Impacts of Individual Patient Language and Neighborhood Ethnic Enclave on COVID-19 Test Positivity Among Hispanic/Latinx Patients in San Francisco

Kim Hanh Nguyen, ScD, MPH,*†‡ Anna D. Rubinsky, PhD,* Mindy C. DeRouen, PhD, MPH,*§
Kathryn E. Kemper, MPH,*† Kirsten Bibbins-Domingo, PhD, MD, MAS,*†‡
and Courtney R. Lyles, PhD*†‡

Background: Given the known disparities in COVID-19 within the Hispanic/Latinx community, we sought to examine the interaction between individual-level and neighborhood-level social determinants of health using linked electronic health record data.

Methods: We examined electronic health record data linked to neighborhood data among Hispanic/Latinx patients tested for COVID-19 between March 1, 2020, and February 28, 2021, from 2 large health care systems in San Francisco. Hispanic/Latinx ethnic enclave is measured using an index of census-tract level indicators of ethnicity, nativity, and language. Multilevel logistic regression models examined associations between ethnic enclave and COVID-19 positivity (COVID-19+), adjusting for patient-level sociodemographic and clinical characteristics and health system. Cross-level interactions were used to test whether associations between ethnic enclave and COVID-19+ differed by patient language preference.

Results: Among 26,871 patients, mean age was 37 years, 56% had Spanish-language preference, and 21% were COVID-19+. In unadjusted models, patients living in the highest versus lowest Hispanic/Latinx enclave had 3.2 higher odds of COVID-19+ (95% CI, 2.45–4.24). Adjusted, the relationship between ethnic enclave and COVID-19+ was attenuated, but not eliminated (odds ratio: 1.4; 95% CI, 1.13–1.17). Our results demonstrated a significant cross-level interaction, such that the influence of ethnic enclave was modified by patient language preference. For individuals with Spanish-language preference, risk of COVID-19+ was high regardless of neighborhood context, whereas for those with English preference, neighborhood ethnic enclave more than doubled the odds of infection.

Conclusions: Findings suggest that a multilevel and intersectional approach to the study of COVID-19 inequities may illuminate dimensions of health inequity that affect marginalized communities and offer insights for targeted clinical and community-based interventions.

Key Words: health disparities, Hispanic health, language, Hispanic or Latinx health, population health, social epidemiology

(Med Care 2023;61: 67–74)

From the *Department of Epidemiology and Biostatistics, University of California San Francisco; †UCSF Center for Vulnerable Populations, Zuckerberg San Francisco General Hospital; ‡Division of General Internal Medicine at Zuckerberg San Francisco General Hospital; and §Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA.
Correspondence to: Kim Hanh Nguyen, ScD, MPH, 550 16th Street, 2nd Floor, San Francisco, CA 94143. E-mail: kim.nguyen2@ucsf.edu.

It is well established that COVID-19 disproportionately impacted racial and ethnic communities in the United States, particularly among Hispanic/Latinx populations. Hispanic/Latinx individuals comprised a third of cumulative COVID-19 cases in the early stages of the pandemic despite representing only 18% of the total US population;[1] moreover, they were 1.3–7.7 times more likely to be hospitalized and 2 times more likely to die from COVID-19 compared with their non-Hispanic/Latinx White counterparts.[1,2]

The COVID-19 pandemic highlighted the underlying structural inequities that have disadvantaged racial and ethnic communities across the United States long before the pandemic.[2–4] Investigations of racial and ethnic inequities suggest a variety of individual-level factors associated with increased susceptibility to the virus including comorbidities, fewer socioeconomic resources, poorer health care access, and immigration status (linked to public benefit access), which are more prevalent in Hispanic/Latinx communities.[3,5] At the same time, higher rates of limited English proficiency among the Hispanic/Latinx population may contribute to COVID-19 inequities by way of reduced access to timely and high-quality public health information about disease transmission, symptoms, and testing, as well as lower receipt of adequate COVID-19–related care and treatment.[6–8] In addition to differing susceptibility to COVID-19 across individual-level factors, unequal distribution of COVID-19 cases in predominantly Hispanic/Latinx geographic areas has been documented.[9,10]

Many Hispanic/Latinx individuals live in ethnic enclaves, or ethnically and culturally distinct geographic areas that have a large share of recent immigrants and ethnic entrepreneurship. Hispanic/Latinx neighborhoods may contribute to racial inequities in COVID-19 because they often have more essential workers, household crowding, and concentrated poverty, which may influence residents' exposure to COVID-19.[11–13] On the other hand, living in an ethnic enclave may provide immigrants economic resources and social connections not easily found outside of the immigrant network[14] and may contribute to better health outcomes.[15] Previous studies have analyzed ethnic composition (percentage of Hispanic residents) and residential segregation (degree to which Hispanic residents are geographically separate from White residents) and are conducted at either the zip or county-level, larger aggregate units that may not reflect neighborhood contexts as accurately as census tracts. Ethnic enclave is a composite representation of ethnic neighborhoods, and to our knowledge, no study thus far has studied its impact on COVID-19 inequities.

Although prior studies have explored the contribution of either individual-level or neighborhood-level factors in racial and ethnic COVID-19 inequities, multilevel studies remain scarce.[16,17] There is growing consensus in health inequities research that an intersectional approach is needed to better understand the synergistic interaction between multiple and multilevel social determinants.[18] Given the predominant role of place in both the spread of COVID-19 and public health mitigation strategies, understanding COVID-19 risk among vulnerable communities requires attention to the interaction between people and the places they live.

This study therefore investigated the role of place, specifically Hispanic/Latinx ethnic enclave, in association with COVID-19 positivity (COVID-19+) among Hispanic/Latinx patients living in San Francisco (SF). We hypothesized that among Hispanic/Latinx individuals who are at high risk of COVID-19 infection, neighborhood context matters, and the degree may vary depending on individual characteristics, specifically patient language. We used multilevel analyses to evaluate the association between Hispanic/Latinx enclave status and COVID-19 infection and investigate cross-level interactions between patient language preference and neighborhood ethnic enclave to identify subgroups that may benefit from targeted interventions. Our study used data derived from electronic health records (EHR), which routinely capture comprehensive patient data including health outcomes and residential address, which, when geocoded and linked to small area-level neighborhood data, can be useful in identifying and quantifying place-based inequities.

METHODS

Data Source and Study Population

We conducted an EHR-based cross-sectional study with patients from 2 large integrated health care systems, each serving different subgroups of the SF population. UCSF Health is a university-affiliated health system serving a mix of privately and publicly insured patients, and San Francisco Heath Network (SFHN) is a county-administered health system serving primarily publicly insured (Medicaid) and uninsured patients.[19] The health care and public health infrastructure in SF is likely more advanced with respect to language accessible services than in other municipalities nationwide, but still face common challenges for non–English-speaking residents. These include the need for bilingual staff and robust communication/marketing in multiple languages, as well as specific barriers in reaching communities that may have distrust or hesitancy of seeking services (such as undocumented residents).[20–22] Throughout the pandemic, COVID-19 testing was offered to patients of each health system at multiple locations throughout the city.

Both UCSF and SFHN have EHR systems that include administrative and clinical data from all patient encounters, including laboratory test results. We obtained patient data including demographics, lab test results, International Classification of Diseases (ICD) diagnostic codes, and residential addresses. Patient EHR data including geocoded addresses from both the health care systems were combined into a single analytic database and linked to data on Hispanic ethnic enclave from the California Neighborhoods Data System.[23] The University of California, San Francisco Institutional Review Board approved the study (#17–22277).

Patients were included if they had documented Hispanic/Latinx ethnicity based on self-identification in the EHR, had a COVID-19 polymerase chain reaction (PCR) diagnostic test at 1 of the 2 health care systems between March 1, 2020, and February 28, 2021, and had a residential address in SF documented in the EHR at the time of the test. The time period was selected to ensure capture of cases before widespread vaccination access, which influenced both COVID-19 prevalence and inequities[24]; therefore, vaccine data were not available for analysis at the time of this study. We excluded patients who had missing information on sex (5.4%) and language (0.9%); however, we retained patients with missing insurance status and included a category for "unknown" because they constituted a large portion of insurance type (33%) and may tend to represent uninsured patients.

Patient-level Outcome and Covariates

The study outcome was COVID-19+, defined as having a positive PCR test result for COVID-19 at either site. ICD diagnoses of COVID-19 without a positive test were not included as such codes were often used to order a laboratory test for COVID-19 and did not indicate a confirmed diagnosis.

Patient characteristics obtained from the EHR included preferred language, age, sex, marital status, Charlson comorbidity index (CCI)[25] and insurance at the time of testing. Preferred language is based on patient self-report, frequently used in previous studies at each health system.[26,27] CCI was constructed based on ICD-10-CM codes documented in outpatient clinic visits, inpatient encounters, or the problem list within the past year.

Neighborhood Exposure Variable

Census-tract level Hispanic ethnic enclave is an established index developed through principal components analysis of 7 indicator variables: percentage of residents who

are linguistically isolated (overall and of those who speak Spanish), recent immigrants, Hispanic/Latinx composition, foreign-born composition, and limited English proficiency (overall and of those who speak Spanish).[23] Data were from 2013 to 2017 American Community Survey 5-year estimates with respect to Census 2010 tract boundaries and categorized into quintiles based on the distribution of index values across SF census tracts. Quintile 1 represents neighborhoods with a low level of Hispanic/Latinx enclave (least ethnically distinct) and quintile 5 represents neighborhoods with the highest level of Hispanic/Latinx enclave (most ethnically distinct). Finally, we also examined neighborhood socioeconomic status (nSES), a previously validated composite measure at the census-tract level also derived from ACS data.[23]

Geocoding and Mapping

Individual patient addresses (most recent documented address on or before their COVID-19 test) from the EHR were geocoded to longitude and latitude coordinates. We merged the geocoded EHR dataset with neighborhood data on ethnic enclave using 2010 census-tract identifiers. Census tracts are geographic areas that average 4000 inhabitants and represent relatively homogenous population characteristics and are often used as a proxy for neighborhood. Census-tract level prevalence of COVID-19+ among our study population was calculated and categorized into quintiles. Choropleth maps of COVID-19+ prevalence and Hispanic ethnic enclave levels across census tracts were created and visual and quantitative comparisons were used to identify geographic areas where COVID-19+ prevalence and enclave quintiles were significantly different. We used the Python programming language (Python Software Foundation, https://www. python.org/), and python package GeoPanda v 0.11.0 for data analysis and map visualization.

Statistical Analysis

Descriptive analyses were performed to show distributions of COVID-19+ and patient characteristics across levels of ethnic enclave; statistically significant differences were assessed using χ^2 and t tests. Multilevel modeling was used to estimate the impact of both neighborhood and patient-level factors simultaneously on COVID-19+. Multilevel logistic regression models were fit with binomial distribution and logit link function and included a random intercept for census tract to account for correlated data within census tracts. Extrabinomial variation was explored systematically in all models, and we found no evidence of underdispersion or overdispersion.

Initially, a null (intercept-only) model was fitted to test whether COVID-19+ varied by census tract.[28] The interclass correlation was 0.10 at the census-tract level suggesting sufficient between-neighborhood variability for further multilevel analyses. Then, to estimate the association between neighborhood Hispanic/Latinx enclave and COVID-19+, we added neighborhood-level variable Hispanic/Latinx ethnic enclave (using the lowest level as reference category). In the final model, we added patient-level variables (language preference, sex, age group, marital status, insurance status, CCI, and health system) as fixed effects to estimate the effect of ethnic enclave on COVID-19+ after adjusting for individual clinical and sociodemographic characteristics. We report odds ratios, 95% CIs, and *P*-values for variables of interest. In a separate sensitivity analysis, we ran the same adjusted model additionally controlling for nSES at the census-tract level—to provide context for the ethnic enclave findings alongside other neighborhood characteristics.

To test whether the strength of associations between neighborhood ethnic enclave and COVID-19+ varied as a function of patient language, a cross-level interaction term was added to the final adjusted model.[29] We calculated stratified odds ratios and CIs by both language and ethnic enclave from the adjusted multilevel model with interaction. To better understand the cross-level associations, we plotted the effect of ethnic enclave on COVID-19+ for English and Spanish-language preference using predicted probability estimates and their confidence limits, with the other covariates set to their reference values from the adjusted interaction model.[30,31] We used SAS version 9.3 macro GLIMMIX to fit all descriptive analyses and regression models (SAS Institute Inc., Cary, NC).[32]

RESULTS

Table 1 describes characteristics of the 26,871 Hispanic/Latinx patients who met study inclusion criteria by Hispanic/Latinx ethnic enclave status. The majority (85%) of patients were from SFHN and over half (56%) indicated Spanish as their preferred language. Most patients living in the highest Hispanic/Latinx enclave (Q5) were SFHN versus UCSF patients (91% vs. 9%). The proportion of patients with Spanish-language preference increased from 24% to 65% across increasing quintiles of Hispanic/Latinx enclave status, and the proportion of patients with public insurance increased from 37% to 58% across enclaves. About 21% (n = 5604) of patients tested positive for COVID-19, and the proportion of patients with COVID-19+ increased with increasing quintiles of Hispanic/Latinx ethnic enclave (10%–24%).

Maps of Hispanic/Latinx Enclave and COVID-19+ Prevalence

Figure 1 displays 2 maps of SF, depicting (1) Hispanic ethnic enclave quintiles and (2) prevalence of COVID-19+ for 193 census tracts. Tracts with higher Hispanic/Latinx ethnic enclave status also tend to have higher COVID-19+ prevalence. The overlap between COVID-19+ prevalence and ethnic enclave was primarily concentrated within the southeast quadrant of SF.

Multilevel Logistic Regression Results

In the unadjusted multilevel regression model, we found that compared with individuals living in the lowest level ethnic enclave (Q1), those living in the highest ethnic enclave quintile (Q5) had 3.22 (CI, 2.45–4.24) times higher odds of COVID-19+ (Table 2).

The association between Hispanic/Latinx enclave and COVID-19+ remained statistically significant after adjusting for language preference and other patient sociodemographic and clinical variables (Table 2; $F_{4,188} = 9.09$, $P < 0.0001$). We found a similar pattern of increasing odds across higher levels of ethnic

TABLE 1. Characteristics of Patients Tested for COVID-19 From March 1, 2020, to February 28, 2021, by Hispanic Ethnic Enclave Quintile

Characteristics, No. (%)	Total (N = 26,871)	Quintile 1 (N = 1424)	Hispanic Ethnic Enclave Quintiles				P*
			Quintile 2 (N = 2469)	Quintile 3 (N = 2049)	Quintile 4 (N = 5697)	Quintile 5 (N = 15,232)	
COVID-19 positive	5604 (21)	143 (10)	346 (14)	313 (15)	1132 (20)	3670 (24)	—
Health system	—	—	—	—	—	—	< 0.001
SFHN	22,757 (85)	900 (63)	1707 (69)	1510 (74)	4755 (83)	13,885 (91)	—
UCSF	4114 (15)	524 (37)	762 (31)	539 (26)	942 (17)	1347 (9)	—
Preferred language							< 0.001
English	11,918 (44)	1077 (76)	1643 (67)	1255 (61)	2652 (47)	5291 (35)	—
Spanish	14,953 (56)	347 (24)	826 (33)	794 (39)	3045 (53)	9941 (65)	—
Age group	—	—	—	—	—	—	< 0.001
< 18	4340 (16)	170 (12)	317 (13)	341 (17)	892 (16)	2620 (17)	—
18–34	8092 (30)	422 (30)	847 (34)	604 (29)	1675 (29)	4544 (30)	—
35–64	12,079 (45)	681 (48)	1069 (43)	886 (43)	2562 (45)	6881 (45)	—
65+	2360 (9)	151 (11)	236 (10)	218 (11)	568 (10)	1187 (8)	—
Sex	—	—	—	—	—	—	—
Male	13,851 (52)	745 (52)	1277 (52)	965 (47)	2918 (51)	7946 (52)	—
Female	13,020 (48)	679 (48)	1192 (48)	1084 (53)	2779 (49)	7286 (48)	—
Marital status	—	—	—	—	—	—	0.02
Married/partnered	6461 (24)	377 (26)	585 (24)	493 (24)	1337 (23)	3669 (24)	—
Divorced/separated	983 (4)	60 (4)	89 (4)	78 (4)	239 (4)	517 (3)	—
Single/widowed	17,173 (64)	890 (63)	1614 (65)	1305 (64)	3647 (64)	9717 (64)	—
Other/unknown	2254 (8)	97 (7)	181 (7)	173 (8)	474 (8)	1329 (9)	—
Insurance type	—	—	—	—	—	—	< 0.001
Private	2656 (10)	398 (28)	521 (21)	328 (16)	549 (10)	860 (6)	—
Public	14,652 (55)	521 (37)	1085 (44)	1027 (50)	3218 (56)	8801 (58)	—
Other/unknown	702 (3)	51 (4)	85 (3)	66 (3)	151 (3)	349 (2)	—
Missing	8861 (33)	454 (32)	778 (32)	628 (31)	1779 (31)	5222 (34)	—
Charlson Comorbidity Index	—	—	—	—	—	—	< 0.001
0	22,390 (83)	1166 (82)	2025 (82)	1707 (83)	4658 (82)	12,834 (84)	—
1	2200 (8)	114 (8)	195 (8)	175 (9)	531 (9)	1185 (8)	—
2+	2281 (8)	144 (10)	249 (10)	167 (8)	508 (9)	1213 (8)	—
Neighborhood SES quintile	—	—	—	—	—	—	< 0.001
Q1 (lowest)	9357 (35)	0	287 (12)	226 (11)	1020 (18)	7824 (51)	—
Q2	7963 (30)	0	0	160 (8)	2189 (38)	5614 (37)	—
Q3	3927 (15)	16 (1)	228 (9)	558 (27)	1331 (23)	1794 (12)	—
Q4	3243 (12)	437 (31)	1061 (43)	937 (46)	808 (14)	0	—
Q5 (highest)	2381 (9)	971 (68)	893 (36)	168 (8)	349 (6)	0	—

*P-values compare patient characteristics by COVID-19 status based on χ^2 test.
SES indicates socioeconomic status; SFHN, San Francisco Health Network; UCSF, University of California, San Francisco.

enclave, though the odds ratios were attenuated at all levels. Individuals in the highest level of ethnic enclaves had 1.81 (CI,1.43–2.29) times higher odds of COVID-19+ compared with individuals in the lowest ethnic enclave quintile. In this adjusted model, individuals with Spanish-language preference had 2.17 (CI, 2.01–2.33) times higher odds of COVID-19+ compared with individuals with English-language preference. In our sensitivity analysis examining the impact of nSES on our findings (Appendix Table 1, Supplemental Digital Content 1, http://links.lww.com/MLR/C571), we found that the inclusion of nSES did not alter our study findings and instead potentially strengthened the associations between ethnic enclave and COVID-19+.

In the adjusted model examining the cross-level interaction between neighborhood-level ethnic enclave and patient-level language on COVID-19+ (not shown), the interaction was statistically significant ($F_{4,183} = 8.70$, $P < 0.0001$), suggesting that the relationship between ethnic enclave and COVID-19+

differed by the language preference of the patient. Table 3 displays stratified odds ratios from the multilevel interaction model adjusting for patient-level variables. Stratified estimates by language preference showed statistically significant increasing odds of COVID-19+ across higher neighborhood ethnic enclave quintiles for English-language but not Spanish-language patients. Estimates of the association between language preference and COVID-19+ stratified by ethnic enclave level showed a more complex pattern. Although individuals with Spanish-language preference had greater odds of COVID-19+ than those with English preference at all levels of ethnic enclave, the association was strongest for patients in the lowest ethnic enclave quintile (4.38; CI, 3.00–6.38) and most attenuated for patients in the highest ethnic enclave quintiles [2.00 (CI, 1.73–2.31) for Q4; 1.98 (CI, 1.81–2.17) for Q5].

The cross-level interaction plot (Fig. 2) depicts the pattern of predicted probability of COVID-19+ across ethnic enclave quintiles (neighborhood level) by language preference
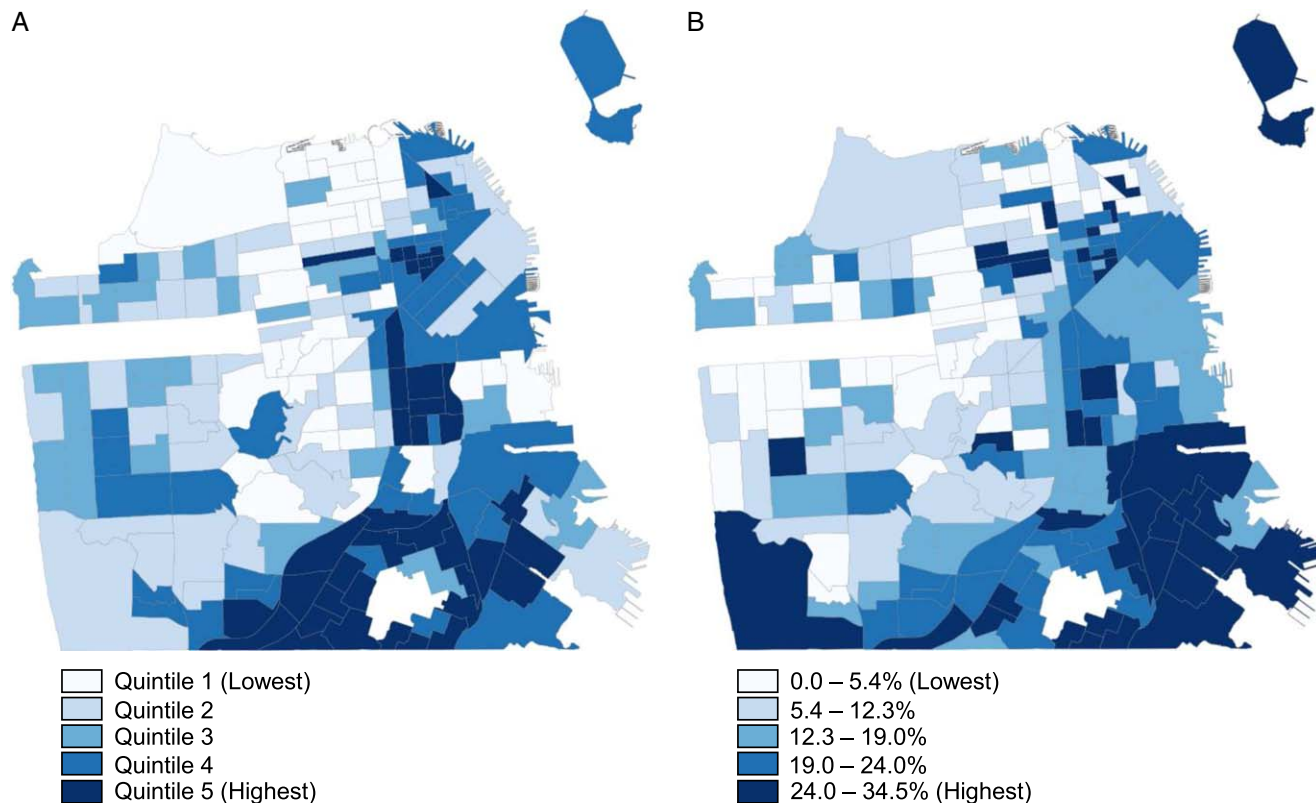
A

B

Quintile 1 (Lowest)
Quintile 2
Quintile 3
Quintile 4
Quintile 5 (Highest)

0.0 – 5.4% (Lowest)
5.4 – 12.3%
12.3 – 19.0%
19.0 – 24.0%
24.0 – 34.5% (Highest)

FIGURE 1. Maps of San Francisco depicting Hispanic ethnic enclave quintile (A) and COVID-19 prevalence (B), by neighborhood. [full color online]

(individual level). The predicted probability is much higher for patients with Spanish-language preference compared with English preference across all Hispanic/Latinx ethnic enclave quintiles. In addition, there are relatively stable probabilities across enclave levels for those with Spanish-language preference (0.16–0.19) but increasing probabilities across enclave levels for those with English preference (0.04–0.11).

DISCUSSION

Our study took a multilevel, intersectional approach in the study of place on COVID-19 inequities. We leveraged EHR data linked to neighborhood data and used multilevel modeling to estimate neighborhood effects on COVID-19+ independent of individual-level characteristics and evaluated the relative contribution of neighborhood effects versus individual factors on COVID-19+. We found that Hispanic/Latinx ethnic enclave was a statistically significant neighborhood influence on COVID-19+. Patients with Spanish-language preference had 2- to over 4-fold greater odds of COVID-19+ compared with those with English-language preference across ethnic enclave status, and with a significant cross-level interaction between language preference and ethnic enclave neighborhood context.

Much of the COVID-19 inequities research thus far has focused on racial residential segregation for African American communities, with overall findings consistent with ours showing a higher risk for adverse COVID-19–related outcomes

with increased racial residential segregation.[5,13,33] The social historical processes that underlie development of Hispanic/Latinx ethnic enclaves and the mechanisms that link ethnic enclave with COVID-19 inequities may differ for Hispanic/Latinx and Black communities, and this needs attention in future research and practice when considering ethnic enclave and racial residential segregation.[34,35]

Although this study aimed to determine the independent association of neighborhood ethnic enclaves on COVID-19 risk across SF, previous community-based work has also demonstrated a need for more attention and explicit evaluation of effective language-relevant and place-based interventions moving forward. For example, future research should investigate further into the role ethnic enclaves may play in structural determinants of health (such as housing, transportation, and service access),[11,12,33,36] as well as social and individual determinants (such as living arrangements and occupation).[37,38]

Another key finding of our study was that ethnic enclaves not only influence overall COVID-19+, but also may have differential effects on certain residents based on individual characteristics—specifically language preference. We observed higher COVID-19+ among patients with Spanish-language versus English-language preference across all ethnic enclave quintiles. In contrast, we observed higher COVID-19+ for patients with English-language preference across increasing ethnic enclave quintiles. It is plausible that Spanish-language preference compared with English-language preference plays a predominant role in COVID-19 risk (eg,

TABLE 2. Adjusted Associations Between Patient and/or Neighborhood Factors and COVID-19 Positivity (Odds Ratios), March 2020 to February 2021[†]

	OR (95% CI)	
Variable	Unadjusted Model	Adjusted Model
Neighborhood factors		
Ethnic enclave quintiles (ref: Q1)		
Q2	1.42** (1.05, 1.91)	1.28* (0.99, 1.67)
Q3	1.49** (1.08, 2.06)	1.27* (0.97, 1.68)
Q4	2.39*** (1.80, 3.16)	1.61*** (1.26, 2.05)
Q5 (highest)	3.22*** (2.45, 4.24)	1.81*** (1.43, 2.29)
Patient characteristics		
Age (ref: 35–64)		
< 18	—	1.19*** (1.09, 1.31)
18–34	—	1.02 (0.95, 1.11)
65+	—	0.67*** (0.59, 0.76)
Sex (ref: male)		
Female	—	0.98 (0.92, 1.04)
Marital Status (ref: single)		
Married/partnered	—	1.31*** (1.21, 1.42)
Divorced/separated	—	0.90 (0.75, 1.09)
Other/unknown	—	0.96 (0.86, 1.08)
Preferred language (ref: English)		
Spanish	—	2.17*** (2.01, 2.33)
Insurance type (ref: Private)		
Public	—	1.08 (0.92, 1.25)
Other/unknown	—	1.22 (0.95, 1.56)
Missing	—	0.98 (0.83, 1.14)
Charlson Comorbidity Index (ref: 0)		
1	—	0.96 (0.85, 1.08)
2+	—	0.92 (0.81, 1.04)
Health system (ref: UCSF)		
SFHN	—	2.31*** (2.00, 2.67)

*$P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.
[†]Results reported in odds ratios obtained from multilevel logistic regression models clustered on census tract.
OR indicates odds ratio; Q, quintile; SFHN, San Francisco Health Network; UCSF, University of California, San Francisco.

inequitable access to Spanish-language COVID-19–related public health information and health services[38]), minimizing the influence of ethnic enclave. Language preference may also reflect other socioeconomic and acculturation factors linked to individual disadvantage which, in turn, play a role in COVID-19 risk.[39] Future studies could decompose dimensions of inequity specific to Hispanic/Latinx populations by language, as well as immigration status or other factors to further elucidate the contribution of neighborhood factors to the widening of these inequities.

Moving forward, it is clear that the relationships between language, ethnic enclave, and COVID-19 are complex, and disentangling multilevel effects is essential to delivering optimal community-based interventions, linking public health and clinical settings.[39,40] Targeted approaches that address underlying structural barriers and social determinants of health (ie, expanded COVID-19 testing in Latinx neighborhoods, linkages across health and social services sectors) and engage community partners are needed.[4,41]

In addition to addressing critical contextual issues of a community, multilevel interventions would also target members with the fewest resources and highest levels of disadvantage who face the greatest COVID-19 risk. For example, health systems and public health departments targeting high ethnic enclave neighborhoods should use public health messaging and communication strategies including written materials, outreach workers, and social media that are culturally and linguistically relevant to and meet the informational requirements of individuals with Spanish-language preference.[21,22,42,43] Our results suggest that a multilevel approach to identify key populations for targeted interventions (such as vaccine outreach or future pandemic response) may increase success to such efforts.[22,44,45]

Strengths and Limitations

Approaches that use multilevel modeling, as well as adding interactions can reveal dimensions of health inequity that may otherwise be obscured using conventional

TABLE 3. Stratified Analyses Showing Odds Ratios and 95% CIs for the Associations Between Patient Language, Neighborhood Ethnic Enclave, and COVID-19 Positivity, Adjusted for Patient-level Factors and Health System[†]

	OR (95% CI)				
Variable	English	Spanish			
Model 1: stratified by language (ref: Ethnic enclave Q1)					
Q2	1.60* (0.97, 2.65)	0.97 (0.59, 1.59)	—	—	—
Q3	1.32 (0.77, 2.28)	1.08 (0.65, 1.79)	—	—	—
Q4	2.34*** (1.47, 3.73)	1.07 (0.68, 1.67)	—	—	—
Q5 (highest)	2.68*** (1.71, 4.20)	1.21 (0.79, 1.87)	—	—	—
Model 2: stratified by ethnic enclave quintiles					
	Q1 OR (95% CI)	Q2 OR (95% CI)	Q3 OR (95% CI)	Q4 OR (95% CI)	Q5 OR (95% CI)
Language (ref: English)					
Spanish	4.38*** (3.00, 6.38)	2.65*** (2.07, 3.39)	3.57*** (2.73, 4.68)	2.00*** (1.73, 2.32)	1.98*** (1.81, 2.17)

*$P < 0.10$.
**$P < 0.05$.
***$P < 0.001$.
[†]Models adjust for all patient-level characteristics.
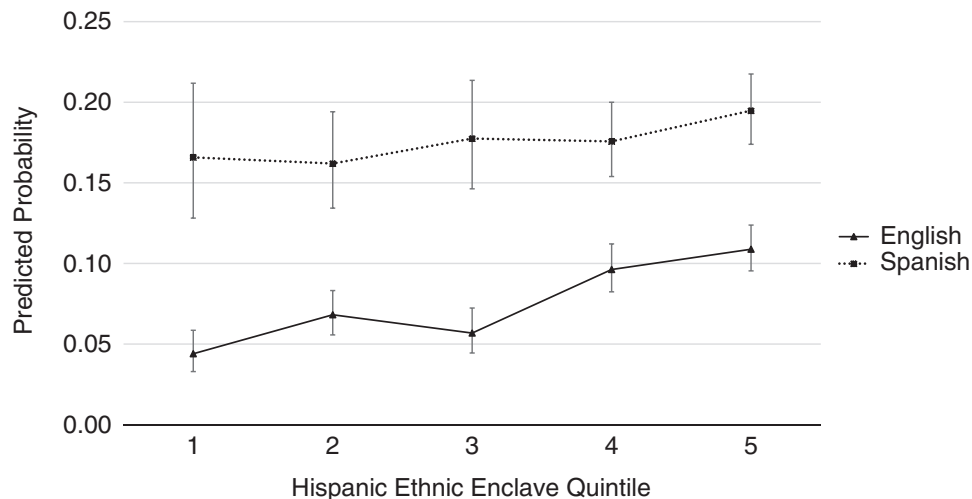OR indicates odds ratio; Q, quintile.

FIGURE 2. Predicted probability for COVID-19 positivity with 95% CIs. Predicted probabilities calculated from a multilevel logistic regression model controlling for patient-level covariates and health system.

methods.[18] Our multilevel modeling approach advances the current ecological literature by evaluating important social risk information of individuals, such as self-identified race/ethnicity, language, or comorbidities, and allowing for greater isolation of neighborhood-level influences from individual-level ones.[46] Our approach can help more accurately identify and inform multilevel interventions that address underlying structural and social factors driving pandemic inequities.

Our analyses controlled for health insurance status, which is a valid proxy for individual SES commonly used in EHR-based population health studies, given the lack of individual-level SES data (such as income) collected within EHRs.[47] We also completed sensitivity analyses adjusted for nSES which confirmed our primary findings. Although it is plausible that omitting different indicators of individual SES such as education and occupation could result in incomplete model bias;[48] our findings for the independent influence of ethnic enclave seem robust.

Whereas the majority of previous studies were conducted at the metropolitan statistical area and county levels, our study reports the impact of residential segregation on COVID-19+ at the census-tract level, which are more granular geographic areas that facilitate precise identification of the most vulnerable areas of the city.[49] To the best of our knowledge, this study is also one of the first attempts to examine social determinants of health associated with increased odds of COVID-19 using EHR data. The use of EHR data in examining relationships between equity-focused predictors, such as race, ethnicity, and language and place-based determinants of health such as ethnic enclave, will help identify priority populations and link clinical decision-making and community-based interventions to accelerate health equity work across fields.[50]

Our study faced several limitations. First, we studied 2 health care systems in SF, which may not be representative of the general SF population, or of Hispanic/Latinx populations in other cities, potentially limiting generalizability. Moreover, ethnic enclave residence was assessed at the index date of COVID-19+ outcome, and we were unable to assess whether

patients moved during the study, which would provide information about neighborhood selection. Finally, we did not control for other neighborhood factors beyond nSES given our specific hypothesis on Hispanic/Latinx ethnic enclaves. Future studies could investigate the joint and interactive effects of ethnic enclave and neighborhood SES.

CONCLUSIONS

The substantial toll exacted by COVID-19 on US minoritized racial and ethnic populations has led health care and public health leaders to call for a refocus on racial and ethnic inequities in health, as well as increased investments in community initiatives and programs to address the social determinants underlying these inequities.[4,6] Barriers to collecting and obtaining sufficient data on individual-level sociodemographic factors may impede public health and health care response to emerging inequities during the pandemic, highlighting the need to strengthen the public health data infrastructure. In addition, the prominent role of geography and place in COVID-19 racial and ethnic inequities compels better understanding of the local environments where minoritized groups live to direct resources and inform community-engaged interventions in these neighborhoods. Our findings demonstrate how important it is for health professionals and policymakers to both understand people within the context of their neighborhoods and consider the intersection of individual and group-level social determinants of health.

REFERENCES

1. Rodriguez-Diaz CE, Guilamo-Ramos V, Mena L, et al. Risk for COVID-19 infection and death among Latinos in the United States: examining heterogeneity in transmission dynamics. *Ann Epidemiol.* 2020;52:46–53 e2.
2. Garcia MA, Homan PA, Garcia C, et al. The color of COVID-19: structural racism and the disproportionate impact of the pandemic on older Black and Latinx adults. *J Gerontol B Psychol Sci Soc Sci.* 2021;76:e75–e80.
3. Webb Hooper M, Napoles AM, Perez-Stable EJ. COVID-19 and racial/ethnic disparities. *JAMA.* 2020;323:2466–2467.

4. Bibbins-Domingo K. This time must be different: disparities during the COVID-19 pandemic. *Ann Intern Med*. 2020;173:233–234.

5. Millett GA, Honermann B, Jones A, et al. White counties stand apart: the primacy of residential segregation in COVID-19 and HIV diagnoses. *AIDS Patient Care STDS*. 2020;34:417–424.

6. Essien UR, Venkataramani A. Data and policy solutions to address racial and ethnic disparities in the COVID-19 pandemic. *JAMA Health Forum*. 2020;1:e200535–e200535.

7. Diamond LC, Jacobs EA, Karliner L. Providing equitable care to patients with limited dominant language proficiency amid the COVID-19 pandemic. *Patient Educ Couns*. 2020;103:1451–1452.

8. Kim HN, Lan KF, Nkyekyer E, et al. Assessment of disparities in COVID-19 testing and infection across language groups in Seattle, Washington. *JAMA Netw Open*. 2020;3:e2021213.

9. Pierce JB, Harrington K, McCabe ME, et al. Racial/ethnic minority and neighborhood disadvantage leads to disproportionate mortality burden and years of potential life lost due to COVID-19 in Chicago, Illinois. *Health Place*. 2021;68:102540.

10. Benitez J, Courtemanche C, Yelowitz A. Racial and ethnic disparities in COVID-19: evidence from six large cities. *J Econ Race Policy*. 2020;3:243–261.

11. Macias Gil R, Marcelin JR, Zuniga-Blanco B, et al. COVID-19 pandemic: disparate health impact on the Hispanic/Latinx population in the United States. *J Infect Dis*. 2020;222:1592–1595.

12. Massey DS, Denton NA. *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press; 1993.

13. Yu Q, Salvador CE, Melani I, et al. Racial residential segregation and economic disparity jointly exacerbate COVID-19 fatality in large American cities. *Ann N Y Acad Sci*. 2021;1494:18–30.

14. Logan JR, Zhang W, Alba RD. Immigrant enclaves and ethnic communities in New York and Los Angeles. *Am Sociol Rev*. 2002;67:299–322.

15. Durazo EM, Mbassa RS, Albert MA. Ethnic enclaves and type II diabetes: a focus on Latino/Hispanic Americans. *Curr Cardiovasc Risk Rep*. 2016;10:36.

16. Van Ingen TA, Brown KA, Buchan SA, et al. Neighbourhood-level socio-demographic characteristics and risk of COVID-19 incidence and mortality in Ontario, Canada: a population-based study. *PLoS One*. 2022;17:e0276507.

17. Borjas GJ. Demographic Determinants of Testing Incidence and COVID-19 Infections in New York City Neighborhoods. IZA Discussion Papers No 131152020.

18. Evans CR. Adding interactions to models of intersectional health inequalities: comparing multilevel and conventional methods. *Soc Sci Med*. 2019;221:95–105.

19. Grumbach K, Knox M, Huang B, et al. A longitudinal study of trends in burnout during primary care transformation. *Ann Fam Med*. 2019;17 (suppl 1):S9–S16.

20. City & County of San Francisco. *Office of Civic Engagement and Immigrant Affairs Language Access in San Francisco: SF Language Access Ordinance Summary Report*: Office of Civic Engagement and Immigrant Affairs; 2021.

21. Khoong EC, Nouri S, Lyles CR, et al. Addressing equity in telemedicine for chronic disease management during the Covid-19 pandemic. *NEJM Catalyst*. 2020.

22. Levins H. A Health Equity Voice From San Francisco's "Latino COVID Pandemic." 2022. Accessed July 25, 2022. https://ldi.upenn.edu/our-work/research-updates/a-health-equity-voice-from-san-franciscos-latino-covid-pandemic/

23. Gomez SL, Glaser SL, McClure LA, et al. The California Neighborhoods Data System: a new resource for examining the impact of neighborhood characteristics on cancer incidence and outcomes in populations. *Cancer Causes Control*. 2011;22:631–647.

24. Njoku A, Joseph M, Felix R. Changing the narrative: structural barriers and racial and ethnic inequities in COVID-19 vaccination. *Int J Environ Res Public Health*. 2021;18:9904.

25. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol*. 1992;45:613–619.

26. Kim NJ, Locke CJ, Park H, et al. Race and hepatitis C care continuum in an underserved birth cohort. *J Gen Intern Med*. 2019;34:2005–2013.

27. Rosenwohl-Mack S, Rubinsky A, Karliner L, et al. Association of patient language with guideline-concordant care for individuals with chronic kidney disease (CKD) in primary care. *J Gen Intern Med*. 2019;34:2348–2350.

28. Snijders T, Bosker R. *Multilevel Analysis An Introduction to Basic and Advanced Multilevel Modeling*, 2nd ed. SAGE Publications Ltd; 1999.

29. Aguinis H, Gottfredson RK, Culpepper SA. Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *J Manag*. 2013;39:1490–1528.

30. Scott Long J, Mustillo SA. Using predictions and marginal effects to compare groups in regression models for binary outcomes. *Sociol Methods Res*. 2018;50:1284–1320.

31. Preacher KJ, Curran PJ, Bauer DJ. Computational Tools for Probing Interactions in Multiple Linear Regression, Multilevel Modeling, and Latent Curve Analysis. *J Educ Behav Stat*. 2006;31:437–448.

32. Kiernan K. *Insights into Using the GLIMMIX Procedure to Model Categorical Outcomes with Random Effects*. SAS Institute; 2018.

33. Yang TC, Emily Choi SW, Sun F. COVID-19 cases in US counties: roles of racial/ethnic density and residential segregation. *Ethn Health*. 2021;26:11–21.

34. Brown TH. Racial stratification, immigration, and health inequality: a life course-intersectional approach. *Social Forces*. 2018;96:1507–1540.

35. Carte GR. Cityscape symposium on the Hispanic housing experience in the United States, part I—Hispanic homelessness, residential segregation, and the neighborhood context of Hispanic housing experiences. *Cityscape*. 2021;23:109–114.

36. Jay J, Bor J, Nsoesie EO, et al. Neighbourhood income and physical distancing during the COVID-19 pandemic in the United States. *Nat Hum Behav*. 2020;4:1294–1302.

37. Liu CY. Ethnic enclave residence, employment, and commuting of Latino workers. *J Policy Anal Manage*. 2009;28:600–625.

38. Olayo-Méndez A, Vidal De Haymes M, García M, et al. Essential, disposable, and excluded: the experience of Latino immigrant workers in the US during COVID-19. *J Poverty*. 2021;25:612–628.

39. Lin Q, Paykin S, Halpern D, et al. Assessment of structural barriers and racial group disparities of COVID-19 mortality with spatial analysis. *JAMA Netw Open*. 2022;5:e220984.

40. Ma KPK, Bacong AM, Kwon SC, et al. The impact of structural inequities on older Asian Americans during COVID-19. *Front Public Health*. 2021;9:690014.

41. Fields J, Gutierrez JR, Marquez C, et al. Community-academic partnerships to address Covid-19 inequities: lessons from the San Francisco Bay area. *NEJM Catalyst*. 2021

42. Esaryk EE, Wesson P, Fields J, et al. Variation in SARS-CoV-2 infection risk and socioeconomic disadvantage among a Mayan-Latinx population in Oakland, California. *JAMA Netw Open*. 2021;4:e2110789.

43. Eliaz A, Blair AH, Chen YH, et al. Evaluating the impact of language concordance on Coronavirus Disease 2019 contact tracing outcomes among Spanish-speaking adults in San Francisco between June and November 2020. *Open Forum Infect Dis*. 2022;9:ofab612.

44. Agurs-Collins T, Persky S, Paskett ED, et al. Designing and assessing multilevel interventions to improve minority health and reduce health disparities. *Am J Public Health*. 2019;109(S1):S86–S93.

45. Chapman LAC, Shukla P, Rodriguez-Barraquer I, et al. Risk factor targeting for vaccine prioritization during the COVID-19 pandemic. *Sci Rep*. 2022;12:3055.

46. Glance LG, Thirukumaran CP, Dick AW. The unequal burden of COVID-19 deaths in counties with high proportions of Black and Hispanic residents. *Med Care*. 2021;59:470–476.

47. Casey JA, Schwartz BS, Stewart WF, et al. Using electronic health records for population health research: a review of methods and applications. *Annual Review of Public Health*. 2016;37:61–81.

48. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socio-economic context and health outcomes: a critical review. *J Epidemiol Community Health*. 2001;55:111–122.

49. Diez Roux AV. Estimating neighborhood health effects: the challenges of causal inference in a complex world. *Soc Sci Med*. 2004;58:1953–1960.

50. Schinasi LH, Auchincloss AH, Forrest CB, et al. Using electronic health record data for environmental and place based population health research: a systematic review. *Ann Epidemiol*. 2018;28:493–502.