma-related health care utilization, asthma care quality and health care costs). The original study used data from subjects who had persistent asthma or those who met Predetermined Asthma Criteria (PAC). In this present report, we limited the primary analysis assessing algorithmic bias to those with persistent asthma in order to focus on a more homogeneous patient group. Details of the original study have been reported.[36] For additional analysis, we used the subjects in the original study who met PAC definition but were not yet diagnosed with asthma at the time of enrollment. This study (IRB number: 15-004435) was approved by the Mayo Clinic Institutional Review Board (IRB).

ML models for estimating AE risk

For A-GPS, we trained and tested 2 ML models: Naïve Bayes (NB) and gradient boosting machine (GBM) for binary classification for estimating 1-year AE risk among children with asthma. We extracted 29 candidate variables based on the literature including sociodemographics, risk factors, and asthma outcomes from EHR over a prior 3-year period. The regularization of NB model selected the following 5 variables as the most informative (previous exacerbation, asthma symptom, hospital visit due to asthma, rescuer medication, and controller medication) and used them when testing the model performance. GBM model used variables with relative influence score at least 1% that included a broader range of variables: most of the variables from NB model (eg, asthma symptoms and previous exacerbation) and sociodemographic factors (eg, race and HOUSES quartiles). The original study included 590 subjects (300 in the training and 290 in the test set) who had persistent asthma or met PAC from a Mayo Clinic pediatric practice panel, respectively. Receiver operating characteristic areas under the curve for NB and GBM model were 0.78 and 0.74 on the testing cohort, respectively. This report calculated differential performance on these models' test performance by SES.

Fairness metrics

We considered common metrics for assessing fairness in model performance: accuracy equality (equal accuracy across groups), equal opportunity (equal false negative rate [FNR] across groups), predictive equality (equal false positive rate [FPR] across groups), and predictive parity (equal precision across groups). As it is impossible for a model to simultaneously satisfy equal opportunity, predictive equality, and predictive parity ('impossibility theorem'),[41,42] and there is no agreed-upon gold standard metric to be used, we prioritized balanced error rate (BER),[43] defined as the unweighted average of the FPR (predictive equality) and FNR (equal opportunity), as the primary metric for assessing bias in this presented work (see more details in Supplementary Table S1). BER was chosen as the primary metric because our focus in the current work was prediction accuracy, which involves both FPR (or 1-specificity) and FNR (or 1-sensitivity). We decided to use the unweighted (ie, equal weights) average for summarizing both metrics, because the relative importance of these metrics will likely depend on the purpose of the studies. While our rationale for the use of BER is supported by literature[41,42] and we use it as the primary measure of fairness, we also present results of each metric separately to see which metric is more meaningful in a given study.

For each metric, we calculated the ratio comparing least privileged group (eg, HOUSES Q1 representing lower SES, see below) with the privileged group (HOUSES Q2–Q4 representing higher SES). For FPR and BER, a ratio >1 means that the model performance is superior for the privileged group, while a ratio >1 for the other 3 metrics (accuracy equality, equal opportunity, and predictive parity) means the model performance is superior for the less privileged group. As a rule of thumb, a ratio that is <0.8 or >1.25 (1/0.8) is considered as meaningful difference, which is implemented in the open source program AI Fairness 360.[44]

Socioeconomic measures

In this study, we included 2 SES measures: HOUSES and area deprivation index (ADI). HOUSES is an individual-level SES measure based on 4 real property data variables of an individual housing unit after principal component factor analysis: *housing value, square footage, number of bedrooms, and number of bathrooms.* An individual's address from the EHR is directly linked to the publicly available assessor's data (which is a basis for property tax and thus is available throughout US counties and cities).[32] We formulated a standardized HOUSES index score by summing these variables after z-score transformation. The greater the HOUSES index, the higher the SES. Since its development, HOUSES has been extensively applied as a validated SES measure that has shown association with numerous (39 different) health-related outcomes, including acute/chronic conditions, health care access issues, health care utilization, and other health-related behaviors such as smoking, and vaccine status as summarized in Supplementary Table S2 representing 23 published reports. As an alternative SES measure, we included ADI in the analysis to compare the relative utility of HOUSES and ADI in assessing bias in model performance by SES. ADI is a widely used aggregate-level SES measure in clinical research, and it can use smaller geographic units (Census Block Groups).[45] We used national-level ADI and categorized subjects into 2 groups: the highest ADI quartile (ADI: 76–100; lower SES) and lower ADI quartiles (ADI: 0–75; higher SES).

Other pertinent variables

While our focus is to quantify bias in model performance by SES, we also considered other readily available demographic characteristics (age, sex, and race/ethnicity), and pediatric chronic conditions defined by Feudtner et al (an accepted measure of pediatric chronic conditions in literature).[46,47] While it is possible that model performance may differ among racial groups (eg, Asians vs African Americans), the current study cohort does not have a large enough sample size of each group to do a separate analysis. Therefore, the race variable is collapsed into "white" and "other". These variables are extracted from patient's EHR. For chronic conditions, ICD-9 diagnostic and procedure codes were used. For simplicity, we dichotomize age (<12 vs ≥12 years) and chronic conditions (yes vs no). To demonstrate the association of SES with completeness of EHR as a potential reason for differential model performance by SES, we compared availability of 7 variables that are clinically relevant to childhood asthma management (health maintenance visit, asthma compliance, asthma severity, asthma type, National Asthma Education and Prevention Program (NAEPP) recommendation, smoking status, and missing school). These variables were extracted from EHR in the 3 years prior to the study index date. Additionally, we assessed data validity by looking at ICD-9 codes for asthma among those who met PAC definition but were not yet diagnosed with asthma at the time of the study.[36] Specifically, we previously reported a significant number of children with undiagnosed asthma

by comparing asthma prevalence by ICD code-based asthma ascertainment with that by natural language processing (NLP)-based ascertainment using PAC (sensitivity: 31% for ICD-9 vs 81% for NLP using criteria-based logic and 85% for NLP using ML).[48–50]

Data analysis

In this presented work, we quantified algorithmic bias for 2 ML models (NB and GBM) for estimating 1-year AE risk among pediatric asthmatics by demographic factors (age, sex, race/ethnicity), SES (HOUSES and ADI), and chronic condition. For race/ethnicity variable, all non-Hispanic Whites were classified as "Others" when assessing algorithmic bias, due to small sample sizes in each minority category. This was done using a separate testing cohort whose data were not used in model training, to avoid overestimates of out-of-sample performance. To see the association of SES with data availability and completeness of EHR, we also calculated proportions of subjects with missing or unknown information for 7 variables relevant to asthma management. This analysis was done using HOUSES only, because the number of subjects with the lowest SES measured by ADI was very small. Based on our earlier work, we focused on assessing one variable as the main measure of data accuracy, diagnosed vs. undiagnosed asthma by ICD codes for those who met PAC.[49,50] This calculation was done in both the training and testing cohorts.

RESULTS

Subject characteristics

The training cohort consisted of subjects with 71% being <12 years old and 57% males. For race/ethnicity, a large portion of subjects (60%) were non-Hispanic White and 14% were African American as shown in Table 1. Roughly 20% of the subjects were in the low-SES (HOUSES Quartile 1, Q1) group and 20% had at least one chronic condition. However, the proportions of subjects with lower SES by ADI were only 7% in training and 8% in testing cohorts. Subject characteristics were similar between training and testing cohorts. Roughly 30% of subjects had AE within 1-year follow-up period (26% in the training cohort and 35% in the testing cohort: Table 3). Table 2 showed that proportion of AE differed by subject characteristics. In general, the proportion was higher in subjects who were younger, male, lower SES by HOUSES, and those with chronic conditions. There was significant discrepancy in the proportion of subjects with a history of AE among lower SES group defined by HOUSES (53%) and ADI (0%) in testing cohort.

Bias in performance

Using the testing cohort, Table 3 summarizes the results of bias in model performance for both NB and GBM models in estimating 1-year AE risk. Overall, model performance was not independent of patient characteristics such as age, sex, and chronic diseases as expected. Also, the 2 models did not have systematically different patterns compared to one another in how their performance differed by these factors. Higher SES as measured by HOUSES index was *greatly* associated with superior model performance. Specifically, children in lower SES groups had higher BERs than those in the higher SES group in both ML models (ratio = 1.35 for NB model and 1.25 for GBM model) which exceed those for race/ethnicity (1.23 and 1.04, respectively). This differential performance by SES was driven more by FNR (=1-sensitivity; ratio = 1.51 by NB and 2.01 by GBM model) than FPR (1.18 by NB and 0.92 by GBM

Table 1. Subject characteristics used in the study

	Training cohort (N = 133)	Testing cohort (N = 113)
Age (in years), *n* (%)		
<12	94 (71%)	80 (71%)
≥12	39 (29%)	33 (29%)
Sex, *n* (%)		
Male	76 (57%)	67 (59%)
Female	57 (43%)	46 (41%)
Race/ethnicity, *n* (%)		
Non-Hispanic Whites	76 (60%)	67 (60%)
African Americans	18 (14%)	9 (8%)
Asians	10 (8%)	13 (12%)
Hispanics	9 (7%)	11 (10%)
Other categories	14 (11%)	12 (11%)
Missing	6	1
HOUSES, *n* (%)		
Q1 (the lowest SES)	22 (18%)	15 (14%)
Q2–Q4	102 (82%)	92 (86%)
Missing	9	6
Chronic condition, *n* (%)		
Yes	30 (23%)	19 (17%)
No	103 (77%)	94 (83%)
National ADI, *n* (%)		
76–100 (the lowest SES)	6 (7%)	6 (8%)
0–75	76 (93%)	65 (92%)
Missing	51	42
Asthma exacerbation, *n* (%)		
Yes	34 (26%)	40 (35%)
No	99 (74%)	73 (65%)

model). This was also true for the equal opportunity (ie, sensitivity) metric. Children in the higher SES group had significantly higher sensitivity in the performance of both models, compared to those in the lower SES group, to a greater extent than the difference by other demographic factors. The bias analysis using ADI was limited due to the lack of children experiencing AE among those having the lowest SES measured by ADI in the testing cohort. For example, 2 of 5 metrics (equal opportunity and BER) used were not computable because the denominator was zero. Also, PPV for those with ADI > 75 was zero because the numerator was zero.

Availability and accuracy of data relevant to asthma management

We compared data availability for the key variables associated with the risk of AE in the training and testing cohorts. As shown in Table 4, compared to children in the higher SES group, those from lower SES background had lower availability of the key variables for asthma (eg, compliance data, severity and smoking exposure) associated with the risk of AE. Additionally, children with lower SES had higher prevalence of undiagnosed asthma (ie, data inaccuracy), compared to those with higher SES, although they met the criteria for asthma.

DISCUSSION

Our study results suggest that lower SES, as measured by the HOUSES index, is associated with worse predictive model performance. A possible mechanism for this bias in performance is *incomplete and inaccurate* EHR data, as AI models perform better with larger amounts of and more accurate data, and we found unavail-

Table 2. Proportion of subjects with asthma exacerbation (AE) by subject characteristics

	Training cohort (N = 133)		Testing cohort (N = 113)	
	Subjects with AE (N = 34)	Subjects without AE (N = 99)	Subjects with AE (N = 40)	Subjects without AE (N = 73)
Age (in years), *n* (%)				
<12	28 (29.8%)	66 (70.2%)	30 (37.5%)	50 (62.5%)
≥12	6 (15.4%)	33 (84.6%)	10 (30.3%)	23 (69.7%)
Sex, *n* (%)				
Male	25 (32.9%)	51 (67.1%)	23 (34.3%)	44 (65.7%)
Female	9 (15.8%)	48 (84.2%)	17 (39.1%)	28 (60.9%)
Race/ethnicity, *n* (%)				
Non-Hispanic Whites	19 (25.0%)	57 (75.0%)	25 (37.3%)	42 (62.7%)
African Americans	5 (27.8%)	13 (72.2%)	4 (44.4%)	5 (55.6%)
Asians	2 (20.0%)	8 (80.0%)	3 (23.1%)	10 (76.9%)
Hispanics	4 (44.4%)	5 (55.6%)	3 (27.3%)	8 (72.7%)
Other categories	4 (28.6%)	10 (71.4%)	4 (33.3%)	8 (66.7%)
HOUSES, *n* (%)				
Q1 (the lowest SES)	6 (27.3%)	16 (72.7%)	8 (53.3%)	7 (46.7%)
Q2–Q4	23 (22.5%)	79 (77.5%)	29 (31.5%)	63 (68.5%)
Chronic condition, *n* (%)				
Yes	10 (33.3%)	20 (66.7%)	7 (36.8%)	12 (63.2%)
No	24 (23.3%)	79 (76.7%)	33 (35.1%)	61 (64.9%)
National ADI, *n* (%)				
76–100 (the lowest SES)	2 (33.3%)	4 (66.7%)	0 (0.0%)	6 (100.0%)
0–75	16 (21.1%)	60 (78.9%)	21 (32.3%)	44 (67.7%)

ability and inaccuracy also associated with lower SES. In turn, this means adopting AI models biased by SES systematically aggravates inequity, alongside greater health risk and lower health care access. One noteworthy finding is disparities in undiagnosed or delayed diagnosed asthma by SES, as the lack of timely diagnosis of asthma will deter access to preventive and therapeutic interventions[51,52] and may influence long-term respiratory outcomes.

As discussed earlier, SES is a key variable for understanding the nature of bias stemming from differential health risk, health care access, and completeness of available EHRs and for assessing and mitigating algorithmic bias in health care. However, objective, scalable, and well-validated individual-level SES measures are unavailable in commonly used data sources for clinical care and research[32] posing a major barrier to health care delivery and research as acknowledged by National Academy of Medicine and National Quality Forum.[9,33,34] In this respect, using the HOUSES index as a measure of individual-level SES can be a useful tool for health care research, including AI research, as it overcomes such unavailability of individual-level SES measures in commonly used data sources such as EHRs.

Our previous work demonstrated that SES defined by HOUSES index correlated with a broad range of health outcomes and care quality as summarized in Supplementary Table S2. Relevant to this present report, we showed that HOUSES was associated with inconsistent self-reporting.[53] We found that lower HOUSES (SES) was associated with higher rates of inconsistency (inaccuracy) in self-reporting a diagnosed disease for the given (documented) diseases between the baseline and 4-year follow-up survey, and the association remained significant after pertinent characteristics such as age and perceived general health (adjusted OR = 1.46; 95% confidence interval [CI] 1.17–1.84 for the lowest compared with the highest HOUSES decile). Given that self-reported information is captured in EHR and often used clinically (eg, a history of pediatric asthma), higher proportion of inconsistent self-

reporting among patients with low SES may produce less reliable ML models (if used). For the findings in Table 4 indicating differential completeness of EHRs pertaining to childhood asthma by SES, it is widely recognized that people with lower SES have greater burden of diseases and poor outcomes compared to those with higher SES,[20] especially childhood asthma.[54–56] It is also well documented that those with lower SES have limited health care access, may not have a usual source of care, or rely more upon safety net care such as emergency department, compared to those with higher SES[57–59] (also see Supplementary Table S2 summarizing differential burden of disease and health care access by SES as measured by HOUSES). For example, our unpublished data showed that the availability of patient's online portal system (a proxy for health care access) was significantly lower among families with lower SES (68% in Q1 [lowest SES]), compared to 74% in Q2, 88% in Q3, and 92% in Q4 (highest SES) (P = .02). As an online portal is an important tool for managing chronic diseases such as childhood asthma (eg, with communications with care providers, patient-reported outcomes [PROs], medication updates, etc, being captured in EHRs), it significantly affected availability of a key PROs on asthma (ie, Asthma Control Test results; 99% for those with portal vs 77% for those without portal) at the end of a clinical trial as supported by this present study (see Table 4). Populations at high risk for poor outcomes are characterized by a mismatch (called cumulative complexity model)[60–62]: despite a higher burden of diseases, families with lower SES often also face limited health care access compared to those of higher SES. This mismatch model provides a useful framework for assessing and mitigating AI bias by SES.

Our study results in Table 3 show the potential association of SES as measured by HOUSES with biases in model performance. For example, BERs were higher for children with lower SES for both algorithms estimating AE risk, compared to those with higher SES, with a disparity larger than those associated with other demographic factors

Table 3. Assessment of algorithmic bias for 2 machine learning models (Naïve Bayes [NB] and gradient boosting machine [GBM]) estimating 1-year asthma exacerbation risk in childhood asthma using 5 commonly used bias metrics

Groups	Accuracy equality NB model	Accuracy equality GBM model	Equal opportunity (sensitivity) NB model	Equal opportunity (sensitivity) GBM model	Predictive parity (PPV) NB model	Predictive parity (PPV) GBM model	Predictive equality (FPR) NB model	Predictive equality (FPR) GBM model	Balanced error rate ([FPR + FNR]/2) NB model	Balanced error rate ([FPR + FNR]/2) GBM model
SES (HOUES)										
Q1 (lowest SES)	0.47	0.47	0.38	0.50	0.50	0.50	0.43	0.57	0.53	0.54
Q2–Q4	0.62	0.50	0.59	0.76	0.43	0.36	0.37	0.62	0.39	0.43
Ratio (Q1/Q2–4) (1 = no diff)	**0.75**	0.93	**0.64**	**0.66**	1.18	**1.39**	1.17	0.92	**1.35**	**1.25**
Age										
<12	0.53	0.45	0.57	0.70	0.41	0.38	0.50	0.70	0.47	0.50
≥12	0.76	0.64	0.40	0.80	0.67	0.44	0.09	0.44	0.34	0.32
Ratio (<12/≥12) (1 = no diff)	**0.69**	**0.71**	**1.42**	0.88	**0.61**	**0.84**	**5.75**	**1.61**	**1.36**	**1.57**
Sex										
Male	0.49	0.45	0.48	0.78	0.33	0.36	0.50	0.73	0.51	0.47
Female	0.74	0.59	0.59	0.65	0.67	0.46	0.17	0.45	0.29	0.40
Ratio (male/female) (1 = no diff)	**0.67**	**0.76**	0.81	**1.21**	**0.50**	**0.79**	**2.90**	**1.62**	**1.75**	1.18
Race/Ethnicity										
Others	0.54	0.39	0.47	0.60	0.35	0.29	0.42	0.71	0.48	0.56
Non-Hispanic White	0.63	0.58	0.56	0.80	0.50	0.47	0.33	0.55	0.39	0.37
Ratio (others/White) (1 = no diff)	0.87	**0.67**	0.83	**0.75**	**0.70**	**0.62**	**1.26**	**1.30**	**1.23**	**1.48**
Chronic condition										
At least one	0.53	0.47	0.20	0.80	0.20	0.33	0.33	0.67	0.57	0.43
None	0.61	0.50	0.59	0.69	0.46	0.39	0.38	0.60	0.39	0.46
Ratio (≥1/none) (1 = no diff)	0.87	0.94	**0.34**	1.16	**0.43**	0.86	0.88	1.11	**1.44**	0.95
ADI										
76–100	0.60	0.60	NC	NC	0.00	0.00	0.40	0.40	NC	NC
0–75	0.64	0.54	0.60	0.80	0.44	0.39	0.35	0.58	0.37	0.39
Ratio (76–100/0–75) (1 = no diff)	0.95	1.11	NC	NC	0.00	0.00	1.15	0.69	NC	NC

NC: not computable.

Ratios either greater than 1.2 or less than 0.8 (ie, an absolute difference between the ratio and 1 being greater than 0.2) were bolded.

Table 4. Summary of data availability for variables relevant to asthma management and data validity by SES for each cohort (training and testing cohort)

	Training Q1 (n = 22)	Training Q2–Q4 (n = 102)	Testing Q1 (n = 15)	Testing Q2–Q4 (92)
Data unavailability, *n* (%)				
Missing health maintenance visit	3 (14%)	12 (12%)	2 (13%)	11 (12%)
Missing asthma care compliance	14 (64%)	43 (42%)	9 (60%)	44 (48%)
Missing asthma severity	9 (41%)	24 (24%)	8 (53%)	23 (25%)
Missing asthma type	22 (100%)	95 (93%)	14 (93%)	79 (86%)
NAEPP recommendation missing	13 (59%)	43 (42%)	8 (53%)	37 (40%)
Missing smoking status	16 (73%)	39 (38%)	8 (53%)	34 (37%)
Missing data on missing school	15 (68%)	41 (40%)	8 (53%)	34 (37%)

	Training Q1 (n = 34)	Training Q2–Q4 (n = 112)	Testing Q1 (n = 37)	Testing Q2–Q4 (n = 121)
Data validity*				
Undiagnosed (ICD) asthma	4 (12%)	11 (9.8%)	3 (8.1%)	8 (6.6%)

*Data validity was calculated for subjects who met PAC criteria but did not have physician diagnosis of asthma.

(age, sex, and race/ethnicity). This was also true for sensitivity. A recent study also showed ML models having differential performance by SES (measured by health insurance, public vs commercial health insurance) in predicting ICU mortality[12] and 30-day psychiatric readmission (people with lower SES had poorer prediction performance of their ML algorithms, compared to those with higher SES).[12] Overall, our study results and the literature suggest that SES may be associated with differential (in)completeness and validity of PROs, which may

subsequently lead to differential algorithmic performance by SES. However, this needs to be further assessed in other health outcomes and for different populations (eg, adults).

It is also important to recognize differential performance of SES measures in predicting health outcomes because researchers routinely use aggregate-level SES measures such as ADI[24,63–65] or other SES measures in research. Aggregate-level SES measures are subject to a significant misclassification of individual-level SES (20–35%)[66,67] and the ecological fallacy[68] and thus, may fail to detect the association of SES with health outcomes. As shown in results, compared to ADI, HOUSES classified more people as low SES, which led to a larger low SES subgroup, which in turn made it possible to compute more bias measures using the HOUSES. For example, there was significant discrepancy in the proportion of subjects with a history of AE among lower SES group defined by HOUSES (53%) and ADI (0%) which contrasts with the widely recognized associations of lower SES with the increased risk of AE in the literature.[54–56,69] In the analysis for algorithmic bias, ADI as an aggregate-level SES measure showed significant limitations and difficulties in applying it to research work assessing algorithmic bias, especially work based on a small sample size requiring precision, due to its imprecision and misclassification of individual-level SES measures. Along these lines, our recent study showed that HOUSES predicted that kidney transplant recipients with lowest HOUSES (Q1) had a significantly higher risk of graft failure than those with highest HOUSES (Q2–4) (adjusted hazard ratio 2.12; 95% CI 1.08–4.16).[70] Importantly, other SES measures such as individual educational levels and census-block group level education and income *failed to predict* outcomes on graft failure. Therefore, in assessing and mitigating algorithmic bias by SES, it is important to a valid measure for individual-level SES measure. The HOUSES index fulfills this requirement and can be a replacement or complement to existing conventional SES measures. As AI models are ultimately being applied to clinical decisions for individual patients, assessing AI model performance and bias using individual-level SES is conceptually and ethically more appropriate than aggregate-level SES measures when individual-level SES measures are available to developers.

The HOUSES index has several conceptual and methodological merits for clinical and translational research, as summarized in the Supplementary section: First, HOUSES is able to capture health effects of SES (defined as 'one's ability to access desired resources')[71] which is associated with 39 health care access, care quality, and health outcomes as summarized in Supplementary Table S2. In this context, HOUSES might be particularly attuned to asthma due to links between housing quality (eg, indoor or outdoor air quality or molds from moisture areas with poor ventilation) and childhood asthma as discussed in the Introduction. Second, it is an external and individual-level SES measure, in contrast to self-reported (eg, income) or aggregate-level (eg, zip-code-based Census data) measures. Third, it can retrospectively measure SES at any given point in time whenever address information at the index date of events is available (not relying on recalls). Fourth, as spatial coordinates are intrinsic to HOUSES, it enables geospatial analysis to identify geographic hotspots of interest (eg, COVID-19 cases) to be used as a feature in predictive models.[72–74] Finally, unlike other SES measures (eg, educational level, which is relatively static), it can capture longitudinal changes as real property data are regularly updated, and relocation of residence often reflects changes in a subject's SES. This feature allows us to use the HOUSES index as a financial outcome across life stages. Taken together, these features highlight how the HOUSES index can help to address issues of fairness in AI adoption,

ultimately helping to achieve greater levels of health equity across populations.

Our study has a few strengths. First, our study is based on a real-world setting where patients have a wide range of EHR completeness, instead of studies based on highly selected subjects. Second, we used an objective individual-level SES measure instead of self-reported or aggregate-level SES measures (eg, Census level data). Therefore, it does not suffer from recall bias or inaccuracy due to aggregation. Third, we assessed data availability and validity for features relevant to AE risk, which is not commonly done in AI research despite its importance. Our study also has limitations. First, the analysis was based on a small sample size. The present study was an exploratory case study based on a small sample size, and thus, findings are preliminary and require confirmation and further assessment from future studies with a larger sample size. In future work, we may also use variability as a way to estimate uncertainty (ie, estimating CIs of point estimates), which would capture uncertainty resulting from small sample size. More importantly, we were not able to do a separate analysis by different minorities due to the lack of samples within minority groups. However, future work can build on this approach of using the HOUSES index as an individual-level SES measure to assess potential bias from adoption of AI systems. Second, our study subjects may not represent the general pediatric population. However, it represents patient population (source population) as this study was based on those who receive care at our institution without involving any recruitment steps. Recognizing the cumulative residential effect from environment,[75] our current work did not include measurement for cumulative residential effect (eg, capturing longitudinal changes of traffic volume associated with changes of address over time) in the analysis. Third, a potentially informative data when using HOUSES as an SES measure is the number of residents in a house. While we recognize its importance, the data source that we use for formulating HOUSES (real property data from counties) does not include this information, and thus, we are unable to investigate its importance. Lastly, while HOUSES was validated in other states such as Missouri and South Dakota,[33,76] HOUSES requires further testing in other areas, including urban cities such as New York or Chicago, to establish validity before applying it across the United States and beyond.

CONCLUSION

Our study findings highlight the important role of SES in assessing potential bias that can result from differential performance of AI models across SES. Understanding the extent to which SES is a dimension along which bias occurs and examining the potential reasons or mechanisms that generate this bias will be crucially important for recognizing and mitigating bias in emerging applications of AI in health care. It will ultimately support efforts to promote health equity and fairness. We believe the HOUSES index, and the approach outlined here, can play an important role in those efforts.

FUNDING

AUTHOR CONTRIBUTIONS

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST STATEMENT

DATA AVAILABILITY

The datasets generated and/or analyzed during the current study are not publicly available as they include protected health information. Access to data could be discussed per the institutional policy after approval of the IRB at Mayo Clinic.

REFERENCES

1. (ONC) TOotNCfHIT. Health IT Dashboard. 2021. https://dashboard.healthit.gov/apps/health-information-technology-data-summaries.php?state=National&cat9=all+data. Accessed January 17, 2021.
2. Partners SG. *The State of Healthcare Automation: Urgent Need, Growing Awareness and Tremendous Potential*. Baltimore, MD: Sage Growth Partners; 2021.
3. Data Science Institute American College of Radiology. FDA Cleared AI Algorithms. https://www.acrdsi.org/DSI-Services/FDA-cleared-ai-algorithms. Accessed January 1, 2020.
4. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med* 2020; 3 (1): 118.
5. Yao X, Rushlow DR, Inselman JW, *et al*. Artificial intelligence–enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. *Nat Med* 2021; 27 (5): 815–9.
6. Pierson E, Cutler DM, Leskovec J, Mullainathan S, Obermeyer Z. An algorithmic approach to reducing unexplained pain disparities in underserved populations. *Nat Med* 2021; 27 (1): 136–40.
7. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019; 366 (6464): 447–53.
8. Institute of Medicine. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington DC: The National Academy of Science; 2003.
9. National Academies of Sciences, Engineering, and Medicine. Accounting for Social Risk Factors in Medicare Payment. Washington, DC: The National Academies Press; 2017.
10. Dzau VJ, McClellan MB, McGinnis J, *et al*. Vital directions for health and health care: Priorities from a national academy of medicine initiative. *JAMA* 2017; 317 (14): 1461.
11. Park Y, Hu J, Singh M, *et al*. Comparison of methods to reduce bias from clinical prediction models of postpartum depression. *JAMA Netw Open* 2021; 4 (4): e213909.
12. Irene Y, Chen PS, Marzyeh G. Can AI help reduce disparities in general medical and mental health care? *AMA J Ethics* 2019; 21 (2): E167–79.
13. Hoskins KF, Danciu OC, Ko NY, Calip GS. Association of race/ethnicity and the 21-gene recurrence score with breast cancer–specific mortality among US women. *JAMA Oncol* 2021; 7: 370–8.
14. Ferryman K. Addressing health disparities in the Food and Drug Administration's artificial intelligence and machine learning regulatory framework. *J Am Med Inform Assoc* 2020; 27 (12): 2016–9.
15. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. *Ann Intern Med* 2018; 169 (12): 866–72.
16. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Inform Assoc* 2018; 25 (8): 1080–8.
17. Lantz PM, Lynch JW, House JS, *et al*. Socioeconomic disparities in health change in a longitudinal study of US adults: the role of health-risk behaviors. *Soc Sci Med* 2001; 53 (1): 29–40.
18. Lantz PM. The medicalization of population health: who will stay upstream? *Milbank Q* 2019; 97 (1): 36–9.
19. Bach PB, Pham HH, Schrag D, Tate RC, Hargraves JL. Primary care physicians who treat Blacks and Whites. *N Engl J Med* 2004; 351 (6): 575–84.
20. Warnecke RB, Oh A, Breen N, *et al*. Approaching health disparities from a population perspective: the National Institutes of Health Centers for Population Health and Health Disparities. *Am J Public Health* 2008; 98 (9): 1608–15.
21. Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies. *Health Aff (Millwood)* 2002; 21 (2): 60–76.
22. Bernheim SM, Ross JS, Krumholz HM, Bradley EH. Influence of patients' socioeconomic status on clinical management decisions: a qualitative study. *Ann Fam Med* 2008; 6 (1): 53–9.
23. Franks P, Fiscella K. Effect of patient socioeconomic status on physician profiles for prevention, disease management, and diagnostic testing costs. *Med Care* 2002; 40 (8): 717–24.
24. Sills MR, Hall M, Colvin JD, *et al*. Association of social determinants with children's hospitals' preventable readmissions performance. *JAMA Pediatr* 2016; 170 (4): 350–8.
25. Roberts ET, Zaslavsky AM, Barnett ML, Landon BE, Ding L, McWilliams J. Assessment of the effect of adjustment for patient characteristics on hospital readmission rates: implications for pay for performance. *JAMA Intern Med* 2018; 178 (11): 1498–507.
26. Baker DW, Chassin MR. Holding providers accountable for health care outcomes. *Ann Intern Med* 2017; 167 (6): 418–23.
27. Jha AK, Zaslavsky AM. Quality reporting that addresses disparities in health care. *JAMA* 2014; 312 (3): 225–6.
28. Snyder-Mackler N, Burger JR, Gaydosh L, *et al*. Social determinants of health and survival in humans and other animals. *Science* 2020; 368 (6493): eaax9553.
29. Belsky DW, Snyder-Mackler N. Invited commentary: integrating genomics and social epidemiology—analysis of late-life low socioeconomic status and the conserved transcriptional response to adversity. *Am J Epidemiol* 2017; 186 (5): 510–3.
30. Martens DS, Janssen BG, Bijnens EM, *et al*. Association of parental socioeconomic status and newborn telomere length. *JAMA Netw Open* 2020; 3 (5): e204057.
31. Phelan JC, Link BG. Controlling disease and creating disparities: a fundamental cause perspective. *J Gerontol B Psychol Sci Soc Sci* 2005; 60 Spec No 2: 27–33.

32. Juhn YJ, Beebe TJ, Finnie DM, *et al.* Development and initial testing of a new socioeconomic status measure based on housing data. *J Urban Health* 2011; 88 (5): 933–44.

33. National Quality Forum Technical Report. Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors. 2014. The report is funded by the DHHS under contract HHSM-500-2012-000091 task order 7.

34. National Quality Forum. Evaluation of NQF's Trial Period for Risk Adjustment for Social Risk Factors. 2017.

35. U.S. Food & Drug Administration: Center for Devices & Radiological Health. Executive Summary for the Patient Engagement Advisory Committee Meeting: Artificial Intelligence (AI) and Machine Learning (ML) in Medical Devices. 2020. https://www.fda.gov/media/142998/download. Accessed December 2, 2020.

36. Seol HY, Shrestha P, Muth JF, *et al.* Artificial intelligence-assisted clinical decision support for childhood asthma management: a randomized clinical trial. *PLoS One* 2021; 16 (8): e0255261.

37. St Sauver JL, Grossardt BR, Yawn BP, *et al.* Data resource profile: the Rochester Epidemiology Project (REP) medical records-linkage system. *Int J Epidemiol* 2012; 41 (6): 1614–24.

38. St Sauver JL, Grossardt BR, Leibson CL, Yawn BP, Melton LJ, Rocca WA. Generalizability of epidemiological findings and public health decisions: an illustration from the Rochester Epidemiology Project. *Mayo Clin Proc* 2012; 87 (2): 151–60.

39. Zhong W, Finnie DM, Shah ND, *et al.* Effect of multiple chronic diseases on health care expenditures in childhood. *J Prim Care Community Health* 2015; 6 (1): 2–9.

40. Yawn BP, Wollan P, Kurland M, Scanlon P. A longitudinal study of the prevalence of asthma in a community population of school-age children. *J Pediatr* 2002; 140 (5): 576–81.

41. Narayanan A. Translation tutorial: 21 fairness definitions and their politics. In: *Paper presented at The Conference on Fairness, Accountability, and Transparency (FAT*)*; 2018; New York, USA.

42. Chouldechova A. Fair prediction with disparate impact: a study of bias in recidivism prediction instruments. *Big Data* 2017; 5 (2): 153–63.

43. Felman MF, Moeller J, Scheidegger C, Venkatasubramanian S. Certifying and removing disparate impact. In: *Knowledge Discovery and Data Mining '15: Proceedings of the 21th Association for Computing Machinery, Special Interest Group Knowledge Discovery and Data Mining International Conference on Knowledge Discovery and Data Mining*; 2015: 259–68.

44. Bellamy RKE, Dey K, Hind M, *et al.* AI fairness 360: an extensible toolkit for detecting and mitigating algorithmic bias. *IBM J Res Dev* 2019; 63 (4/5): 4:1–15.

45. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible—the neighborhood atlas. *N Engl J Med* 2018; 378 (26): 2456–8.

46. Bjur KA, Wi CI, Ryu E, Crow SS, King KS, Juhn YJ. Epidemiology of children with multiple complex chronic conditions in a mixed urban-rural US community. *Hosp Pediatr* 2019; 9 (4): 281–90.

47. Feudtner C, Feinstein JA, Zhong W, Hall M, Dai D. Pediatric complex chronic conditions classification system version 2: updated for ICD-10 and complex medical technology dependence and transplantation. *BMC Pediatr* 2014; 14: 199.

48. Wu ST, Sohn S, Ravikumar KE, *et al.* Automated chart review for asthma cohort identification using natural language processing: an exploratory study. *Ann Allergy Asthma Immunol* 2013; 111 (5): 364–9.

49. Wi CI, Sohn S, Rolfes MC, *et al.* Application of a natural language processing algorithm to asthma ascertainment. An automated chart review. *Am J Respir Crit Care Med* 2017; 196 (4): 430–7.

50. Wi CI, Sohn S, Ali M, *et al.* Natural language processing for asthma ascertainment in different practice settings. *J Allergy Clin Immunol Pract* 2018; 6 (1): 126–31.

51. Bisgaard H, Szefler S. Prevalence of asthma-like symptoms in young children. *Pediatr Pulmonol* 2007; 42 (8): 723–8.

52. Bloom CI, Franklin C, Bush A, Saglani S, Quint JK. Burden of preschool wheeze and progression to asthma in the UK: Population-based cohort 2007 to 2017. *J Allergy Clin Immunol* 2021; 147 (5): 1949–58.

53. Ryu E, Olson JE, Juhn YJ, *et al.* Association between an individual housing-based socioeconomic index and inconsistent self-reporting of health conditions: a prospective cohort study in the Mayo Clinic Biobank. *BMJ Open* 2018; 8 (5): e020054.

54. Harris MN, Lundien MC, Finnie DM, *et al.* Application of a novel socioeconomic measure using individual housing data in asthma research: an exploratory study. *NPJ Prim Care Respir Med* 2014; 24: 14018.

55. Akinbami LJ, Simon AE, Rossen LM. Changing trends in asthma prevalence among children. *Pediatrics* 2016; 137 (1): 1–7.

56. Cardet JC, Louisias M, King TS, *et al.* Income is an independent risk factor for worse asthma outcomes. *J Allergy Clin Immunol* 2018; 141 (2): 754–60.e3.

57. Flores G, Snowden-Bridon C, Torres S, *et al.* Urban minority children with asthma: substantial morbidity, compromised quality and access to specialists, and the importance of poverty and specialty care. *J Asthma* 2009; 46 (4): 392–8.

58. Cooper S, Rahme E, Tse SM, Grad R, Dorais M, Li P. Are primary care and continuity of care associated with asthma-related acute outcomes amongst children? A retrospective population-based study. *BMC Prim Care* 2022; 23 (1): 5.

59. Johnson LH, Chambers P, Dexheimer JW. Asthma-related emergency department use: current perspectives. *Open Access Emerg Med* 2016; 8: 47–55.

60. Shippee ND, Shah ND, May CR, Mair FS, Montori VM. Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice. *J Clin Epidemiol* 2012; 65 (10): 1041–51.

61. Grembowski D, Schaefer J, Johnson KE, *et al.*; AHRQ MCC Research Network. A conceptual model of the role of complexity in the care of patients with multiple chronic conditions. *Med Care* 2014; 52 (Suppl 3): S7–14.

62. Boehmer KR, Shippee ND, Beebe TJ, Montori VM. Pursuing minimally disruptive medicine: disruption from illness and health care-related demands is correlated with patient capacity. *J Clin Epidemiol* 2016; 74: 227–36.

63. Ash AS, Mick EO, Ellis RP, Kiefe CI, Allison JJ, Clark MA. Social determinants of health in managed care payment formulas. *JAMA Intern Med* 2017; 177 (10): 1424–30.

64. Knighton AJ, Savitz L, Belnap T, Stephenson B, VanDerslice J. Introduction of an area deprivation index measuring patient socioeconomic status in an integrated health system: implications for population health. *eGEMs* 2016; 4 (3): 9.

65. Chien AT, Wroblewski K, Damberg C, *et al.* Do physician organizations located in lower socioeconomic status areas score lower on pay-for-performance measures? *J Gen Intern Med* 2012; 27 (5): 548–54.

66. Narla NP, Pardo-Crespo MR, Beebe TJ, *et al.* Concordance between individual vs. area-level socioeconomic measures in an urban setting. *J Health Care Poor Underserved* 2015; 26: 1157–72.

67. Pardo-Crespo MR, Narla NP, Williams AR, *et al.* Comparison of individual-level versus area-level socioeconomic measures in assessing health outcomes of children in Olmsted County, Minnesota. *J Epidemiol Community Health* 2013; 67 (4): 305–10.

68. Geronimus AT. Invited commentary: using area-based socioeconomic measures–think conceptually, act cautiously. *Am J Epidemiol* 2006; 164 (9): 835–40. discussion 841–33.

69. Harris MN, Lundien MC, Finnie DM, *et al.* Application of a novel socioeconomic measure using individual housing data in asthma research: an exploratory study. *NPJ Prim Care Respir Med* 2014; 24: 14018.

70. Stevens MA, Beebe TJ, Wi C-I, Taler SJ, St., Sauver JL, Juhn YJ. HOUSES index as an innovative socioeconomic measure predicts graft failure among kidney transplant recipients. *Transplantation* 2020; 104 (11): 2383–92.

71. Oakes JM, Rossi PH. The measurement of SES in health research: current practice and steps toward a new approach. *Soc Sci Med* 2003; 56 (4): 769–84.

72. Juhn YJ, Wheeler P, Wi CI. Role of geographic risk factors in COVID-19 epidemiology: longitudinal geospatial analysis. *Mayo Clin Proc Innov Qual Outcomes* 2021; 5: 916–27.

73. Wi C-I, Wheeler PH, Kaur H, Ryu E, Kim D, Juhn Y. Spatio-temporal comparison of pertussis outbreaks in Olmsted County, Minnesota, 2004–2005 and 2012: a population-based study. *BMJ Open* 2019; 9 (5): e025521.

74. Patel AA, Wheeler PH, Wi C-I, *et al*. Mobile home residence as a risk factor for adverse events among children in a mixed rural–urban community: a case for geospatial analysis. *J Clin Trans Sci* 2020; 4 (5): 443–50.

75. Clarke P, Morenoff J, Debbink M, Golberstein E, Elliott MR, Lantz PM. Cumulative exposure to neighborhood context: consequences for health transitions over the adult life course. *Res Aging* 2014; 36 (1): 115–42.

76. Harris MN, Lundien MC, Finnie DM, *et al*. Application of a novel socioeconomic measure using individual housing data in asthma research: an exploratory study. *NPJ Prim Care Respir Med* 2014; 24: 14018.