RESEARCH ARTICLE

Comparing the performance of screening surveys versus predictive models in identifying patients in need of health-related social need services in the emergency department

Olena Mazurenko[1,2]*, Adam T. Hirsh[3], Christopher A. Harle[1,2], Joanna Shen[2,4], Cassidy McNamee[1], Joshua R. Vest[1,2]

1 Department of Health Policy & Management, Indiana University Richard M. Fairbanks School of Public Health–Indianapolis, Indianapolis, Indiana, United States of America, 2 Regenstrief Institute, Indianapolis, Indiana, United States of America, 3 School of Science, Indiana University–Indianapolis, Indianapolis, Indiana, United States of America, 4 Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States of America

* omazuren@iu.edu

Abstract

Background

Health-related social needs (HRSNs), such as housing instability, food insecurity, and financial strain, are increasingly prevalent among patients. Healthcare organizations must first correctly identify patients with HRSNs to refer them to appropriate services or offer resources to address their HRSNs. Yet, current identification methods are suboptimal, inconsistently applied, and cost prohibitive. Machine learning (ML) predictive modeling applied to existing data sources may be a solution to systematically and effectively identify patients with HRSNs. The performance of ML predictive models using data from electronic health records (EHRs) and other sources has not been compared to other methods of identifying patients needing HRSN services.

Methods

A screening questionnaire that included housing instability, food insecurity, transportation barriers, legal issues, and financial strain was administered to adult ED patients at a large safety-net hospital in the mid-Western United States (n = 1,101). We identified those patients likely in need of HRSN-related services within the next 30 days using positive indications from referrals, encounters, scheduling data, orders, or clinical notes. We built an XGBoost classification algorithm using responses from the screening questionnaire to predict HRSN needs (screening questionnaire model). Additionally, we extracted features from the past 12 months of existing EHR, administrative, and health information exchange data for the survey respondents. We built ML predictive models with these EHR data using XGBoost (ML EHR model). Out of concerns of potential bias, we built both the screening

question model and the ML EHR model with and without demographic features. Models were assessed on the validation set using sensitivity, specificity, and Area Under the Curve (AUC) values. Models were compared using the Delong test.

Results

Almost half (41%) of the patients had a positive indicator for a likely HRSN service need within the next 30 days, as identified through referrals, encounters, scheduling data, orders, or clinical notes. The screening question model had suboptimal performance, with an AUC = 0.580 (95%CI = 0.546, 0.611). Including gender and age resulted in higher performance in the screening question model (AUC = 0.640; 95%CI = 0.609, 0.672). The ML EHR models had higher performance. Without including age and gender, the ML EHR model had an AUC = 0.765 (95%CI = 0.737, 0.792). Adding age and gender did not improve the model (AUC = 0.722; 95%CI = 0.744, 0.800). The screening questionnaire models indicated bias with the highest performance for White non-Hispanic patients. The performance of the ML EHR-based model also differed by race and ethnicity.

Conclusion

ML predictive models leveraging several robust EHR data sources outperformed models using screening questions only. Nevertheless, all models indicated biases. Additional work is needed to design predictive models for effectively identifying all patients with HRSNs.

Introduction

Healthcare organizations in the United States face growing expectations to address patients' health-related social needs (HRSNs) as issues such as housing instability, food insecurity, and financial strain affect health, well-being, and healthcare costs [1]. Patients with HRSNs require referrals to social service organizations, connections with social workers, or access to relevant resources to meet their specific HRSNs, like food or transportation vouchers [2]. Interventions to address HRSNs appear to be a promising solution to reducing unnecessary healthcare utilization and improving health [3].

Healthcare organizations must first correctly identify patients with HRSNs to refer them to appropriate services or offer resources to address their HRSNs [4,5]. This has proven challenging [6]. For one, identifying patients with HRSNs requires changes to clinical workflows by increasing clinician and patient data collection burdens and changes to information technology systems to record and ultimately leverage HRSN information [6]. Additionally, even successful interventions to identify patients with HRSNs will leave some patients unscreened due to non-responses, language barriers, breakdowns in workflow, time constraints, or uncertainty over screening responsibilities [7–10]. Moreover, the scope and scale of collecting HRSNs on patients are immense and costly [11]. While no optimal frequency for HRSN has been established [12], annual or even universal screening is a common recommendation [4] and the agency's new quality reporting metrics require HRSN screening during every admission [13]. Such reporting expectations would be a sizable data collection effort for many organizations. Notably, the actual performance of many existing screening instruments to collect HRSNs is unknown [14] or even potentially poor [15].

Machine learning (ML) predictive models may be an alternative that addresses many of the practical challenges of systematically identifying patients with HRSNs who need services and resources. The predictive modeling process can be automated, eliminating time constraints, workflow issues, or staff availability that often impede the current collection and use of HRSN data. In addition, automated predictive modeling operates as a universal screening program: it is not subject to intentional or unintentional biases that lead individuals to selectively administer questionnaires [16] or to missing data due to patient nonresponse. In addition, ML techniques can capitalize on the ever-growing amount of longitudinal electronic health records (EHR), health information exchange (HIE), and non-healthcare organization data reflecting patient social circumstances and factors [17,18]. These data can provide a longitudinal and comprehensive view of the patient and are not dependent upon a single organization for data collection. ML predictive modeling has already demonstrated promise in identifying patients with HRSNs [19–21]. Nevertheless, the performance of ML predictive models has not yet been compared to other methods of identifying patients needing HRSN services.

Objective

This study's objective was to compare the predictive power of EHR features versus screening questionnaire features in identifying HRSNs in identifying emergency department (ED) patients needing services to address HRSNs. The ED is an appropriate setting for such a study. First, HRSNs are highly prevalent among ED patients [10]. Second, patient HRSNs drive adverse outcomes such as repeat ED visits [22] and additional ED services [23]. Relatedly, patients' HRSNs complicate care delivery, inhibit treatment adherence, and act as a barrier to follow-up with primary and specialty care [22,24]. Third, the ED is an essential source of care for minoritized and under-resourced patient populations [25]. This is important because prior ML predictive models have been biased against these populations [26].

This analysis focuses on predicting a patient's likely need for HRSN services and not identifying the actual presence of individual HRSNs. The rationale for this choice is to focus on the most actionable outcome for providers. Specifically, at the individual patient level, screening for HRSNs is primarily used to inform clinicians about the need for appropriate services [4,5,27]–most often to initiate a referral to social service professionals (e.g., social workers, case managers, and navigators) or community-based social service organizations best equipped to address HRSNs [5,28]. This is true regardless of the method of identification used (ML predictive modeling or questionnaire). Moreover, the difference between the presence of a HRSN and the need for HRSN services is of practical importance for health care organizations. Not all patients responding to HRSNs screening want or receive services [6,29,30]. By focusing our prediction modeling on those with a likely need for HRSN services, we are more directly targeting the resources and activities that healthcare providers and organizations will need to supply.

Methods

Below, we describe the process of comparing two approaches to predict the likely need for an HRSN service (see the study pipeline in Fig 1). We built XGBoost classification algorithms using responses from a screening questionnaire (screening questionnaire model). Additionally, we built XGBoost classification algorithms using existing EHR, administrative, and health information exchange data (ML EHR model).
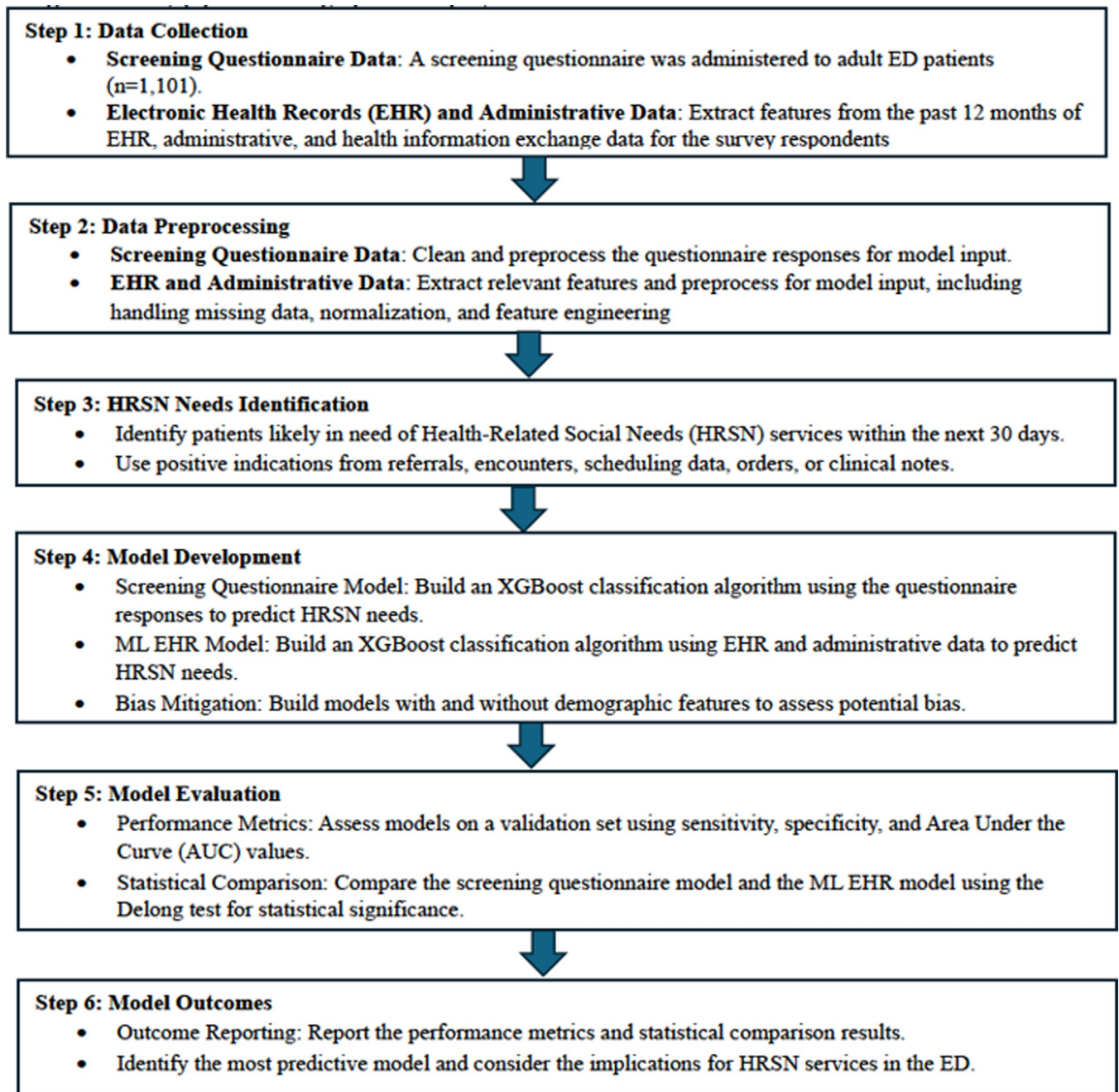
Step 1: Data Collection
- **Screening Questionnaire Data**: A screening questionnaire was administered to adult ED patients (n=1,101).
- **Electronic Health Records (EHR) and Administrative Data**: Extract features from the past 12 months of EHR, administrative, and health information exchange data for the survey respondents

Step 2: Data Preprocessing
- **Screening Questionnaire Data**: Clean and preprocess the questionnaire responses for model input.
- **EHR and Administrative Data**: Extract relevant features and preprocess for model input, including handling missing data, normalization, and feature engineering

Step 3: HRSN Needs Identification
- Identify patients likely in need of Health-Related Social Needs (HRSN) services within the next 30 days.
- Use positive indications from referrals, encounters, scheduling data, orders, or clinical notes.

Step 4: Model Development
- Screening Questionnaire Model: Build an XGBoost classification algorithm using the questionnaire responses to predict HRSN needs.
- ML EHR Model: Build an XGBoost classification algorithm using EHR and administrative data to predict HRSN needs.
- Bias Mitigation: Build models with and without demographic features to assess potential bias.

Step 5: Model Evaluation
- Performance Metrics: Assess models on a validation set using sensitivity, specificity, and Area Under the Curve (AUC) values.
- Statistical Comparison: Compare the screening questionnaire model and the ML EHR model using the Delong test for statistical significance.

Step 6: Model Outcomes
- Outcome Reporting: Report the performance metrics and statistical comparison results.
- Identify the most predictive model and consider the implications for HRSN services in the ED.

Fig 1. Study pipeline. The graphical display of the data elements and model outcomes.

https://doi.org/10.1371/journal.pone.0312193.g001

Setting and sample

The study sample included 1,101 adult (18 years old) patients who sought care at the Eskenazi Health ED in Indianapolis, IN, between June 2021 and March 2023. Eskenazi Health is the community's safety-net provider. Patients were ineligible for study inclusion if: 1) age <18 years old; 2) lack of decision capacity (e.g., psychosis, altered mental state, dementia) or 3) critical illness/injury impairing the patient's ability to consent/participate. It is the policy of Indiana University that all research involving human subjects shall be subject to review or granted

exemption by an appropriate Indiana University Institutional Review Board or designee before project initiation and without respect to funding or the source of funding. The Indiana University institutional review board approved all study procedures before the study started, and written informed consent was collected by all subjects.

Data sources

Screening questionnaire.   We collected primary data on patients' HRSNs using a screening questionnaire administered during an ED encounter between June 2021 and March 2023. The questionnaire contained the Epic EHR's social determinants screening module and social worker assessment forms questions, such as housing instability, food insecurity, transportation barriers, legal issues, and financial strain [31–35]. Trained data collection staff recruited patients to complete the questionnaire during their ED encounter. We also included an invitation to complete the questionnaire online within the discharge paperwork. Data collection was offered in both English and Spanish. Subjects self-reported their HRSNs by completing the questionnaire in REDCap [36,37]. Subjects provided written informed consent to link their responses to existing EHR, administrative, and health information exchange data. We created binary indicators of the presence or absence of each HRSN [31–35].

EHR and other clinical data sources.   Eskenazi Health's EHR was the primary source of demographic, clinical, encounter, and billing history data. For each patient completing the survey, we extracted the past 12 months of encounter types, procedure codes, orders, diagnoses, prescriptions, payer history, address history, contact information, age, and gender. Using all patients' clinical documents and notes from the past 12 months, we applied our previously validated natural language processing (NLP) algorithms to identify housing instability, financial strain, history of legal problems, food insecurity, transportation barriers, and unemployment [38,39]. We also linked historical billing and payment data.

In addition, we linked responses to the data available for each respondent in the Indiana Network for Patient Care (INPC). The INPC is one of the largest and oldest multi-institutional clinical repositories [40], with over 13 billion data elements on 15 million patients from more than 117 different hospitals and 18,000 practices across the state. These data provided additional information on patient encounters and diagnoses outside Eskenazi Health. Lastly, using a patient address, we linked to the most recent area deprivation index scores [41]. The transformation of these data sources into specific features is detailed in Features for EHR models, below. All data linkage was completed between April 2023 and December 2023. The authors had no access to information that could identify individual participants during or after data collection.

Prediction target

We defined our prediction as the patient's likely need for an HRSN service within the next 30 days of survey completion. Patients who met this definition were identified by a positive indication from the following data sources: referrals, encounters, scheduling data, orders, or clinical notes [21,42–44]. Eskenazi Health offers social workers, financial counseling, or medical/legal partnerships "in-house" through referrals. A positive indicator of a likely HRSN needs to include any referral to these services (whether resulting in an appointment or not), any encounter with these services (whether scheduled, kept, missed, or "no-show"), or any order mentioning these services or HRSNs. Order text or clinical notes that mentioned referrals to, or the need for, social workers, case managers, financial counselors, or case conferences were identified through keyword searches [42]. Data collection staff and study team members did

not share results from the questionnaires with clinicians, so any decisions or actions did not reflect patients' answers to study data collection.

Features for ML EHR models

Following our prior work, we extracted and engineered over 100 features that could indicate a patient with a likely need for HRSNs services [45]. The individual patient features were reflected in the following categories. *Demographics and contact information* included features derived from patient's addresses or contact information. *Encounter history* included all features representing visit types and frequencies. The *Clinical* class included all features resulting from diagnoses, documentation, or care delivery processes. The *HRSN screeners* class included all prior social needs screening results. The *Text* class included the indicators extracted via NLP. The *Geospatial* class included the area deprivation index. The *Financial* class included features derived from billing, payment, and insurance data. All measures were for the 12 months before the completion of the screening instrument. Throughout the paper, we collectively refer to all these features as the EHR-based data and detail in S1 File.

Modeling and comparative analyses

To develop the ML EHR models, we built ML classifiers using XGBoost [46], an ensemble-based classification algorithm that employs gradient boosting to add decision trees to address errors in prior predictions, resulting in a robust decision model. Models were developed in Python using 5-fold cross-validation and a grid search for hyperparameter tuning. To support interpretation, we extracted XGBoost's feature importance scores (based on F scores) and utilized the SHAP (SHapley Additive exPlanations) method to summarize the contributions of features to the models [47]. Models were assessed on the validation set using sensitivity, specificity, and Area Under the Curve (AUC) values. The HRSN screening questions were not included in the ML EHR models.

For the screening question models, we also built an XGBoost classification algorithm. The input features were the binary indicator of a positive screen for each of the 5 assessed HRSN. We followed the same modeling approach as above and calculated sensitivity, specificity, and AUC values. We adopted the approach purposefully. For one, screening questionnaires with multiple HRSNs do not include guidelines for creating an overall risk score [48]. As our prediction target includes services that potentially reflect multiple HRSNs or professionals capable of addressing various HRSNs, we required the screening questions to produce a single prediction result. Additionally, we wanted a consistent comparison between the screening question models and the ML EHR models; using alternative classification methodologies would have introduced potential alternative explanations for performance differences.

We recognize that ML modeling of healthcare data may demonstrate or perpetuate biases against certain populations [26] and that differences in access to care or other structural barriers could influence the distribution of our outcome, screening questionnaire responses, or measures of prior utilization by demographics (S2 File). Therefore, we took several steps in the analysis to mitigate bias. First, we ran all models with and without age and gender to identify potential differences for these groups of patients. Second, race and ethnicity have been used with uncertain purpose in past algorithms [49]. Therefore, we did not include race or ethnicity as input features in any model. Instead, we stratified all models by patient self-reported race and ethnicity as a check on model fairness (i.e., consistent performance).

We described the sample using frequencies and percentages. We used the equality of proportions test to compare AUC values using the Delong test and the difference in sensitivity, specificity, and PPV [50].

Results

The study sample was predominately female (58.9%) and reflective of an urban, safety-net ED population, with less than a third of the sample being non-Hispanic White (31.5%). The HRSNs were common, with half of the patients reporting food insecurity (46.0%) and housing instability (46.0%) (Table 1). Almost half (41%) of the patients had a positive indicator for the likely HRSN service needs within the next 30 days, as identified through referrals, encounters, scheduling data, orders, or clinical notes (Table 1).

Patients with an indicator for the likely HRSN service needs in the next 30 days differed from those without on several characteristics. Patients with an indicator for the HRSN service needs were older, had more comorbidities, and had higher prior utilization. Additionally, the distribution of patients with and without an indicator for the HRSN service needs in the next 30 days differed by race and ethnicity (p = 0.003). Finally, patients with an indicator for the HRSN service needs had higher reported financial strain (25.3% vs. 17.5%; p = 0.002) and transportation barriers (44.6% vs. 33.1%; p<0.001).

ML EHR data model

The model using the EHR-based data sources (without demographics) had an AUC = 0.765 (95%CI = 0.737, 0.792) (Table 2). The model was more specific (0.790; 95%CI = 0.753, 0.822) than sensitive (0.600; 95%CI = 0.560, 0.640), and the positive predictive value (PPV) was 0.756 (95%CI = 0.715, 0.793). Including patient age and gender in the models resulted in a small but

Table 1. Adult emergency department patients by indication of health-related social need (HRSN) service need, Indianapolis, IN.

	Total	No HRSN service	HRSN service need in next 30 days	p-value
	n=1,101	n=650	n=451	
Age (mean, sd)	41.5 (15.1)	38.7 (14.5)	45.6 (15.1)	<0.001
Female gender	649 (58.9%)	397 (61.1%)	252 (55.9%)	0.084
Race & ethnicity				0.003
Asian	11 (1.0%)	8 (1.2%)	3 (0.7%)	
Black non-Hispanic	511 (46.4%)	322 (49.5%)	189 (41.9%)	
Hispanic	203 (18.4%)	116 (17.8%)	87 (19.3%)	
Multiple	14 (1.3%)	7 (1.1%)	7 (1.6%)	
Other / unknown	15 (1.4%)	14 (2.2%)	1 (0.2%)	
White non-Hispanic	347 (31.5%)	183 (28.2%)	164 (36.4%)	
Language other than English	129 (11.7%)	68 (10.5%)	61 (13.5%)	0.12
Encounter history[a]				
Inpatient admissions (mean, sd)	0.8 (3.0)	0.5 (1.7)	1.3 (4.2)	<0.001
Emergency department visits (mean, sd)	5.2 (14.0)	4.2 (9.2)	6.5 (18.9)	0.007
Primary care visits (mean, sd)	3.3 (5.4)	2.8 (4.6)	4.0 (6.4)	<0.001
Elixhauser co-morbidity score (mean, sd)	1.5 (1.9)	1.0 (1.3)	2.3 (2.2)	<0.001
Health-related social needs[b]				
Financial strain	228 (20.7%)	114 (17.5%)	114 (25.3%)	0.002
Food insecurity	692 (62.9%)	408 (62.8%)	284 (63.0%)	0.95
Housing instability	506 (46.0%)	287 (44.2%)	219 (48.6%)	0.15
Transportation barriers	416 (37.8%)	215 (33.1%)	201 (44.6%)	<0.001
Legal problems	239 (21.7%)	128 (19.7%)	111 (24.6%)	0.052

[a]during 12 months prior to survey date.

[b]per Epic SDOH screening questions.

https://doi.org/10.1371/journal.pone.0312193.t001

Table 2. Comparison of models using screening questionnaire and EHR data in predicting need for health-related social service within 30 days of emergency department visit.

	ML EHR model	ML EHR model with age & gender	Screening questionnaire model	Screening questionnaire model with age & gender
AUC	0.765 (0.737, 0.792)	0.772 (0.744, 0.800)	0.580 (0.546, 0.611)	0.640 (0.609, 0.672)
Sensitivity	0.600 (0.560, 0.640)	0.646 (0.606, 0.684)	0.578 (0.537, 0.617)	0.613 (0.572, 0.652)
Specificity	0.790 (0.753, 0.822)	0.765 (0.727, 0.799)	0.566 (0.524, 0.608)	0.583 (0.541, 0.625)
PPV	0.756 (0.715, 0.793)	0.749 (0.709, 0.785)	0.591 (0.550, 0.631)	0.615 (0.574, 0.654)

AUC = area under the curve.

PPV = positive predictive value.

https://doi.org/10.1371/journal.pone.0312193.t002

not statistically significant increase in overall performance (AUC = 0.772; 95%CI = 0.744, 0.800; p = 0.5000).

Screening questionnaire model. The screening questionnaire model performed worse than the ML HER model (Table 2). Using only the screening questions, the AUC value of 0.580 (95%CI = 0.546, 0.611) was lower than the ML model (p < 0.0001). Additionally, while sensitivity was not statistically different between the two models, the specificity of the screening questionnaire model (0.566; 95%CI = 0.524, 0.608) was lower than the ML EHR model (p<0.0001). The positive predictive value for the screening questionnaire model was 0.591 (95%CI = 0.550, 0.631), which was also lower than the ML EHR model (p<0.0001). Adding patient age and gender to the screening questionnaire model resulted in slight improvements in overall performance (AUC = 0.640; 95%CI = 0.609, 0.672; p = 0.0117). However, this model still had lower performance than the ML EHR model with demographics (p < 0.0001).

The differential performance is illustrated in Fig 2, with the largest ROC curves reflecting the ML predictive models. For the ML predictive models without controlling for demographics (S3 File Panel A), the five most important features in predicting an indication for the HRSN need for services were the patient's Elixhauser comorbidity score, a prior housing instability
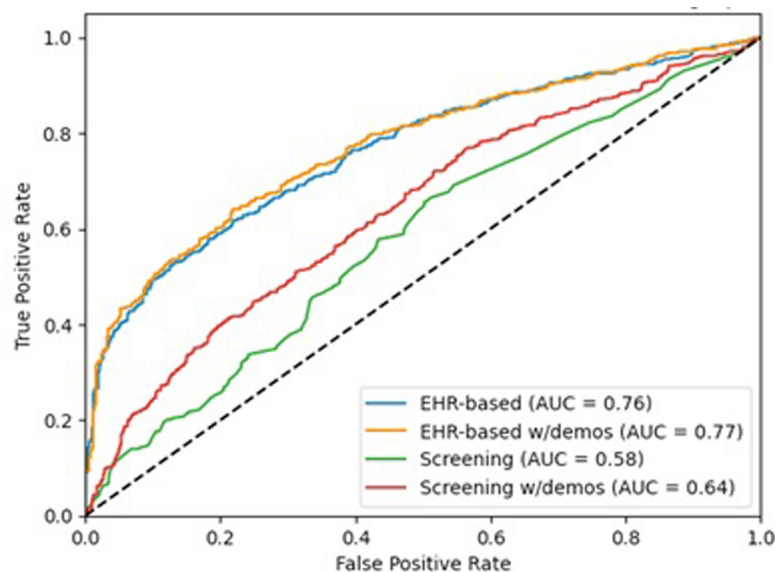

Fig 2. ROC Curves of EHR and screening question models predicting the need for HRSN service within 30 days of an emergency department visit, with and without age and gender.

https://doi.org/10.1371/journal.pone.0312193.g002

ICD-10 Z code, number of different therapeutic classes prescribed, legal problems documented in clinical notes, and total primary care visits. Including patient age and gender in the ML predictive models resulted in the same top five features, with some variation in ordering (S3 File Panel B). Notably, age was the sixth most important feature in the model with demographics, outranking other direct measures of HRSN. The most important feature of the screening questionnaire model was the transportation barrier (S3 File Panel C). Including demographics in the screening questionnaire model made age the most important feature (Fig 3 Panel D).

Stratification by race and ethnicity.   The ML predictive and screening questionnaire models stratified by race and ethnicity varied in their overall performance (Fig 3). The stratified AUCs varied significantly by the patient's race and ethnicity for the ML predictive models (p = 0.0241) and the screening questionnaire model (p = 0.0392). Differences by race and ethnicity persisted even if the minor " *other* " group was omitted from the analyses. Beyond overall performance, notable differences existed in other performance measures (Table 3).

In both the ML predictive and screening questionnaire models, the sensitivity was roughly equal to or higher than the specificity for White, non-Hispanic patients. However, for all other respondents, both models had higher specificity than sensitivity. In other words, both models were better at finding White, non-Hispanic patients with HRSN service needs than it was at finding patients from other races and ethnicities with HRSN service needs. Additionally, which features were most important changed when each model was stratified by race and ethnicity (S3 File).

Discussion

In predicting which adult ED patients would need HRSN services within the next 30 days, ML predictive models leveraging several robust EHR data sources outperformed models using
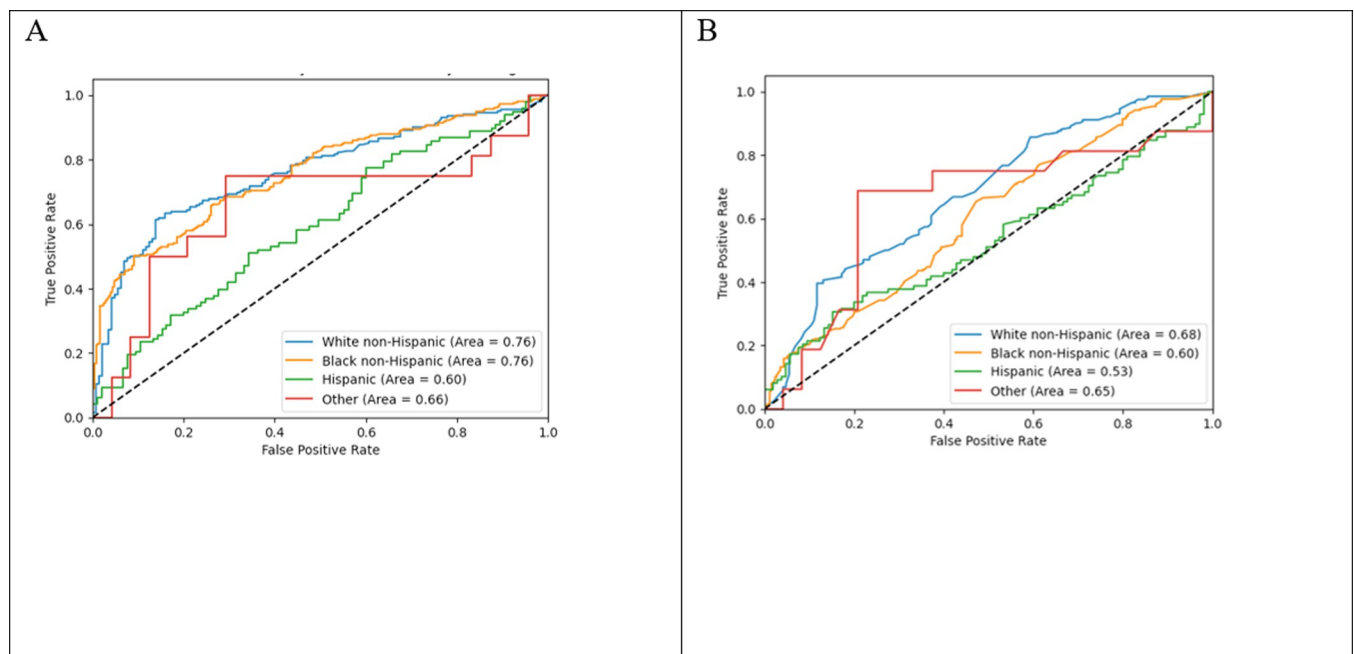

Fig 3. ROC Curves of (A) EHR-based and (B) screening question models predicting need for HRSN service within 30 days of emergency department visit by race and ethnicity.

https://doi.org/10.1371/journal.pone.0312193.g003

Table 3. Comparison of models using screening questionnaire and EHR data in predicting need for health-related social service within 30 days of emergency department visit stratified by self-reported race and ethnicity.

	ML EHR model				Screening questionnaire model			
	White Non-Hispanic	Black Non-Hispanic	Hispanic	Other Non-Hispanic	White Non-Hispanic	Black Non-Hispanic	Hispanic	Other Non-Hispanic
	n=347	**n=511**	**n=203**	**n=40**	**n=347**	**n=511**	**n=203**	**n=40**
AUC	0.766	0.753	0.601	0.577	0.659	0.536	0.454	0.404
Sensitivity	0.703	0.591	0.357	0.188	0.678	0.370	0.276	0.063
Specificity	0.697	0.748	0.676	0.875	0.545	0.756	0.657	0.875
PPV	0.763	0.704	0.507	0.500	0.675	0.605	0.429	0.250

https://doi.org/10.1371/journal.pone.0312193.t003

screening questions only. While comparatively better, the ML predictive models using EHR data sources did not exhibit particularly strong performance, and all models demonstrated bias.

Collecting HRSN information is a growing necessity for healthcare organizations due to quality reporting requirements as part of CMS' Hospital Inpatient Quality Reporting (IQR) program [51] and the Merit-based Incentive Payment System (MIPS) [52], new National Committee on Quality Assurance (NCQA) HEDIS quality measures [53], and Joint Commission accreditation requirements [54]. The collection of HRSN data can be used to drive individual patient referrals, help inform providers' decision-making processes, and support organizational quality improvement and measurement efforts [27]. However, our findings indicate that models using the HRSN screening questions were poor predictors of future HRSN service needs. Indeed, these models were only slightly better than a coin flip. The low performance of models using screening questions could stem from the overwhelming number of positive results, making it hard to identify who requires services in a universally high-risk group. This criticism is not unique to HRSN questionnaires; although ML predictive models using EHR data sources performed better, they failed to reach accepted thresholds for clinical usefulness [55]. Nevertheless, the limited ability of the HRSN information to predict who needs HRSN services can pose a challenge, as HRSNs are often highly prevalent among patients at many healthcare organizations [10,56].

The widely utilized general HRSNs screening questionnaires, such as PRAPARE, Epic, HealthLeads or CMS's Accountable Health Communities tool, do not account for demographics. Although, PRAPARE's tally method of scoring for overall social risk includes race and ethnicity as contributing factors [57]. However, in this study, adding gender and age led to minor improvements in the performance of the ML EHR model and significant improvements in the screening questionnaire model. HRSNs vary across patient age groups [58], and when added to the screening question models, age was the most important feature. Thus, one potential improvement to HRSN screening questions could be incorporating age. Some domain-specific screening instruments, like the Consumer Financial Protection Bureau's Financial Well-Being Scale [59], account for differences between age groups. Additionally, women are likely to experience poverty at higher rates than men [58] but adding gender was not as important in the models. A potential path for future work could be the appropriate and equitable integration of age and other demographics into screening question interpretation.

Overall, the ML EHR model performed better than the screening questionnaire model. The ML EHR modeling approach has the obvious advantage of drawing upon more information and, thus, is an increasingly preferred approach to HRSN measurement [19–21]. While model performance still requires significant improvement, our findings highlight the potential value of several EHR data elements already accessible to healthcare organizations. Prior ICD-10 Z-codes and HRSN documented in clinical notes were important model features. As healthcare

organizations already collect and maintain the HRSN data, it can be readily leveraged to support existing HRSN measurement activities as potential means of imputing missing responses or by driving decision support systems on patients for additional HRSN screening.

Problematically, the ML EHR and the screening questionnaire models exhibited differential performance and bias across patient race and ethnicity. Biased models perpetuate health inequities [59,60]. The bias observed in our models is likely due to both measurement bias and representation bias [61]. Measurement bias can affect the performance of the ML predictive and screening questionnaire models. For example, while data points like ICD-10 Z codes are potentially informative, they are often underutilized and inconsistent [62]. Even though the screening questions are not perfect instruments, they have psychometric properties that vary according to specific HRSNs [15]. Representation bias may be more influential for the ML EHR models, as many of the most important features were correlated with access to care and services. To have a higher Elixhauser score or more medication classes or prior diagnoses requires access to the healthcare system for diagnoses and services. This is important because access to care is not equivalent across populations [63]. Although we recruited a diverse sample, that may not account for potential differences in the patients' underlying data due to differential experiences with the healthcare system.

This study reinforces the need for a detailed examination of the differential performance of risk prediction models. Simply focusing on the differential overall performance could miss the actual effects such models can have in practice. Of particular concern is the fact that the observed bias predominately affected model sensitivity, which is the proportion of true positives that are correctly identified as positive [64]. As such, the models, including those using the screening questions, would under-identify members of certain race and ethnic groups needing HRSN services. If applied to care practices, under-identification would run contrary to patients', physicians', staff members', and organizations' expectations for HRSN measurement and intervention activities [2,65]. Such differential under-identification would likely perpetuate health inequities. Obviously, before implementing practice, models must be checked for bias. In addition, the effectiveness and impact of any bias mitigation strategies requires evaluation [66].

Limitations

This study faces several limitations, particularly in terms of external validity. First, our models predict a future event–the likely need for HRSN services. However, the availability of the data, the nature of the services, and workflow processes unique to the setting may limit the generalizability of the outcome to other settings. Our models may not be generalizable either. We can draw on a large set of EHR and health information exchange data, including financial information and data extracted via NLP. Likewise, our choice of screening questions may differ from other available screening tools. Additionally, our findings may not hold over time as screening practices tend to degrade [7] and prediction models drift [67] over time.

Lastly, HRSN screening via questionnaires aims to drive referrals to services to address patient needs [4,68]. However, our choice of outcome measure may have limited the potential predictive ability of our models. It is possible the screening questions better capture the concept of social "risk" [69], that is, the presence of a potentially adverse factor to health. By contrast, the predicted outcome may have better reflected the concept of social "need", that is, an immediate concern that aligns with patient's priorities and preferences [69,70]. These are related, but distinct, concepts.

Conclusions

The ML predictive model performed better than the screening questionnaire model, yet both models demonstrated biases. Additional work is needed to design predictive models that effectively identify all patients with HRSNs.

Supporting information

S1 File. Engineered features from electronic health record (EHR) and health information exchange (HIE) data (past 12 months).
(DOCX)

S2 File. Comparison of features.
(DOCX)

S3 File. Feature importance values by race and ethnicity.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Olena Mazurenko, Adam T. Hirsh, Christopher A. Harle, Joanna Shen, Cassidy McNamee, Joshua R. Vest.

Data curation: Olena Mazurenko, Joanna Shen.

Formal analysis: Olena Mazurenko, Joanna Shen.

Funding acquisition: Joshua R. Vest.

Methodology: Adam T. Hirsh, Joshua R. Vest.

Software: Joanna Shen.

Supervision: Olena Mazurenko, Joshua R. Vest.

Validation: Olena Mazurenko, Joshua R. Vest.

Writing – original draft: Olena Mazurenko, Adam T. Hirsh, Christopher A. Harle, Cassidy McNamee, Joshua R. Vest.

Writing – review & editing: Olena Mazurenko, Adam T. Hirsh, Christopher A. Harle, Joshua R. Vest.

References

1. Schiavoni KH, Helscel K, Vogeli C, Thorndike AN, Cash RE, Camargo CA, et al. Prevalence of social risk factors and social needs in a Medicaid Accountable Care Organization (ACO). BMC Health Serv Res. 2022; 22: 1375. https://doi.org/10.1186/s12913-022-08721-9 PMID: 36403024

2. Beidler LB, Razon N, Lang H, Fraze TK. "More than just giving them a piece of paper": Interviews with Primary Care on Social Needs Referrals to Community-Based Organizations. J Gen Intern Med. 2022; 37: 4160–4167. https://doi.org/10.1007/s11606-022-07531-3 PMID: 35426010

3. Yan AF, Chen Z, Wang Y, Campbell JA, Xue Q-L, Williams MY, et al. Effectiveness of Social Needs Screening and Interventions in Clinical Settings on Utilization, Cost, and Clinical Outcomes: A Systematic Review. Health Equity. 2022; 6: 454–475. https://doi.org/10.1089/heq.2022.0010 PMID: 35801145

4. O'Gurek DT, Henke C. A Practical Approach to Screening for Social Determinants of Health. Fam Pract Manag. 2018; 25: 7–12. Available: http://www.ncbi.nlm.nih.gov/pubmed/29989777. PMID: 29989777

5. Gurewich D, Garg A, Kressin NR. Addressing Social Determinants of Health Within Healthcare Delivery Systems: a Framework to Ground and Inform Health Outcomes. J Gen Intern Med. 2020; 35: 1571–1575. https://doi.org/10.1007/s11606-020-05720-6 PMID: 32076989

6. Trochez RJ, Sharma S, Stolldorf DP, Mixon AS, Novak LL, Rajmane A, et al. Screening Health-Related Social Needs in Hospitals: A Systematic Review of Health Care Professional and Patient Perspectives. Popul Health Manag. 2023; 26: 157–167. https://doi.org/10.1089/pop.2022.0279 PMID: 37092962

7. Savitz ST, Nyman MA, Kaduk A, Loftus C, Phelan S, Barry BA. Association of Patient and System-Level Factors With Social Determinants of Health Screening. Med Care. 2022; 60: 700–708. https://doi.org/10.1097/MLR.0000000000001754 PMID: 35866557

8. Patel S, Moriates C, Valencia V, Garza K de la, Sanchez R, Leykum LK, et al. A Hospital-Based Program to Screen for and Address Health-Related Social Needs for Patients Admitted with COVID-19. J Gen Intern Med. 2022; 37: 2077. https://doi.org/10.1007/s11606-022-07550-0 PMID: 35411525

9. Dolce M, Keedy H, Chavez L, Boch S, Zaim H, Dias B, et al. Implementing an EMR-based Health-related Social Needs Screen in a Pediatric Hospital System. Pediatr Qual Saf. 2022; 7: e512. https://doi.org/10.1097/pq9.0000000000000512 PMID: 35071955

10. Wallace AS, Luther B, Guo J-W, Wang C-Y, Sisler S, Wong B. Implementing a Social Determinants Screening and Referral Infrastructure During Routine Emergency Department Visits, Utah, 2017–2018. Prev Chronic Dis. 2020; 17: E45. https://doi.org/10.5888/pcd17.190339 PMID: 32553071

11. Drake C, Reiter K, Weinberger M, Eisenson H, Edelman D, Trogdon JG, et al. The Direct Clinic-Level Cost of the Implementation and Use of a Protocol to Assess and Address Social Needs in Diverse Community Health Center Primary Care Clinical Settings. J Health Care Poor Underserved. 2021; 32: 1872–1888. https://doi.org/10.1353/hpu.2021.0171 PMID: 34803048

12. Boch S, Keedy H, Chavez L, Dolce M, Chisolm D. An Integrative Review of Social Determinants of Health Screenings used in Primary Care Settings. J Health Care Poor Underserved. 2020; 31: 603–622. Available: https://muse.jhu.edu/pub/1/article/756664. https://doi.org/10.1353/hpu.2020.0048 PMID: 33410796

13. Department of Health & Human Services. Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long-Term Care Hospital Prospective Payment System and Policy Changes and Fiscal Year 2023 Rates; Quality Programs and Medicare Promoting Interoperability Program Requirements for Eligible Hospitals and Critical Access Hospitals; Costs Incurred for Qualified and Non-Qualified Deferred Compensation Plans; and Changes to Hospital and Critical Access Hospital Conditions of Participation. Fed Regist. 2022;87: 48780–49499. Available: https://www.federalregister.gov/documents/2022/08/10/2022-16472/medicare-program-hospital-inpatient-prospective-payment-systems-for-acute-care-hospitals-and-the.

14. Henrikson NB, Blasi PR, Dorsey CN, Mettert KD, Nguyen MB, Walsh-Bailey C, et al. Psychometric and Pragmatic Properties of Social Risk Screening Tools: A Systematic Review. Am J Prev Med. 2019; 57: S13–S24. https://doi.org/10.1016/j.amepre.2019.07.012 PMID: 31753276

15. Harle CA, Wu W, Vest JR. Accuracy of Electronic Health Record Food Insecurity, Housing Instability, and Financial Strain Screening in Adult Primary Care. JAMA. 2023; 329: 423–424. https://doi.org/10.1001/jama.2022.23631 PMID: 36749341

16. Garg A, Boynton-Jarrett R, Dworkin PH. Avoiding the Unintended Consequences of Screening for Social Determinants of Health. JAMA. 2016; 316: 813–4. https://doi.org/10.1001/jama.2016.9282 PMID: 27367226

17. Bazemore AW, Cottrell EK, Gold R, Hughes LS, Phillips RL, Angier H, et al. "Community vital signs": incorporating geocoded social determinants into electronic records to promote patient and population health. J Am Med Inf Assoc. 2016; 23: 407–412. https://doi.org/10.1093/jamia/ocv088 PMID: 26174867

18. Cantor MN, Chandras R, Pulgarin C. FACETS: using open data to measure community social determinants of health. J Am Med Inform Assoc JAMIA. 2017; 25: 419–422. https://doi.org/10.1093/jamia/ocx117 PMID: 29092049

19. Hatef E, Rouhizadeh M, Tia I, Lasser E, Hill-Briggs F, Marsteller J, 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. https://doi.org/10.2196/13802 PMID: 31376277

20. Holcomb J, Oliveira LC, Highfield L, Hwang KO, Giancardo L, Bernstam EV. Predicting health-related social needs in Medicaid and Medicare populations using machine learning. Sci Rep. 2022; 12: 4554. https://doi.org/10.1038/s41598-022-08344-4 PMID: 35296719

21. Vest JR, Menachemi N, Grannis SJ, Ferrell JL, Kasthurirathne SN, Zhang Y, et al. Impact of Risk Stratification on Referrals and Uptake of Wraparound Services That Address Social Determinants: A Stepped Wedged Trial. Am J Prev Med. 2019; 56: e125–e133. https://doi.org/10.1016/j.amepre.2018.11.009 PMID: 30772150

22. Anderson ES, Hsieh D, Alter HJ. Social Emergency Medicine: Embracing the Dual Role of the Emergency Department in Acute Care and Population Health. Ann Emerg Med. 2016; 68: 21–25. https://doi.org/10.1016/j.annemergmed.2016.01.005 PMID: 26921967

23. Capp R, Kelley L, Ellis P, Carmona J, Lofton A, Cobbs-Lomax D, et al. Reasons for Frequent Emergency Department Use by Medicaid Enrollees: A Qualitative Study. Acad Emerg Med. 2016; 23: 476–481. https://doi.org/10.1111/acem.12952 PMID: 26932230

24. Axelson DJ, Stull MJ, Coates WC. Social Determinants of Health: A Missing Link in Emergency Medicine Training. AEM Educ Train. 2018; 2: 66–68. https://doi.org/10.1002/aet2.10056 PMID: 30051070

25. Parast L, Mathews M, Martino S, Lehrman WG, Stark D, Elliott MN. Racial/Ethnic Differences in Emergency Department Utilization and Experience. J Gen Intern Med. 2022; 37: 49–56. https://doi.org/10.1007/s11606-021-06738-0 PMID: 33821410

26. Parikh RB, Teeple S, Navathe AS. Addressing Bias in Artificial Intelligence in Health Care. JAMA. 2019; 322: 2377–2378. https://doi.org/10.1001/jama.2019.18058 PMID: 31755905

27. National Academies of Sciences Engineering, Medicine. Integrating Social Care into the Delivery of Health Care: Moving Upstream to Improve the Nation's Health. Washington, DC: The National Academies Press; 2019. doi:10.17226/25467.

28. Chagin K, Choate F, Cook K, Fuehrer S, Misak JE, Sehgal AR. A Framework for Evaluating Social Determinants of Health Screening and Referrals for Assistance. J Prim Care Community Health. 2021; 12: 21501327211052204. https://doi.org/10.1177/21501327211052204 PMID: 34905991

29. Tubbs RE, Warner L, Shy BD, Manikowski C, Roosevelt GE. A descriptive study of screening and navigation on health-related social needs in a safety-net hospital emergency department. Am J Emerg Med. 2023; 74: 65–72. https://doi.org/10.1016/j.ajem.2023.09.007 PMID: 37778164

30. Kulie P, Steinmetz E, Johnson S, McCarthy ML. A health-related social needs referral program for Medicaid beneficiaries treated in an emergency department. Am J Emerg Med. 2021; 47: 119–124. https://doi.org/10.1016/j.ajem.2021.03.069 PMID: 33799141

31. Hall MH, Matthews KA, Kravitz HM, Gold EB, Buysse DJ, Bromberger JT, et al. Race and Financial Strain are Independent Correlates of Sleep in Midlife Women: The SWAN Sleep Study. Sleep. 2009; 32: 73–82. https://doi.org/10.5665/sleep/32.1.73 PMID: 19189781

32. Puterman E, Haritatos J, Adler N, Sidney S, Schwartz JE, Epel E. Indirect effect of financial strain on daily cortisol output through daily negative to positive affect index in the Coronary Artery Risk Development in Young Adults Study. Psychoneuroendocrinology. 2013; 38: 10.1016/j.psyneuen.2013.07.016. https://doi.org/10.1016/j.psyneuen.2013.07.016 PMID: 23969421

33. Children's HealthWatch. The Hunger Vital SignTM: Best practices for screening and intervening to alleviate food insecurity. 2016 [cited 16 Nov 2020]. Available: https://childrenshealthwatch.org/wp-content/uploads/CHW_HVS_whitepaper_FINAL.pdf.

34. Hager ER, Quigg AM, Black MM, Coleman SM, Heeren T, Rose-Jacobs R, et al. Development and validity of a 2-item screen to identify families at risk for food insecurity. Pediatrics. 2010; 126: e26–32. https://doi.org/10.1542/peds.2009-3146 PMID: 20595453

35. National Association of Community Health Centers, Association of Asian Pacific Community Health Organizations, Oregon Primary Care Association. PRAPARE: Implementation and action toolkit. 2019 [cited 1 Feb 2023]. Available: https://prapare.org/wp-content/uploads/2022/09/Full-Toolkit_June-2022_Final.pdf.

36. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009; 42: 377–381. https://doi.org/10.1016/j.jbi.2008.08.010 PMID: 18929686

37. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform. 2019; 95: 103208. https://doi.org/10.1016/j.jbi.2019.103208 PMID: 31078660

38. Allen KS, Hood DR, Cummins J, Kasturi S, Mendonca EA, Vest JR. Natural language processing-driven state machines to extract social factors from unstructured clinical documentation. JAMIA Open. 2023; 6: ooad024. https://doi.org/10.1093/jamiaopen/ooad024 PMID: 37081945

39. Magoc T, Allen KS, McDonnell C, Russo J-P, Cummins J, Vest JR, et al. Generalizability and portability of natural language processing system to extract individual social risk factors. Int J Med Inf. 2023; 177: 105115. https://doi.org/10.1016/j.ijmedinf.2023.105115 PMID: 37302362

40. McDonald CJ, Overhage JM, Barnes M, Schadow G, Blevins L, Dexter PR, et al. The Indiana Network For Patient Care: A Working Local Health Information Infrastructure. Health Aff (Millwood). 2005; 24: 1214–1220. https://doi.org/10.1377/hlthaff.24.5.1214 PMID: 16162565

41. Kind AJH, Buckingham WR. Making Neighborhood-Disadvantage Metrics Accessible - The Neighborhood Atlas. N Engl J Med. 2018; 378: 2456–2458. https://doi.org/10.1056/NEJMp1802313 PMID: 29949490

42. 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 Inf. 2017; 107: 101–106. https://doi.org/10.1016/j.ijmedinf.2017.09.008 PMID: 29029685

43. Kasthurirathne SN, Vest J, Menachemi N, Halverson PK, Grannis SJ. Assessing the capacity of social determinants of health data to augment predictive models identifying patients in need of wraparound social services. J Am Med Inform Assoc. 2018; 25: 47–53. https://doi.org/10.1093/jamia/ocx130 PMID: 29177457

44. Kasthurirathne SN, Grannis S, Halverson PK, Morea J, Menachemi N, Vest JR. Precision Health–Enabled Machine Learning to Identify Need for Wraparound Social Services Using Patient- and Population-Level Data Sets: Algorithm Development and Validation. JMIR Med Inform. 2020; 8: e16129. https://doi.org/10.2196/16129 PMID: 32479414

45. Vest JR, Adler-Milstein J, Gottlieb LM, Bian J, Campion TR, Cohen GR, et al. Assessment of structured data elements for social risk factors. Am J Manag Care. 2022; 28: e14–e23. https://doi.org/10.37765/ajmc.2022.88816 PMID: 35049262

46. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: Association for Computing Machinery; 2016. pp. 785–794. https://doi.org/10.1145/2939672.2939785

47. Lundberg S, Lee S-I. A Unified Approach to Interpreting Model Predictions. arXiv; 2017. https://doi.org/10.48550/arXiv.1705.07874

48. Moen M, Storr C, German D, Friedmann E, Johantgen M. A Review of Tools to Screen for Social Determinants of Health in the United States: A Practice Brief. Popul Health Manag. 2020; 23: 422–429. https://doi.org/10.1089/pop.2019.0158 PMID: 31910355

49. Vyas DA, Eisenstein LG, Jones DS. Hidden in Plain Sight—Reconsidering the Use of Race Correction in Clinical Algorithms. N Engl J Med. 2020; 383: 874–882. https://doi.org/10.1056/NEJMms2004740 PMID: 32853499

50. StataCorp. Stata 17 Base Reference Manual. College Station, TX: Stata PRess; 2021.

51. Centers for Medicare & Medicaid Services. Screening for Social Drivers of Health. In: Centers for Medicare and Medicaid Services Measures Inventory Tool [Internet]. 2023 [cited 25 Mar 2024]. Available: https://cmit.cms.gov/cmit/#/MeasureView?variantId=10405&sectionNumber=1.

52. The Physicians Foundation. Quality ID #487: Screening for Social Drivers of Health. In: Quality Payment Program [Internet]. 2022 [cited 25 Mar 2024]. Available: https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2023_Measure_487_MIPSCQM.pdf.

53. Reynolds A. Social Need: New HEDIS Measure Uses Electronic Data to Look at Screening, Intervention. In: NCQA [Internet]. 2 Nov 2022 [cited 20 Dec 2023]. Available: https://www.ncqa.org/blog/social-need-new-hedis-measure-uses-electronic-data-to-look-at-screening-intervention/.

54. The Joint Commission. New Requirements to Reduce Health Care Disparities. 2022. Report No.: Issue 36. Available: https://www.jointcommission.org/-/media/tjc/documents/standards/r3-reports/r3_disparities_july2022-6-20-2022.pdf.

55. de Hond AAH, Steyerberg EW, van Calster B. Interpreting area under the receiver operating characteristic curve. Lancet Digit Health. 2022; 4: e853–e855. https://doi.org/10.1016/S2589-7500(22)00188-1 PMID: 36270955

56. Kusnoor SV, Koonce TY, Hurley ST, McClellan KM, Blasingame MN, Frakes ET, et al. Collection of social determinants of health in the community clinic setting: a cross-sectional study. BMC Public Health. 2018; 18: 550. https://doi.org/10.1186/s12889-018-5453-2 PMID: 29699539

57. National Association of Community Health Centers. PRAPARE® Risk Tally Score Quick Sheet. 5 Oct 2022 [cited 28 May 2024]. Available: https://prapare.org/knowledge-center/data-documentation-and-clinical-integration-resources/https-prapare-org-wp-content-uploads-2022-10-prapare-risk-tally-score-quick-sheet-1-pdf/.

58. Sharma M. Poverty and Gender: Determinants of Female- and Male-Headed Households with Children in Poverty in the USA, 2019. Sustainability. 2023; 15: 7602. https://doi.org/10.3390/su15097602

59. Tan M, Hatef E, Taghipour D, Vyas K, Kharrazi H, Gottlieb L, et al. Including Social and Behavioral Determinants in Predictive Models: Trends, Challenges, and Opportunities. JMIR Med Inform. 2020; 8. https://doi.org/10.2196/18084 PMID: 32897240

60. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. JAMA Intern Med. 2018; 178: 1544–1547. https://doi.org/10.1001/jamainternmed.2018.3763 PMID: 30128552

61. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A Survey on Bias and Fairness in Machine Learning. ACM Comput Surv. 2022; 54: 1–35. https://doi.org/10.1145/3457607

62. Maksut J, Hodge C, Van C, Razmi A, Khau M. Utilization of Z Codes for Social Determinants of Health among Medicare Fee-for-Service Beneficiaries, 2019. Baltimore, MD: Centers for Medicare & Medicaid Services; 2021. Report No.: 24. Available: https://www.cms.gov/files/document/z-codes-data-highlight.pdf.

63. Mahajan S, Caraballo C, Lu Y, Valero-Elizondo J, Massey D, Annapureddy AR, et al. Trends in Differences in Health Status and Health Care Access and Affordability by Race and Ethnicity in the United States, 1999–2018. JAMA. 2021; 326: 637–648. https://doi.org/10.1001/jama.2021.9907 PMID: 34402830

64. Altman DG, Bland JM. Statistics Notes: Diagnostic tests 1: sensitivity and specificity. BMJ. 1994; 308: 1552. https://doi.org/10.1136/bmj.308.6943.1552 PMID: 8019315

65. Mazurenko O, Hirsh AT, Harle CA, McNamee C, Vest JR. Health-related social needs information in the emergency department: clinician and patient perspectives on availability and use. BMC Emerg Med. 2024; 24: 45. https://doi.org/10.1186/s12873-024-00959-2 PMID: 38500019

66. Huang J, Galal G, Etemadi M, Vaidyanathan M. Evaluation and Mitigation of Racial Bias in Clinical Machine Learning Models: Scoping Review. JMIR Med Inform. 2022; 10: e36388. https://doi.org/10.2196/36388 PMID: 35639450

67. Rahmani K, Thapa R, Tsou P, Casie Chetty S, Barnes G, Lam C, et al. Assessing the effects of data drift on the performance of machine learning models used in clinical sepsis prediction. Int J Med Inf. 2023; 173: 104930. https://doi.org/10.1016/j.ijmedinf.2022.104930 PMID: 36893656

68. Fichtenberg CM, Alley DE, Mistry KB. Improving Social Needs Intervention Research: Key Questions for Advancing the Field. Am J Prev Med. 2019; 57: S47–S54. https://doi.org/10.1016/j.amepre.2019.07.018 PMID: 31753279

69. Green K, Zook M. When Talking About Social Determinants, Precision Matters | Health Affairs. In: Health Affairs Blog [Internet]. 29 Oct 2019 [cited 3 Dec 2019]. Available: https://www.healthaffairs.org/do/10.1377/hblog20191025.776011/full/.

70. Alderwick H, Gottlieb LM. Meanings and Misunderstandings: A Social Determinants of Health Lexicon for Health Care Systems. Milbank Q. 2019; 97: 1–13. https://doi.org/10.1111/1468-0009.12390 PMID: 31069864