Assessing the Impact of Social Needs and Social Determinants of Health on Health Care Utilization: Using Patient- and Community-Level Data

Elham Hatef, MD, MPH,[1,2] Xiaomeng Ma, MS,[1] Masoud Rouhizadeh, MS, PhD,[3]
Gurmehar Singh, MS,[1] Jonathan P. Weiner, DrPH,[1] and Hadi Kharrazi, MD, PhD[1,4]

Abstract

As the US health care system moves to expand access to and quality of medical care, the importance of addressing patient-level social needs and community-level social determinants of health (SDOH) is increasingly being recognized. This study evaluates individual- and community-level needs of housing (one of the SDOH domains) across the patient population of an academic medical center and explores how the level of housing needs impacts health care utilization. The authors performed a descriptive analysis of housing issues identified in both structured and unstructured (eg, clinical notes) data extracted from the electronic health record (EHR) and compared this to community-level characteristics of patients' neighborhood as measured by the Area Deprivation Index. Multivariate analyses were performed to assess the association between these and other factors on the frequency of service encounters. Among the 1,034,683 study participants, 59,703 (5.8%) had at least 1 housing issue identified in their EHR from structured or unstructured data combined. After adjusting for other factors, patients with housing instability and homelessness had 49% and 34% more encounters with the health care system compared to patients without housing issues ($P < 0.00001$). Patients living in the most disadvantaged neighborhoods had 55% more encounters with the health care system compared to those living in the most advantaged neighborhoods ($P < 0.00001$). This data collection approach and findings can inform health care systems aiming to make use of their EHRs and community-level SDOH information to provide a full assessment of patients' social needs and challenges.

Keywords: electronic health records, health care utilization, patient-level social data, community-level social data, social determinants of health, social needs

Introduction

As the US health care system moves to expand access to and quality of medical care, health care providers are increasingly recognizing the importance of addressing patient-level social needs and community-level social determinants of health (SDOH).[1–5] Increasing health care costs and worsening life expectancy in the United States are in part the direct results of these unmet needs.[1] Although these needs affect patients' health and life expectancy long before medical providers and health plans get involved, American health care organizations are increasingly acknowledging that they have ethical and fiscal responsibilities to address these social issues.[1]

The Department of Health and Human Services has highlighted the distinction between individual-level "social needs" and community-level "social determinants" (ie, SDOH).[6] At the individual level, social needs play a significant role in a person's health care use and overall health.[7–13] At the community level, a neighborhood's

[1]Department of Health Policy and Management, Center for Population Health Information Technology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.
[2]Department of Health Policy and Management at Bloomberg School of Public Health, Johns Hopkins Center for Health Disparities Solutions, Baltimore, Maryland, USA.
[3]Center for Clinical Data Analysis, Institute for Clinical and Translational Research, Johns Hopkins School of Medicine, Baltimore, Maryland, USA.
[4]Division of Health Sciences Informatics, Johns Hopkins School of Medicine, Baltimore, Maryland, USA.

socioeconomic composition and context are associated with self-rated health and risk of diseases[14,15] as well as health care utilization.[16,17]

Currently, clinical providers' efforts, if any, are focused mostly on meeting individual patients' social needs. Such efforts can be beneficial,[2,3] but 1-off social need interventions are generally limited to a small segment of the providers' patient population and often target the "sickest and most expensive" patients.[1] Health care delivery systems often do not comprehensively address the community-level social conditions of the neighborhoods they serve. Rather, providers usually rely on local, state, or national government policies to address population-level SDOH challenges.[4,5] Evidence from other countries has proven that a higher ratio of spending on social services to spending on health care has the best population health impact.[7]

Linking and analyzing available information on both individual-level social needs and community-level SDOH will enable health care systems, working in close consort with public health and social services agencies, to understand and address community-level policy actions in tandem with services targeted directly at both the social and medical needs of individual patients. To achieve these goals, it is critical to identify patient-level social needs in association with community-level SDOH and to assess their combined impact on health care utilization and health outcomes.

This study attempted to accomplish this for a large population of patients from across Johns Hopkins Health System (JHHS), an intersectional academic health care system with a large affiliated ambulatory provider network across Maryland. The study focused on housing issues as a proxy for individual-level social needs. The study team chose a composite measure of SDOH risk factors at the community level to assess the characteristics of the neighborhoods where the patients reside. The team measured the degree to which social needs and neighborhood SDOH impact some measure of health care utilization within JHHS. The team also applied geographic information system-based map visualization to identify neighborhoods with high social needs in support of community-level interventions.

Methods

Data sources

To assess patient-level housing needs and representative measures of health care utilization, the study team extracted both structured and unstructured data (providers' free-text notes) from the electronic health record (EHR) of all providers practicing at JHHS. The EHR data contained information collected from 1,187,956 unique patients captured between July 2016 and May 2018. For community-level SDOH the team used data from the US Census American Community Survey (ACS) 5-Year Estimates from 2017.[18]

Variable selection

The study team assessed patients' housing issues as an important social need and defined 3 related domains: homelessness, housing instability, and structural challenges associated with patients' residence (eg, characteristics of a building such as existing lead paint or structural damage).

In the structured EHR, the study team defined the 3 housing variables as follows: the team counted patients with a homeless shelter as their personal address, those with a positive response to the homelessness question in the available EHR screening questionnaires, and patients with the *International Classification of Diseases, Tenth Revision* (ICD-10) code of Z590 as homeless. Patients with ICD-10 codes of Z591 and/or Z598 were counted as those with housing instability. The team was unable to identify any ICD-10 codes related to the structural challenges associated with patients' residence in the JHHS-EHR.

In order to identify housing issues in the unstructured EHR data, the team applied statistical natural language processing (NLP) and other text mining techniques (eg, Regular Expression pattern matching) to determine different domains of housing problems identifiable within the EHR unstructured data. The technical details of unstructured EHR information extraction and assessment of the accuracy of information retrieval is fully described in a separate paper.[19] In brief, to identify notes containing selected domains of housing issues, the team used handcrafted linguistic patterns that a team of experts developed using ICD-10 and other coding systems[20,21] and the description of those domains in common public health surveys and instruments.[18,22] The team also reviewed phrases derived from a literature review and the results of a manual annotation process from a previous study.[23] A patient was defined as having housing issues if there was at least 1 provider's note in her/his EHR containing any of the phrases developed.

The community-level SDOH measure was the Area Deprivation Index (ADI) constructed for neighborhoods in Maryland and Baltimore City. ADI is a composite measure, allowing for ranking of neighborhoods by their socioeconomic disadvantage.[24] The ADI construction in this study was based on ACS 5-Year Estimates in 2017[18] and the method introduced by Singh et al.[24] The study team selected ADI percentiles as the national ranks for different neighborhoods in Maryland, comparing them to other neighborhoods across the country. A higher ADI percentile represented more disadvantaged neighborhoods. The team defined 3 levels of neighborhood disadvantage as ≤10th percentile, 11th to 89th percentile, ≥90th percentile, following the order from low to high for further analyses.

To assess utilization of health care services, the study team identified the number of face-to-face encounters within JHHS for each patient during the study period (2016–2018). Each patient encounter was categorized as follows: 0–5, 6–10, 11–20, ≥21. Charlson comorbidity score[25] was applied to group patients' multimorbidity into one of 4 categories: 0, 1, 2, ≥3. The team also extracted EHR data on patients' age, sex, race, and insurance type.

Statistical analysis and spatial-based analysis

The team performed a descriptive analysis exploring the association between housing issues in structured and unstructured EHR data in relation to other patient-level and community-level characteristics. Logistic regression was used to assess the impact of community-level characteristics on utilization of health care services in the entire population and for Medicaid patients while adjusting for patient-level characteristics. The team defined the utilization of health care services as

the number of encounters with the health care system, comparing those at the 95th percentile (number of encounters ≥31) with the rest of the sample. The Akaike Information Criterion is reported for each model to show the model fitness.

Community-level SDOH issues were spatially explored by displaying the distribution of each domain of housing issues in Baltimore City in association with characteristics of the neighborhoods arranged by ADI strata. The number of patients residing in each neighborhood area represented the denominator population in the geoanalyses. The team identified the number of patients with housing issues in each of the 3 domains by block groups and presented adjusted counts on maps. All analyses were conducted and map visualizations generated using R version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria).[26] The Institutional Review Board of Johns Hopkins Bloomberg School of Public Health approved this study.

Results

Of 1,187,956 unique patients in the cohort 1,034,683 (87.1%) had complete data on their demographic characteristics, clinical comorbidities, health care utilization, and residence location. This population was used to link JHHS-EHR data to community-level ACS data[18] for ADI calculation. Of 1,034,683 patients, 59,703 (5.8%) had at least 1 housing issue identified in their EHR from structured or unstructured data combined. Table 1 presents the characteristics of the study population and those with housing issues.

Overall, more patients had housing issues in the unstructured data of the EHR than in the structured data (Figure 1). The team also assessed the breakdown of the different domains of housing issues, including homelessness, housing instability, and characteristics of the building of patients' residence, in unstructured EHR data (Supplementary Table S1).

The team compared characteristics of patients with documented housing issues in their structured and unstructured EHR. Compared to unstructured EHR data, patients identified in structured EHR data were younger (eg, 17.5% aged 65–85 years in structured vs. 22.8% in unstructured EHR group), and more likely to be African American (35.4% vs 27.8%), to reside in the most disadvantaged neighborhoods with ADI at the 90th percentile (9.2% vs 5.0%), more likely to be on Medicaid (18.1% vs 7.7%), and to have a high health care utilization rate (ie, more than 20 encounters during the study period; 30.2% vs 10.0%), but to have fewer comorbidities (Charlson Comorbidity Score ≥3: 4.9% vs 8.8%). All of these differences were statistically significant ($P < 0.00001$).

The team assessed the number of encounters with the JHHS network during the study period (2016–2018) among all JHHS patients and those with housing issues. Overall, patients with housing issues had more encounters with JHHS (Table 1) and the differences were statistically significant ($P < 0.00001$). After adjusting for other factors among all study patients, those with homelessness and housing instability had 34% and 49% more encounters with the health care system compared to patients without housing issues ($P < 0.00001$). Medicaid patients with homelessness or housing instability had 92% and 47% more encounters with JHHS compared to patients without housing issues ($P < 0.00001$). Controlling for housing issues and other factors among all study patients, those living in neighborhoods at the 90th percentile ADI had 55% more encounters

with the health care system compared to those living in neighborhoods at the 10th percentile ADI ($P < 0.00001$). Medicaid patients living in neighborhoods at the 90th percentile ADI had 60% more encounters with the health care system compared to those living in neighborhoods at the 10th percentile ADI ($P < 0.00001$) (Table 2).

TABLE 1. CHARACTERISTICS OF THE OVERALL STUDY POPULATION AND THOSE WITH HOUSING ISSUES AT JOHNS HOPKINS HEALTH SYSTEM BETWEEN 2016 AND 2018

Characteristics	Overall (N=1,034,683) n	%	Housing issues* (N=59,674) n	%
Age (years)				
0–17	196,252	19.0%	11,388	19.1%
18–29	128,229	12.4%	8070	13.5%
30–49	254,088	24.6%	14,866	24.9%
50–64	220,109	21.3%	12,771	21.4%
65–85	236,005	22.8%	12,579	21.1%
Sex				
Male	457,985	44.3%	26,468	44.4%
Female	576,607	55.7%	33,206	55.6%
Race				
White	592,795	57.3%	32,197	54.0%
African American	279,802	27.0%	18,177	30.5%
Asian	53,471	5.2%	2704	4.5%
Other	108,615	10.5%	6596	11.1%
Neighborhood Characteristics (ADI National Rank)[†]				
Below the 10th Percentile	268,001	25.9%	13,821	23.2%
Between 11th & 89th Percentiles	717,002	69.3%	42,046	70.5%
Above the 90th Percentile	49,680	4.8%	3807	6.4%
Insurance Type				
Medicare	206,178	19.9%	11,960	20.0%
Medicaid	76,956	7.4%	6513	10.9%
Commercial	522,962	50.5%	27,051	45.3%
No Insurance	221,179	21.4%	13,739	23.0%
International Patients	787	0.1%	46	0.1%
Other	6621	0.6%	365	0.6%
Health Care Utilization[‡]				
0–5	656,755	63.5%	32,366	54.2%
6–10	159,432	15.4%	9498	15.9%
11–20	119,969	11.6%	8407	14.1%
≥21	98,527	9.5%	9403	15.8%
Charlson Comorbidity Score				
0	806,173	77.9%	44,195	74.1%
1	177,552	17.2%	11,145	18.7%
2	34,983	3.4%	2772	4.6%
≥3	15,975	1.5%	1562	2.6%

*Patient-level housing issues (identified in EHR structured or unstructured data) included homelessness, housing instability, and building characteristics/structure of the patients' residence.

[†]Using ADI national rank for patient's place of residence. Higher ADI national rank represents a more disadvantaged neighborhood.

[‡]Number of encounters with the health care system during the study period (2016–2018).

ADI, Area Deprivation Index; EHR, electronic health record.
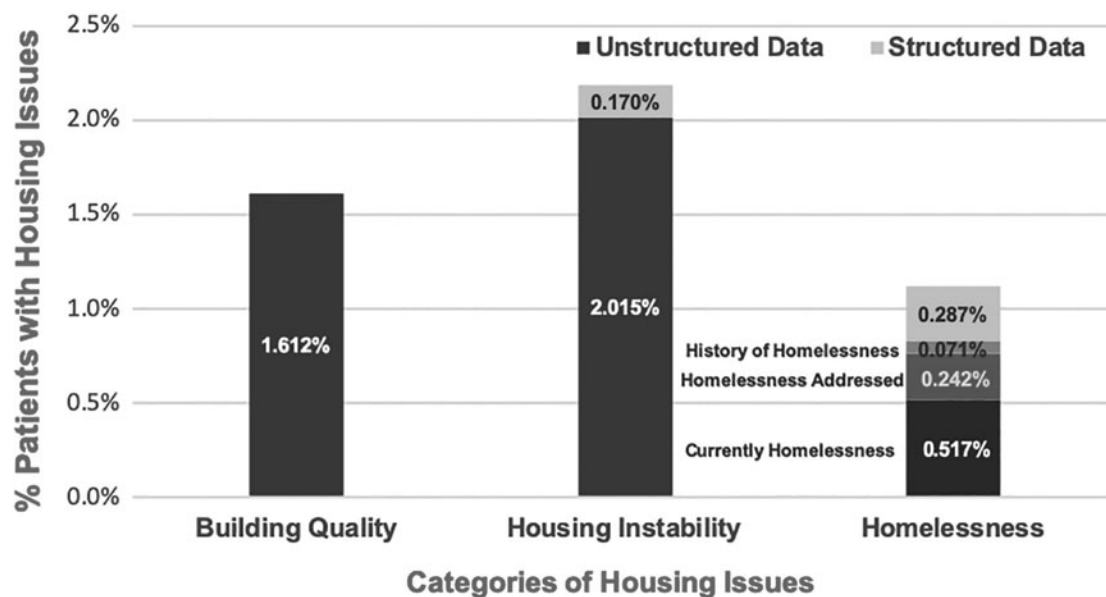
FIG. 1. Percentage of patients with housing issues in their EHR at Johns Hopkins Health System between 2016 and 2018; comparing structured* and unstructured data.
Notes:
* Patients with no entry in their address box, those with a positive response to the homelessness question in EHR questionnaires, and patients with an ICD-10 code of Z590 were counted as homeless in the structured data. Patients with ICD-10 codes of Z591 and/or Z598 were counted as patients with housing instability in the structured data.
Patients with at least 1 provider's note containing any of the developed phrases referring to housing issues in their unstructured EHR data were counted as those with housing issues.
EHR, electronic health record; ICD-10, International Classification of Diseases, Tenth Revision.

Figures 2a–c present maps of housing issues across different neighborhoods in Baltimore City using structured and unstructured EHR data for patient-level housing needs, and census block group ADI for community-level SDOH. From a to c, Figure 2 displays the density of homelessness, housing instability, and characteristics of the buildings. Overall, housing issues were more prevalent in disadvantaged neighborhoods (larger gray dots/circles in neighborhoods with higher ADI presented in dark gray polygons of block groups). East Baltimore, a disadvantaged neighborhood with high ADI and in close proximity to JHHs' main hospital had the highest number of housing issues across the 3 housing domains.

Discussion

Health care providers are increasingly recognizing the importance of addressing nonmedical challenges facing their patients and the communities where they reside. To achieve these goals, it is critical to identify social challenges of patients and their communities and to assess their combined impact on health care utilization and health outcomes. This study attempted to accomplish this by using JHHS-EHR data linked to ADI, a composite measure of SDOH risk factors at the community level, and assessed how the combination of these factors affected the number of encounters at JHHS.

Within the medical community, addressing social needs and SDOH challenges has recently started and is gaining significant momentum.[27–30] A number of health care sys-

tems have already begun designing and implementing interventions that target such needs and challenges.[31] For instance, Gold et al outlined the capability of EHRs to improve care coordination,[32] and to support referrals for the social needs identified.[33] In addition to using EHR data for care coordination, Cincinnati Children's Hospital Medical Center established community partnerships to address legal needs surrounding SDOH (eg, housing).[34]

To successfully accomplish the goals behind this momentum, health care systems must improve the identification of social needs of their patients using EHR-embedded questionnaires and screening tools and linking such data to characteristics of patients' neighborhoods. Integrating patient-level EHR data with community-level data will provide a more robust system for addressing social needs and SDOH at the patient and community levels.[28] Possible uses of such data include risk prediction and identification, care coordination, and contextualized medical care, as well as community-based interventions.

For instance, including social needs and SDOH variables in risk prediction models will help health care systems to predict greater utilization of health care services (eg, hospitalizations, emergency department visits) in subsets of their patient population experiencing such challenges. In addition, using data on social needs and SDOH challenges helps to prioritize patients with such risk factors for care coordination interventions. An example would be a care manager or physician identifying patients with homelessness and housing insecurity and referring them to a homeless shelter or available housing assistance programs.

TABLE 2. LOGISTIC REGRESSION ASSESSING FACTORS ASSOCIATED WITH HEALTH CARE UTILIZATION AMONG JOHNS HOPKINS PATIENTS BETWEEN 2016–2018*

	Overall population			Medicaid population		
Variables	OR[†]	95% CI	P	OR[†]	95% CI	P
Housing Issues						
Homelessness	1.336	1.261–1.416	<0.00001	1.902	1.576–2.296	<0.00001
Housing Instability	1.489	1.380–1.607	<0.00001	1.473	1.227–1.769	<0.00001
Characteristics of the Building	0.888	0.818–0.964	0.00469	0.847	0.640–1.121	0.24500
Age						
	1.001	1.001–1.002	<0.00001	1.010	1.008–1.011	<0.00001
Sex (Male as reference)						
Female	1.437	1.409–1.467	<0.00001	1.563	1.458–1.675	<0.00001
Race (White as reference)						
Asian	0.741	0.701–0.782	<0.00001	0.493	0.385–0.631	<0.00001
African American	0.959	0.937–0.981	0.00039	0.794	0.734–0.858	<0.00001
Others	0.646	0.620–0.674	<0.00001	0.553	0.488–0.627	<0.00001
Neighborhood Characteristics (ADI National Rank, neighborhoods below the 10th percentile as reference)[‡]						
Between 11th & 89th Percentiles	1.442	1.404–1.481	<0.00001	1.466	1.239–1.734	<0.00001
Above the 90th Percentile	1.549	1.474–1.627	<0.00001	1.598	1.325–1.926	<0.00001
Insurance Type (commercial insurance as reference)						
Medicare	1.489	1.447–1.532	<0.00001	-	-	-
Medicaid	2.078	1.997–2.162	<0.00001	-	-	-
No insurance	2.416	2.356–2.477	<0.00001	-	-	-
International	1.909	1.358–2.683	0.00020	-	-	-
Others	0.394	0.301–0.516	<0.00001	-	-	-
Charlson Comorbidity Score (score of 0 as reference)						
1	5.636	5.503–5.771	<0.00001	4.219	3.912–4.550	<0.00001
2	18.674	18.100–19.267	<0.00001	16.357	14.518–18.427	<0.00001
≥3	55.444	53.333–57.639	<0.00001	38.497	32.447–45.675	<0.00001
AIC[§]						
		312,925			27,005	

*Health care utilization was defined as the number of encounters with the health care system during the study period (2016–2018). Logistic regression compared those at the 95th percentile (number of encounters ≥31) with the rest of the sample. Total sample size for overall the study population was 1,034,592 and for Medicaid patients was 76,946.

[†]Odds ratios associated with being a top 5% user among study patients between 2016–2018.

[‡]Using ADI national rank for patient's place of residence. Higher ADI national rank represents a more disadvantaged neighborhood.

[§]An estimator of the relative quality of statistical models for a given set of data. Lower values of AIC present a model with a better fit.

ADI, Area Deprivation Index; AIC, Akaike information criterion; CI, confidence enterval; OR, odds ratio; -, not applicable.

Integrating patient-level EHR data with community-level data will provide the opportunity for contextualized medical care. For instance, a patient with uncontrolled asthma who lives in a poor-quality building with structural damage might need some intervention beyond medical management of her/his disease. Identifying such environmental challenges will help the health care system to plan for a house visit by a care manager or community health worker to identify environmental risk factors that affect the disease course and its medical management. Linkage of EHR data and community-level data will help to prioritize community-based interventions in neighborhoods with greater challenges. An example would be identification of communities with greater numbers of buildings with existing lead paint and then conducting selective lead poisoning screening among children in such neighborhoods. A challenge in this context would be to identify at-risk communities served by multiple health care systems and the limitations related to the separate EHRs of those health care systems. Using community-level data from publicly available sources such as ACS would help to identify such at-risk communities.

Despite ongoing efforts to address the nonmedical needs of patients, several challenges related to data on social needs should be addressed. For instance, many sources of data on social needs can be found in a typical EHR (eg, notes, diagnoses, assessments) and some EHR vendors have started adding specific data fields for collecting information on social needs.[32,33] However, no universally accepted or standardized format exists for documenting social need information[35] and the majority of social need variables are not discretely represented or captured in structured formats such as ICD codes in EHRs.[36] Also, extracting data from EHR unstructured text (where most social needs and social needs information now resides) requires time-consuming and subjective methods (ie, manual chart review), which is not a feasible approach when trying to screen large populations of patients.[36,37]

NLP and text mining have become increasingly appealing techniques to mine unstructured data for social needs. Specifically, NLP techniques have been used to find different domains of social needs such as social connection and isolation.[23] In spite of ongoing efforts to use NLP
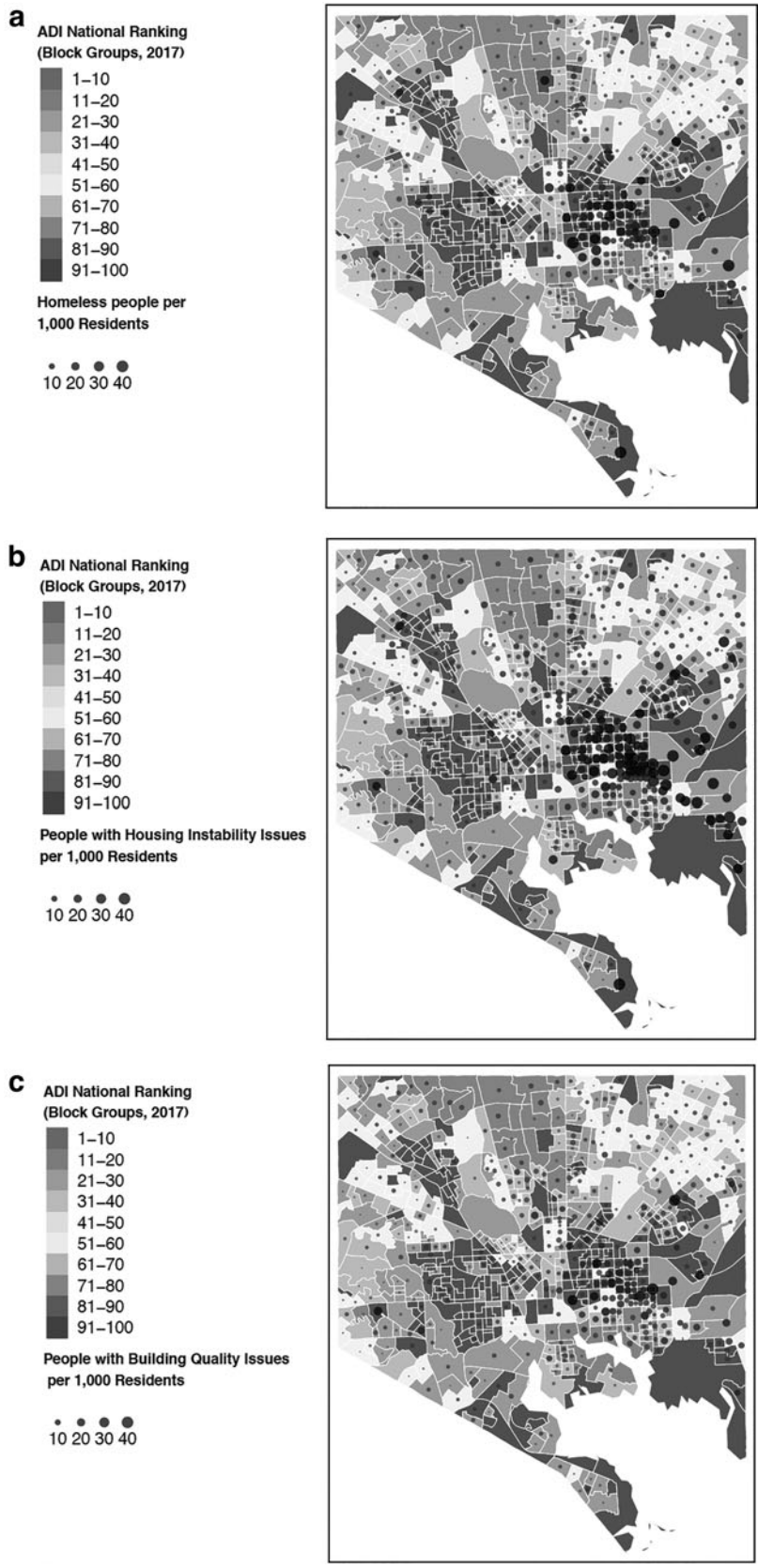
FIG. 2. Housing issues across different neighborhoods in Baltimore City using EHR structured and unstructured data on patient-level and community-level ADI (Block Groups): **(a)** homelessness, **(b)** housing instability, and **(c)** characteristics of the building. ADI, Area Deprivation Index; EHR, electronic health record.

techniques for data extraction on social needs from unstructured EHRs, off-the-shelf data extraction solutions lack social needs data, in contrast to clinical diagnostic codes and their standardized terminology.[36]

The study team's prior work evaluated the use of rule-based text mining methods and explored the utility of pattern-based techniques to extract selected domains of social needs from JHHS-EHR. The team investigated the coverage and accuracy of these methods among various clinical notes authored by different providers.[19] In this study, the same techniques were applied to extract information on housing issues from JHHS-EHR. Study findings were comparable with other studies and presented higher number of patients with social issues documented in their unstructured EHR data comparing to structured EHR data (eg, 0.3% of homelessness in structured vs 0.8% in unstructured EHR data).[37]

Social needs and SDOH challenges do not represent isolated problems that randomly distribute across a patient population or a community. Such needs and challenges are interconnected and affect the health of individuals and communities in aggregate. Hence, underserved populations often experience a number of social needs and SDOH challenges[38] that are believed to be critical factors in explaining many health-related disparities, ultimately leading to differences in mortality.[38] In the present study, patients with housing issues were more likely to be African American (30.5% vs 27.0%), on Medicaid (10.9% vs 7.4%), or to have no insurance (23.0% vs 21.4%), and lived in the most disadvantaged neighborhoods as measured by high ADI (6.4% vs 4.8%) compared to overall study population (Table 1).

Linking patient-level EHR data to community-level data using ADI helped the team to assess the impact of patients' individual needs in association with characteristics of their neighborhoods on health care utilization, at least within the large system where this study took place. The literature varies on how social needs and SDOH might affect outcomes of care (eg, hospitalizations, emergency department visits). Some studies have found that community-level SDOH may not contribute much more to risk prediction above and beyond what is already provided by EHR data,[39] while others have found these data to be predictive of care outcomes such as hospitalizations.[28] The present study showed that both patient-level social needs (ie, homelessness, housing instability) and community-level SDOH (ie, living in disadvantaged neighborhoods with high ADI) resulted in more encounters with the health care system among all study patients and those on Medicaid. The team performed modeling using the Medicaid population because they present similar socioeconomic challenges and poverty level beyond what was able to be measured in the study. Mapping housing issues and ADI also presented the same results, with more housing issues identified in disadvantaged neighborhoods (Figure 2a–c).

This study had some limitations. First, results were driven by the underlying EHR data of a large integrated academic health care system. Other health care organizations may find data on social needs captured and collected at different rates depending on the characteristics of their patient population, workflow, EHR use, and other system or policy factors. Second, ICD codes were used to identify information stored as structured data (as other coding terminologies were not found in JHHS-EHR); however, other coding terminologies (eg, LOINC, SNOMED) also have addressed various determinants of health. Investigation of information captured in EHRs using different coding systems might help identify more information stored as structured data. Third, this study focused on data captured before 2018; however, because of the trends in value-based payment models and policy requirements, an increase in collection of social needs information in EHR settings likely has already begun. And fourth, the number of encounters with the JHHS network during the study period was used to assess health care utilization. Utilization rates only included encounters documented in JHHS-EHR and not insurance claims or other ancillary databases. Utilization rates did not include encounters outside the JHHS network. In addition, this study did not differentiate among encounter types. The number of encounters at JHHS might not be a strong representation of health care utilization. Encounters may have varied and included a wide range of incidents with different levels of severity or impact on health, such as receiving a preventive service, or a primary care or emergency department visit, as well as more serious incidents such as an overnight hospitalization.

Conclusions

To the study team's knowledge, this study is one of the first attempts by a major academic health care system to assess the impact of patient-level social needs in association with community-level SDOH. The team assessed the rates of housing issues collected in the structured and unstructured EHR data of approximately 1.2 million patients and in association with their individual characteristics (ie, demographics and insurance type), clinical characteristics (ie, Charlson comorbidity score), and characteristics of their neighborhoods (ie, ADI). Data also were collected from a variety of health care settings, which helped avoid the possibility that physicians in one setting might have habitually failed to collect social data.

This data collection approach can inform health care systems aiming to use their EHRs to capture social needs and integrate them into patient care and population health management interventions. Findings of this study can serve as a baseline for future studies using patient-level data from EHRs linked to population-level data to assess the combination of social needs and SDOH impact on health outcomes and health care utilization.

Acknowledgment

Author Disclosure Statement

The authors declare that there are no conflicts of interest.

Supplementary Material

Supplementary Table S1

References

1. Castrucci B, Auerbach J. Meeting individual social needs falls short of addressing social determinants of health. 2019. https://www.healthaffairs.org/do/10.1377/hblog2019 0115.234942/full/ Accessed January 21, 2020.

2. Sulo S, Feldstein J, Partridge J, Schwander B, Sriram K, Summerfelt WT. Budget impact of a comprehensive nutrition-focused quality improvement program for malnourished hospitalized patients. Am Health Drug Benefits 2017;10:262–270.

3. Pruitt Z, Emechebe N, Quast T, Taylor P, Bryant K. Expenditure reductions associated with a social service referral program. Popul Health Manag 2018;21:469–476.

4. Centers for Disease Control and Prevention; Office of the Associate Director for Policy and Strategy. The HI-5 Interventions. 2016. https://www.cdc.gov/policy/hst/hi5/ interventions/index.html Accessed January 21, 2020.

5. de Beaumont Foundation and Kaiser Permanente. cityhealth website. https://www.cityhealth.org Accessed January 21, 2020.

6. Azar AM. The Root of the Problem: America's Social Determinants of Health Hatch Foundation for Civility and Solutions. 2018. https://www.hhs.gov/about/leadership/ secretary/speeches/2018-speeches/the-root-of-the-problem-americas-social-determinants-of-health.html Accessed January 21, 2020.

7. Institute of Medicine. Capturing social and behavioral domains and measures in electronic health records: phase 2. Washington, DC: National Academies Press, 2014.

8. Chetty R, Stepner M, Abraham S, et al. The association between income and life expectancy in the United States, 2001–2014. JAMA 2016;315:1750–1766.

9. Balakrishnan MP, Herndon JB, Zhang J, Payton T, Shuster J, Carden DL. The association of health literacy with preventable ED visits: a cross-sectional study. Acad Emerg Med 2017;24:1042–1050.

10. Mackelprang JL, Qiu Q, Rivara FP. Predictors of emergency department visits and inpatient admissions among homeless and unstably housed adolescents and young adults. Med Care 2015;53:1010–1017.

11. Wright BJ, Vartanian KB, Li H-F, Royal N, Matson JK. Formerly homeless people had lower overall health care expenditures after moving into supportive housing. Health Aff (Millwood) 2016;35:20–27.

12. Sadowski LS, Kee RA, VanderWeele TJ, Buchanan D. Effect of a housing and case management program on emergency department visits and hospitalizations among chronically ill homeless adults: a randomized trial. JAMA 2009;301:1771–1778.

13. Butler DC, Petterson S, Phillips RL, Bazemore AW. Measures of social deprivation that predict health care access and need within a rational area of primary care service delivery. Health Serv Res 2013;48(2 pt1):539–559.

14. Leonardi C, Simonsen NR, Yu Q, Park C, Scribner RA. Street connectivity and obesity risk: evidence from electronic health records. Am J Prev Med 2017;52(1 suppl 1): S40–S47.

15. Rohrer J, Pierce J, Denison A. Walkability and self-rated health in primary care patients. BMC Fam Pract 2004;5:29.

16. Bhargava V, Lee JS. Food insecurity and health care utilization among older adults in the United States. J Nutr Gerontol Geriatr 2016;35:177–192.

17. Weiser SD, Hatcher A, Frongillo EA, et al. Food insecurity is associated with greater acute care utilization among HIV-infected homeless and marginally housed individuals in San Francisco. J Gen Intern Med 2013;28:91–98.

18. The United States Census Bureau. American Community Survey (ACS). https://www.census.gov/programs-surveys/ acs/ Accessed January 21, 2020.

19. Hatef E, Rouhizadeh M, Tia I, et al. Assessing the availability of data on social and behavioral determinants in structured and unstructured electronic health records: a retrospective analysis of a multilevel health care system. JMIR Med Inform 2019;7:e13802.

20. Arons A, DeSilvey S, Fichtenberg C, Gottlieb L. Compendium of medical terminology codes for social risk factors. 2018. https://sirenetwork.ucsf.edu/tools-resources/ mmi/compendium-medical-terminology-codes-social-risk-factors Accessed January 21, 2020.

21. Richard M, Aimé X, Krebs MO, Charlet J. Enrich classifications in psychiatry with textual data: an ontology for psychiatry including social concepts. Stud Health Technol Inform 2015;210:221–223.

22. National Association of Community Health Centers. The Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences (PRAPARE). www.nachc.org/ research-and-data/prapare/ Accessed January 21, 2020.

23. Kharrazi H, Anzaldi LJ, Hernandez L, et al. The value of unstructured electronic health record data in geriatric syndrome case identification. J Am Geriatr Soc 2018;66:1499–1507.

24. Singh GK. Area deprivation and widening inequalities in US mortality, 1969–1998. Am J Public Health 2003;93: 1137–1143.

25. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–383.

26. The R Foundation for Statistical Computing. The R project for statistical computing. www.r-project.org Accessed January 21, 2020.

27. Hewner S, Casucci S, Sullivan S, et al. Integrating social determinants of health into primary care clinical and informational workflow during care transitions. EGEMS (Wash DC) 2019;5:2.

28. Hatef E, Searle KM, Predmore Z, et al. The impact of social determinants of health on hospitalization in the Veterans Health Administration. Am J Prev Med 2019;56:811–818.

29. Vest JR, Grannis SJ, Haut DP, Halverson PK, Menachemi N. Using structured and unstructured data to identify patients' need for services that address the social determinants of health. Int J Med Inform 2017;107:101–106.

30. Gottlieb LM, Tirozzi KJ, Manchanda R, Burns AR, Sandel MT. Moving electronic medical records upstream: incorporating social determinants of health. Am J Prev Med 2015;48:215–218.

31. Anugu M, Braksmajer A, Huang J, Yang J, Ladowski KL, Pati S. Enriched medical home intervention using community health worker home visitation and ED use. Pediatrics 2017;139:e20161849.

32. Gold R, Cottrell E, Bunce A, et al. Developing electronic health record (EHR) strategies related to health center patients' social determinants of health. J Am Board Fam Med 2017;30:428–447.

33. Gold R, Bunce A, Cowburn S, et al. Adoption of social determinants of health EHR tools by community health centers. Ann Fam Med 2018;16:399–407.

34. Henize AW, Beck AF, Klein MD, Adams M, Kahn RS. A road map to address the social determinants of health through community collaboration. Pediatrics 2015;136: e993–e1001.

35. National Academies of Sciences, Engineering, and Medicine. Integrating social care into the delivery of health care: moving upstream to improve the nation's health. Washington, DC: National Academies Press, 2019.

36. Winden TJ, Chen ES, Monsen KA, Wang Y, Melton GB. Evaluation of flowsheet documentation in the electronic health record for residence, living situation, and living conditions. AMIA Jt Summits Transl Sci Proc 2018;2018: 236–245.

37. Navathe AS, Zhong F, Lei VJ, et al. Hospital readmission and social risk factors identified from physician notes. Health Serv Res 2018;53:1110–1136.

38. Cunningham TJ, Croft JB, Liu Y, Lu H, Eke PI, Giles WH. Vital signs: racial disparities in age-specific mortality among blacks or African Americans—United States, 1999–2015. MMWR Morb Mortal Wkly Rep 2017;66:444–456.

39. Bhavsar NA, Gao A, Phelan M, Pagidipati NJ, Goldstein BA. Value of neighborhood socioeconomic status in predicting risk of outcomes in studies that use electronic health record data. JAMA Netw Open 2018;1:e182716–e182716.

Address correspondence to:
Elham Hatef, MD, MPH, FACPM
Department of Health Policy and Management
Center for Population Health Information Technology
Johns Hopkins Bloomberg School of Public Health
624 North Broadway, Room 502
Baltimore, MD 21205
USA

E-mail: ehatef1@jhu.edu