Effects of Smoking on SARS-CoV-2 Positivity: A Study of a Large Health System in Northern and Central California

Tobacco Use Insights
Volume 15: 1–10

DOI: 10.1177/1179173X221114799
SAGE

Jiang Li[1], Meghan C Martinez[1], Dominick L Frosch[1] and Georg E Matt[2]

[1]Palo Alto Medical Foundation Research Institute, Center for Health Systems Research, Sutter Health, Palo Alto, CA, USA. [2]College of Sciences, San Diego State University, San Diego, CA, USA.

ABSTRACT

INTRODUCTION: COVID-19 continues to impact vulnerable populations disproportionally. Identifying modifiable risk factors could lead to targeted interventions to reduce infections. The purpose of this study is to identify risk factors for testing positive for SARS-CoV-2.

METHODS: Using electronic health records collected from a large ambulatory care system in northern and central California, the study identified patients who had a test for SARS-CoV-2 between 2/20/2020 and 3/31/2021. The adjusted effect of active and passive smoking and other risk factors on the probability of testing positive for SARS-CoV-2 were estimated using multivariable logistic regression. Analyses were conducted in 2021.

RESULTS: Of 556 690 eligible patients in our sample, 70 564 (12.7%) patients tested positive for SARS-CoV-2. Younger age, being male, racial/ethnic minorities, and having mild major comorbidities were significantly associated with a positive SARS-CoV-2 test. Current smokers (adjusted OR: 0.69, 95% CI: 0.66-0.73) and former smokers (adjusted OR: 0.92, 95% CI: 0.89-0.95) were less likely than nonsmokers to be lab-confirmed positive, but no statistically significant differences were found when comparing passive smokers with non-smokers. The patients with missing smoking status (25.7%) were more likely to be members of vulnerable populations with major comorbidities (adjusted OR ranges from severe: 2.52, 95% CI = 2.36-2.69 to mild: 3.28, 95% CI = 3.09-3.48), lower income (adjusted OR: 0.85, 95% CI: 0.85-0.86), aged 80 years or older (adjusted OR: 1.11, 95% CI: 1.07-1.16), have less access to primary care (adjusted OR: 0.07, 95% CI: 0.07-0.07), and identify as racial ethnic minorities (adjusted OR ranges from Hispanic: 1.61, 95% CI = 1.56-1.65 to Non-Hispanic Black: 2.60, 95% CI = 2.5-2.69).

CONCLUSIONS: Our findings suggest that the odds of testing positive for SARS-CoV-2 were significantly lower in smokers compared to non-smokers. Other risk factors include missing data on smoking status, being under 18, being male, being a racial/ethnic minority, and having mild major comorbidities. Since those with missing data on smoking status were more likely to be members of vulnerable populations with higher smoking rates, the risk of testing positive for SARS-CoV-2 among smokers may have been underestimated due to missing data on smoking status. Future studies should investigate the risk of severe outcomes among active and passive smokers, the role that exposure to tobacco smoke constitutes among nonsmokers, the role of comorbidities in COVID-19 disease course, and health disparities experienced by disadvantaged groups.

KEYWORDS: active smoking, passive smoking, COVID-19, infectious diseases, documentation of smoking status, missing smoking status, racial ethnic disparities

RECEIVED: February 15, 2022. **REVISED:** July 3, 2022. **ACCEPTED:** July 3, 2022.

TYPE: Original Research

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Tobacco-Related Disease Research Program Emergency COVID-19 Funding (10.13039/100005188, R00RG2429).

DATA AVAILABILITY: The data that support the findings of this study are available from the corresponding author, JL, upon reasonable request. The data are not publicly available due to their containing patient health information that could compromise the privacy of research participants. A Data Use Agreement will be implemented for any outside researcher interested in using the data for replication of research findings or for additional areas of research.

SUPPLEMENTAL MATERIAL: Supplemental material for this article is available online.

CORRESPONDING AUTHOR: Jiang Li, Palo Alto Medical Foundation Research Institute, Center for Health Systems Research, Sutter Health, 795 El Camino Real, Ames Building, Palo Alto, CA 94301, USA. Email: Lij14@sutterhealth.org

Introduction

In two and a half years, over 557 million people worldwide have been infected with coronavirus disease 2019 (COVID-19), caused by the SARS-CoV-2 virus (severe acute respiratory syndrome coronavirus 2), and over 6.3 million people have died.[1]

While increasing evidence suggests that the elderly, racial/ethnic minorities, and those with certain comorbidities are at significantly higher risk of adverse outcomes from COVID-19,[2-4] little else is known about what impacts COVID-19 infection. Smoking is a particularly concerning health behavior because it suppresses immune function in the lungs, forces the exhalation of significant quantities and high concentrations of droplet particles, and impacts the amount of contact people may have with contaminated surfaces (from contaminated hands touching cigarettes and cigarettes touching lips) which may increase the risk of lung infection. Additionally, individuals regularly exposed to secondhand smoke (SHS), known as "passive

smokers," may face similar respiratory problems and immune system effects[5-7] as smokers and therefore an increased likelihood of COVID-19 infection. Interestingly, published research to date has shown mixed results on the association of smoking and COVID-19, with some studies suggesting a slight protective effect of smoking on COVID-19 infection.[8-10] However, one of the key challenges for studies on COVID-19 infection is having sufficient sample sizes to allow adjustment for confounding risk factors, such as comorbidities that are closely associated with tobacco smoking.[11] Thus, the need remains for well-designed population-based studies to examine the association between smoking, comorbidities, race/ethnicity, and COVID-19 in terms of the likelihood of infection.

These population-based studies are hindered by another challenge–the incomplete documentation of smoking status in the electronic health record (EHR). Certain patient subgroups such as older adults, racial/ethnic minorities, patients with language barriers, patients with fewer office visits, and light smokers may face a lower likelihood of having smoking history documentation in the EHR.[12] For patients with COVID-19, collecting information about tobacco use and SHS exposure is difficult during any emergency admission and likely to lead to significant reporting errors and potential biases.

In light of these gaps, we sought to explore the differences between laboratory-confirmed COVID-19 positive and negative cases to identify risk factors for testing positive for SARS-CoV-2 and to understand whether smoking (active or passive), comorbidities, or race/ethnicity contributes independently to predisposition of COVID-19 infection, which could guide efforts aimed at minimizing these disparities for COVID-19. We hypothesized that racial and ethnic minority patients who are active or passive smokers and have more severe comorbidities are more likely than other patients to be laboratory-confirmed COVID-19 positive cases. Moreover, we also examined the differences between patients with and without smoking history in the EHR among those who had a SARS-CoV-2 test to understand the generalizability of the findings as well as disparities in documentation of smoking status in the EHR when testing for COVID-19.

Methods

Study Sample

The study setting was Sutter Health, a not-for-profit organization serving over 100 communities in northern and central California, including San Francisco and many of its affluent suburbs, rural communities from Eureka to Modesto, and disadvantaged urban/inner city residents in Oakland and San Jose. In 2018, Sutter Health's foundation-affiliated providers served more than 3 million patients who represent diverse groups of socio-demographic and cultural backgrounds. As of 2020, Sutter patients self-identified their race/ethnicity in the electronic health record (EHR) as: 45.6% non-Hispanic White, 15.6% Hispanic, 16.5% non-Hispanic Asian, 4.7% non-Hispanic Black/African American, and 17.4% other.[13] Sutter Health's electronic health record system (EPIC) is deployed across the enterprise in 3 settings: acute, ambulatory, and community-based practices.

The COVID-19 Universal Registry for Vital Evaluations (CURVE) database, a semi-real-time registry of all confirmed and suspected cases of COVID-19 Sutter Health patients, serves as a centralized resource for this cross-sectional study. Patients were included in the study sample if they had a SARS-CoV-2 test performed at a Sutter Health facility between February 20, 2020 and March 31, 2021. There was no exclusion criterion based on age as COVID-19 can infect individuals of all ages and secondhand smoke exposure may be higher among children and adolescents. A total of 556 690 patients were identified. COVID-19 confirmed cases (N = 70 564) are defined as having one or more positive reverse transcription polymerase chain reaction (RT-PCR) tests for SARS-CoV-2. The comparison group included patients whose COVID-19 lab results were all negative (N = 486 126).

Measures

COVID-19 confirmed cases and the comparison group were linked longitudinally at the patient-encounter level to billing, diagnosis, procedure, clinical encounter records, and provider notes (text fields). Using linked data, we ascertained the prior exposures of subjects in each group, focusing on smoking habits, the severity of comorbidities, and race/ethnicity, and looked into details of care and missing patterns that may be relevant to understanding variation in the COVID-19 lab results. Specifically, detailed information about smoking status, age when the patient first smoked, age when patient stopped smoking, type of tobacco smoked, duration of smoking, and quantity of cigarettes smoked per day were extracted from EHR for both groups. We identified whether the smoking status of each person has been documented in the EHR (missing or non-missing), and if non-missing, we further identified whether s/he is a user of tobacco (current/former/passive/non-smoker) using the most recent self-reported smoking status found in the Social History part of Epic EHR at the time of the earliest lab order for SARS-CoV-2. For example, "Passive smoker" was operationalized using the question "Exposure to secondhand smoke (Yes or No)?" on the adult patient questionnaire and "Does anyone who lives with your child smoke (Yes or No)?" on the children's patient questionnaire. "Pack years of smoking" was calculated by multiplying the number of packs of cigarettes smoked per day by the number of years the person has smoked, and was used to quantify the level of smoking: mild (<10 pack-years); moderate (10-20 pack-years); or heavy (>20 pack- years). Charlson Comorbidity Index (CCI) scores[14] at the time of the earliest lab order for SARS-CoV-2 were calculated to determine the severity of major comorbidities, and patients were divided into four groups: no major comorbidity (CCI = 0); mild (CCI = 1-2); moderate (CCI = 3-4); or severe (CCI ≥ 5). Sutter Health's EHR system also captures race/ethnicity, categorized as

non-Hispanic White, non-Hispanic Black, Hispanic, non-Hispanic Asian, or other. Other covariates include age at the time of the earliest lab order for SARS-CoV-2 and sex. In addition to age, sex, race/ethnicity, and severity of major comorbidities, having received primary care at Sutter Health, defined as having a visit to internal medicine, family medicine, or OB/GYN in the 12 months prior at the time of the earliest lab order for SARS-CoV-2, and median household income, estimated through census data and linked to geocoded home address, were included to examine the documentation of smoking status.

Statistical Analysis

We audited the data for quality and completeness before any analyses were carried out, including missing data patterns. We evaluated distributions to ensure that they met the assumptions of planned analyses and examined the variable distributions to detect outliers. All inferential tests were carried out at a two-tailed alpha level of 0.05. Unadjusted and adjusted Odds Ratios (ORs) were estimated for measuring the effect, and Chi-square tests were used for testing the association between the SARS-CoV-2 positivity and race/ethnicity, smoking status, and severity of major comorbidities among patients who had a SARS-CoV-2 test performed at a Sutter Health facility. Variables were then assessed as potential independent predictors using multivariable logistic regression. Additionally, patients with missing smoking status were compared to those with documented smoking status to understand the generalizability of the findings using Chi-square tests. Logistic regression analysis was used to examine the association of independent variables (i.e., age, sex, race/ethnicity, median household income, having received primary care at Sutter Health, and severity of major comorbidities) with missing smoking status data. All analyses were conducted in 2021 and performed using SAS, version 9.4. This work was reviewed and approved by the Sutter Health IRB and granted a Waiver of Health Insurance Portability and Accountability Act Authorization and a Waiver of Consent as a data-only study.

Results

Sample Characteristics

There were 70 564 (12.7%) patients who tested positive for SARS-CoV-2 out of 556 690 patients who had a test for SARS-CoV-2 between February 20, 2020, and March 31, 2021 (Table 1). The majority were female (57%) and 40 years or older (64.4%). Severity of major comorbidities data were missing in 27.6% of the sample, while 11.9%, 10%, 30.3%, and 20.2% had severe, moderate, mild, and no major comorbidities, respectively. Mean age was 49 years (SD = 21), ranging from 0 to 114 years old. Race and ethnicity data were missing in 6.5% of the sample. Non-Hispanic Whites made up half of the sample, while Hispanics/Latinos, non-Hispanic Blacks, Non-Hispanic Asians, and other ethnicities made up 19.8%, 6.2%, 12.9%, and 4.3% of the sample, respectively.

Of our sample, 5.1% (n = 28 270) were current smokers, 17.3% (n = 96 563) former smokers, and 0.7% (n = 4175) passive smokers. Smoking status data were missing in one quarter of the sample (n = 142 902). Among the lab-confirmed positive patients, 2447 (3.5%) were current smokers, 9106 (12.9%) were former smokers, 643 (0.9%) were passive smokers, and 22 025 (31.2%) did not have their smoking status recorded. Less than 50% of current and former smokers provided detailed information on their smoking habits that enabled us to calculate the pack-years. Twenty-two percent of the smokers were light smokers with fewer than 10 pack-years smoking history, 12.8% of the smokers were moderate smokers with 10-20 pack-years smoking history, while 12% of the smokers were heavy smokers with over 20 pack-years smoking history. The average age of smoking initiation was 22 years (SD = 12), ranging from 7 to 83 years old. The average age of smoking cessation was 39 years (SD = 15), ranging from 12 to 98 years old. Only 3.1% of all patients used smokeless tobacco (0.8% current users and 2.3% former users). Statistically significant differences in years of quitting, level of smoking, and smokeless tobacco use were found when comparing lab-confirmed SARS-CoV-2 positive and negative cases ($P < .0001$).

Risk Factors for Testing Positive for SARS-CoV-2

As shown in Table 2, patients younger than 18 years old were more likely than adults aged 18-39 years old to be lab-confirmed positive (adjusted OR: 1.24, 95% CI: 1.18-1.3) while patients aged 40-59 years old (adjusted OR: 0.93, 95% CI: 0.91-0.96), aged 60-79 years old (adjusted OR: 0.69, 95% CI: 0.67-0.71), or aged 80 years or older (adjusted OR: 0.81, 95% CI: 0.77-0.85) were less likely than adults aged 18-39 years old to be lab-confirmed positive. Males (vs. female: adjusted OR = 1.12, 95% CI = 1.09-1.14), racial ethnic minorities (vs Non-Hispanic White; Hispanic or Latino: adjusted OR = 2.70, 95% CI = 2.64-2.77; Non-Hispanic Black: adjusted OR = 1.31, 95% CI = 1.25-1.36; Other: adjusted OR = 1.90, 95% CI = 1.81-1.99), and patients with mild comorbidities (vs no comorbidities: adjusted OR = 1.10, 95% CI = 1.07-1.13) were more likely to be lab-confirmed positive cases. Patients with missing smoking status (adjusted OR: 1.11, 95% CI: 1.07-1.14) were more likely than non-smokers to be lab-confirmed positive, while current smokers (adjusted OR: 0.69, 95% CI: 0.66-0.73) and former smokers (adjusted OR: 0.92, 95% CI: 0.89-0.95) were less likely than non-smokers to be lab-confirmed positive. There were no significant differences (adjusted OR: 1.05, 95% CI: 0.93-1.17) in passive smokers compared to non-smokers to be lab-

Table 1. Participant characteristics among patients with positive vs negative SARS-CoV-2 tests.

	LAB-CONFIRMED				ALL		*P* VALUE
	NEGATIVE (N = 486 126)		POSITIVE (N = 70 564)		(N = 556 690)		
	N	%	N	%	N	%	
Sex							<.0001
Female	278 868	57.4	38 465	54.5	317 333	57	
Male	207 258	42.6	32 099	45.5	239 357	43	
Age group							<.0001
<18	30 977	6.4	7743	11	38 720	7	
18-39	135 740	27.9	23 741	33.6	159 481	28.6	
40-59	143 234	29.5	21 968	31.1	165 202	29.7	
60-79	140 125	28.8	13 421	19	153 546	27.6	
80+	36 050	7.4	3691	5.2	39 741	7.1	
Race/Ethnicity							<.0001
Missing	30 267	6.2	5684	8.1	35 951	6.5	
Hispanic	85 088	17.5	25 393	36	110 481	19.8	
NH Black	30 789	6.3	3994	5.7	34 783	6.2	
NH Asian	64 667	13.3	6994	9.9	71 661	12.9	
NH White	255 255	52.5	24 407	34.6	279 662	50.2	
Other	20 060	4.1	4092	5.8	24 152	4.3	
Charlson comorbidity index (CCI)							
Missing	127 043	26.1	26 778	37.9	153 821	27.6	
Severe: CCI >= 5	60 101	12.4	5986	8.5	66 087	11.9	
Moderate: CCI = 3-4	50 462	10.4	5313	7.5	55 775	10	
Mild: CCI = 1-2	148 454	30.5	20 150	28.6	168 604	30.3	
No comorbidity: CCI = 0	100 066	20.6	12 337	17.5	112 403	20.2	
Smoking status							<.0001
Missing	120 877	24.9	22 025	31.2	142 902	25.7	
Current smoker	25 823	5.3	2447	3.5	28 270	5.1	
Former smoker	87 457	18	9106	12.9	96 563	17.3	
Passive smoker	3532	0.7	643	0.9	4175	0.7	
Non-smoker	248 437	51.1	36 343	51.5	284 780	51.2	
Age start smoking							0.98156
Missing	110 477	97.5	11 291	97.7	121 768	97.5	
<25 years old	2139	1.9	201	1.7	2340	1.9	
>=25 years old	664	0.6	61	0.5	725	0.6	
Years of quitting smoking							<.0001
Missing	48 602	42.9	5407	46.8	54 009	43.3	
Quit less than 15 years	25 695	22.7	2750	23.8	28 445	22.8	
Quit at leasst 15 years	38 983	34.4	3396	29.4	42 379	33.9	
Level of smoking							<.0001
Missing	59 521	52.5	6838	59.2	66 359	53.2	
>20 pack-years	13 868	12.2	1073	9.3	14 941	12	
10-20 pack-years	14 769	13	1175	10.2	15 944	12.8	
<10 pack-years	25 122	22.2	2467	21.4	27 589	22.1	
Smokeless tobacco use							<.0001
Missing	141 851	29.1	25 374	35.9	167 225	30	
Current user	4020	0.8	581	0.8	4601	0.8	
Former user	11 330	2.3	1283	1.8	12 613	2.3	
Never used	328 925	67.7	43 326	61.4	372 251	66.9	

Note: SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; NH, non-Hispanic; CCI, Charlson Comorbidity Index.

Table 2. Odds ratio for SARS-CoV-2 positivity.

EFFECT	UNCONDITIONAL OR	UNCONDITIONAL 95% CI	MULTIPLE OR	MULTIPLE 95% CI
Age				
<18 vs 18-39 years old	**1.43**	**1.39-1.47**	**1.24**	**1.18-1.3**
Age 40-59 vs 18-39 years old	**0.88**	**0.86-0.89**	**0.93**	**0.91-0.96**
Age 60-79 vs18-39 years old	**0.55**	**0.54-0.56**	**0.69**	**0.67-0.71**
80+ vs 18-39 years old	**0.59**	**0.56-0.61**	**0.81**	**0.77-0.85**
Race/Ethnicity				
Hispanic vs NH White	**3.12**	**3.06-3.18**	**2.70**	**2.64-2.77**
NH Black vs NH White	**1.36**	**1.31-1.41**	**1.31**	**1.25-1.36**
NH Asian vs NH White	**1.13**	**1.1-1.16**	**1.03**	**0.99-1.06**
Other vs NH White	**2.13**	**2.06-2.21**	**1.90**	**1.81-1.99**
Sex				
Male vs Female	**1.12**	**1.11-1.14**	**1.12**	**1.09-1.14**
Smoking status				
Missing vs Non-smoker	**1.25**	**1.22-1.27**	**1.11**	**1.07-1.14**
Current smoker vs Non-smoker	**0.65**	**0.62-0.68**	**0.69**	**0.66-0.73**
Former smoker vs Non-smoker	**0.71**	**0.69-0.73**	**0.92**	**0.89-0.95**
Passive smoker vs Non-smoker	**1.24**	**1.14-1.35**	**1.05**	**0.93-1.17**
Severity of major comorbidities				
Severe vs No comorbidities	**0.81**	**0.78-0.83**	**1.01**	**0.97-1.05**
Moderate vs No comorbidities	**0.85**	**0.83-0.88**	**1.02**	**0.98-1.06**
Mild vs No comorbidities	**1.10**	**1.08-1.13**	**1.10**	**1.07-1.13**

Note: Boldface indicates statistical significance ($P < .05$).
SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; NH, non-Hispanic; CCI, Charlson Comorbidity Index.

confirmed positive after adjustment for age, sex, race/ethnicity, and comorbidities.

Differences Between Patients With and Without Smoking History in the Electronic Health Record Among Patients Who had a SARS-CoV-2 Test

As shown in Table 3, the proportion of lab-confirmed COVID-19 cases was 11.7% and 15.4% for patients with and without documentation of smoking status, respectively ($P < .0001$). The majority (65%) of patients with documented smoking status have received primary care at Sutter Health, while over 90% of patients with missing smoking status have never received primary care at Sutter Health (91%, $P < .0001$). The majority (58.7%) of patients with documentation of smoking status were female, and half (51.9%) of patients without documentation of smoking status were female ($P < .0001$). There are significant differences in age (<18 years old: 6.9% with documentation vs 7.1% without, and 18-39 years old: 26.7% with vs 34.3% without, $P < .0001$) and race ethnicity (Hispanic: 17.7% vs 25.9%, NH Black: 4.6% vs 10.9%, NH Asian: 13.4% vs 11.3%, NH White: 53.6% vs 40.4%, other: 3.7% vs 6.2%, $P < .0001$) when comparing patients with and without documented smoking status. The majority (54.9%) of patients with missing smoking status also had missing data on the major co-morbidities, while only 18.2% of patients with documented smoking status had missing data on the major comorbidities ($P < .0001$).

Risk Factors for Missing Smoking Status in the Electronic Health Record Among Patients Who had a SARS-CoV-2 Test

The odds of having missing smoking status among patients who have received primary care at Sutter Health was only 0.07 times that of those who never received primary care at Sutter Health (adjusted OR: 0.07, 95% CI: 0.07-0.07) (Table 4). Patients younger than 18 years old were nearly one third as likely as patients aged 18-39 years old to have missing smoking status (adjusted OR: 0.35, 95% CI: 0.32-0.37). Similarly, age was associated with a decreased odds of having missing smoking status after age 40 (Age 40-59 vs 18-39 years old: adjusted OR = 0.95, 95% CI = 0.92-0.98 and Age 60-79 vs 18-39 years old: adjusted OR = 0.93, 95% CI = 0.9-0.96) until age 80 or older when the odds of having missing smoking status increased (Age 80 + vs 18-39 years old: adjusted OR = 1.11, 95% CI = 1.07-1.16). Being male (adjusted OR = 1.41, 95% CI = 1.38-1.44), racial ethnic minorities (Hispanic: adjusted OR = 1.61, 95% CI = 1.56-1.65, Non-Hispanic Black: adjusted OR = 2.60, 95%

Table 3. Participant characteristics among patients who received SARS-CoV-2 tests with or without documentation of smoking status in the EHR.

	DOCUMENTATION OF SMOKING STATUS				ALL (N = 556 690)		*P* VALUE
	YES (N = 413 788)		NO (N = 142 902)				
	N	%	N	%	N	%	
Lab-confirmed COVID-19 positive							<.0001
No	365 249	88.3	120 877	84.6	486 126	87.3	
Yes	48 539	11.7	22 025	15.4	70 564	12.7	
Having received primary care at Sutter health							<.0001
No	143 524	34.7	129 987	91	273 511	49.1	
Yes	270 264	65.3	12 915	9	283 179	50.9	
Sex							
Female	243 097	58.7	74 236	51.9	317 333	57	
Male	170 691	41.3	68 666	48.1	239 357	43	
Age group							<.0001
<18	28 642	6.9	10 078	7.1	38 720	7	
18-39	110 455	26.7	49 026	34.3	159 481	28.6	
40-59	127 776	30.9	37 426	26.2	165 202	29.7	
60-79	117 756	28.5	35 790	25	153 546	27.6	
80+	29 159	7	10 582	7.4	39 741	7.1	
Race/Ethnicity							<.0001
Missing	28 507	6.9	7444	5.2	35 951	6.5	
Hispanic	73 413	17.7	37 068	25.9	110 481	19.8	
NH Black	19 172	4.6	15 611	10.9	34 783	6.2	
NH Asian	55 470	13.4	16 191	11.3	71 661	12.9	
NH White	221 984	53.6	57 678	40.4	279 662	50.2	
Other	15 242	3.7	8910	6.2	24 152	4.3	
Charlson comorbidity index (CCI)							<.0001
Missing	75 376	18.2	78 445	54.9	153 821	27.6	
1. Severe: CCI >= 5	51 399	12.4	14 688	10.3	66 087	11.9	
2. Moderate: CCI = 3-4	43 339	10.5	12 436	8.7	55 775	10	
3. Mild: CCI = 1-2	133 256	32.2	35 348	24.7	168 604	30.3	
4. No comorbidity: CCI = 0	110 418	26.7	1985	1.4	112 403	20.2	

Note: SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; NH, non-Hispanic; CCI, Charlson Comorbidity Index.

CI = 2.5-2.69, Non-Hispanic Asian: adjusted OR = 1.68, 95% CI = 1.63-1.74, other: adjusted OR = 1.80, 95% CI = 1.71-1.88), and those with major comorbidities (severe: adjusted OR = 2.52, 95% CI = 2.36-2.69, moderate: adjusted OR = 3.08, 95% CI = 2.89-3.29, mild: adjusted OR = 3.28, 95% CI = 3.09-3.48) were more likely to have missing smoking status than being female, Non-Hispanic White, and those with no major comorbidities. The median household income was associated with lower odds of having missing smoking status (adjusted OR: 0.85, 95% CI: 0.85-0.86) (Table 4).

Discussion

This is one of the first and largest cross-sectional analyses to assess risk factors for testing positive for SARS-CoV-2 in primary care in the United States using EHR data from a large, integrated healthcare system. Moreover, we sought to understand the prevalence and causes of missing smoking data among patients tested for SARS-CoV-2. This study includes patients of all ages who had a test for SARS-CoV-2 which is uncommon among published COVID-19 research to-date. Including patients younger than 18 years old also provided sufficient number of passive smokers to allow assessment of secondhand smoke exposure, a possible COVID-19 risk factor that has not been well studied.

Our findings suggest that current and former smokers are less likely to be lab-confirmed SARS-CoV-2 positive. To date, data on whether smokers have a higher risk of SARS-CoV-2 infection is contradictory and inconclusive.[11] A recent report from a primary care network in the UK showed that active smoking was linked with decreased odds of a positive test result,[15] which is consistent with our findings. Another study on risk of COVID-19 in health-care workers in Denmark showed no significant difference between current, former, and never smokers

Table 4. Odds Ratio for missing smoking status.

EFFECT	UNCONDITIONAL OR	UNCONDITIONAL 95% CI	MULTIPLE OR	MULTIPLE 95% CI
Age				
<18 vs 18-39 years old	**0.79**	**0.77-0.81**	**0.35**	**0.32-0.37**
Age 40-59 vs 18-39 years old	**0.66**	**0.65-0.67**	**0.95**	**0.92-0.98**
Age 60-79 vs 18-39 years old	**0.69**	**0.67-0.7**	**0.93**	**0.9-0.96**
80+ vs 18-39 years old	**0.82**	**0.8-0.84**	**1.11**	**1.07-1.16**
Race/Ethnicity				
Hispanic vs NH White	**1.94**	**1.91-1.97**	**1.61**	**1.56-1.65**
NH Black vs NH White	**3.13**	**3.06-3.21**	**2.60**	**2.5-2.69**
NH Asian vs NH White	**1.12**	**1.1-1.15**	**1.68**	**1.63-1.74**
Other vs NH White	**2.25**	**2.19-2.31**	**1.80**	**1.71-1.88**
Sex				
Male vs Female	**1.32**	**1.3-1.33**	**1.41**	**1.38-1.44**
Income				
Median Household Income (1 unit= $10k)	**0.81**	**0.81-0.81**	**0.85**	**0.85-0.86**
Access to healthcare				
Received primary care at Sutter Health vs not	**0.05**	**0.05-0.05**	**0.07**	**0.07-0.07**
Severity of major comorbidities				
Severe vs No comorbidities	**15.90**	**15.15-16.68**	**2.52**	**2.36-2.69**
Moderate vs No comorbidities	**15.96**	**15.2-16.76**	**3.08**	**2.89-3.29**
Mild vs No comorbidities	**14.76**	**14.09-15.45**	**3.28**	**3.09-3.48**

Note: Boldface indicates statistical significance ($P < .05$).
NH, non-Hispanic.

when comparing prevalence of antibodies against SARS-CoV-2.[16] There is no significant association between passive smokers and being lab-confirmed SARS-CoV-2 positive after adjustment for age, sex, race/ethnicity, and comorbidities. The possible reasons for the lower odds of being lab-confirmed SARS-CoV-2 positive among active smokers but not passive smokers remain unclear. However, Kashyap et al[17] hypothesized that a current smoker's immune systems is less responsive to a COVID-19 infection than that of a never smoker whose immune system would rapidly trigger a cytokine release syndrome. Moreover, when smokers exhale, they are intentionally and forcefully pushing air out of their lungs. The deliberate, deep exhalation of tobacco smoke expels large quantities and high concentrations of particles. The massive rush of particulate matter (PM) may provide opportunities for viral particle to be expelled from the lung and transmitted to passive smokers. Exhaled smoke PM is likely to travel farther than regular exhaling. Consequently, secondhand smoke deposits on surfaces and in dust to become thirdhand smoke residue. Viral particles and droplets with viral particles also deposit on surfaces and in dust. Third-hand smoke constituents include compounds with known microbial activity. Thirdhand smoke may affect the survival of COVID-19 on surfaces and in dust.[18]

An important finding is that missing smoking status was associated with an increased odds of being lab-confirmed SARS-CoV-2 positive after adjustment for age, sex, race/ethnicity, and comorbidities. In our sample, the patients with missing smoking status were more likely to be male, aged 80 years or older, a racial ethnic minority, have lower income, less access to primary care, and have major comorbidities. Smokers are at high risk of having or developing other chronic diseases,[19] and smoking is more prevalent among economically disadvantaged groups and certain racial/ethnic minority groups.[20] This suggest that those with missing smoking status (25.7% of the sample) were more likely to be smokers. Therefore, the overall SARS-CoV-2 positivity rate of smokers in the entire sample may have been underestimated due to missing smoking status data. Additionally, selection bias may skew the results because smokers are more likely to experience respiratory symptoms including cough, expectoration, and sore throat, which could lead to more frequent testing and increasing proportion of smokers with negative SARS-CoV-2 results.[15]

Approximately 17% of U.S. adults currently smoke cigarettes[21] and, while shelter-in-place restrictions have been put in place to slow the spread of COVID-19, these restrictions may also increase feelings of social isolation and mental distress, both of which have been linked to increased motivation to smoke.[22]

As smoking is ultimately a modifiable behavior, understanding the impact of active and passive smoking on COVID-19 infection is especially important during the COVID-19 pandemic. Evidence suggests that individuals aware of the emerging infectious diseases are more likely to practice preventive behavior,[23-25] providing a potential opportunity to decrease smoking rates and protect non-smokers from tobacco smoke exposure.

Previous studies suggested that COVID-19 has affected more adults than children, and more men than women.[26,27] With the more recent sample, we found an increased risk of a positive SARS-CoV-2 test in patients under 18 years old and in men. The number of children contracting COVID-19 in the U.S. made up only around 3% of the U.S. total in 2020, but children now account for more than a fifth of new coronavirus cases in states that release data by age, according to the American Academy of Pediatrics.[28] The steadily increasing child COVID-19 cases have raised alarms about a surge in COVID infection among children. By March 31, 2021 (i.e., end of the study period), children were the remaining population ineligible for the vaccine and the main vectors of virus spread creating risk to both themselves and the rest of the population. Full approval of the coronavirus vaccine for children 5-11 year-old was released on October 29, 2021,[29] which is essential to help protect children from COVID-19 infection.

We also found that Hispanic/Latino patients had 2.70 times the odds of a positive SARS-CoV-2 test result than non-Hispanic White patients, which remained significant after adjusting for smoking status, age, sex, and comorbidities. Consistent with previous studies,[30,31] this finding highlights the disproportionate incidence of COVID-19 among the Hispanic population and the burden of the pandemic on racial and ethnic minority communities across the country.[32,33] Employment in high-risk positions, education, income, and structural barriers to healthcare are all factors likely contributing to this association.[34,35] According to data from Johns Hopkins and the American Community Survey, the COVID-19 infection rate is 3-fold higher in counties with a majority Black population vs predominately White counties, and the death rate is 6-fold higher in Black counties.[36] The higher proportion of positive SARS-CoV-2 tests in racial and ethnic minority groups highlights differential access to resources and underlying systemic racism which are contributing to and exacerbating the racial ethnic disparities seen in the COVID-19 pandemic.

Increasing evidence shows that a number of chronic comorbidities such as hypertension, diabetes, and cardiovascular disease are associated with an increased risk of progressing to severe COVID-19 disease.[2-4] Comorbidities that coexist with COVID-19 may lead to a delayed diagnosis, be confounders in analysis of association between COVID-19 and other risk factors, and increase morbidity and mortality. Therefore, it appears desirable to summarize one or multiple comorbidities into a single score in an efficient manner, using comorbidity indices from the EHR data. The CCI is the most commonly used comorbidity indicx and includes sixteen diseases with different weights based on the strength of their association with mortality. Our study is the first to adapt the CCI to a large health care database to examine the association between comorbidities and COVID-19. We found an association between severity of major comorbidities and a positive SARS-CoV-2 test as hypothesized. Specifically, patients with mild comorbidities were about 1.10 times as likely to have the positive SARS-CoV-2 test result as those without comorbidities. It may add prognostic information to the COVID-19 diagnosis as a predictor of SARS-CoV-2 positivity.

Over a quarter of patients tested for SARS-CoV-2 have no documentation of smoking status in the EHR. The influence of the missing data on the association between smoking and SARS-CoV-2 positivity remains unknown. Healthcare systems should enhance documentation of smoking status among all patients and especially among COVID-19 patients. Vulnerable populations with major comorbidities, lower income, less access to primary care, aged 80 years or older, and identifying as racial ethnic minorities were more likely to have missing smoking data, suggesting opportunities to target subgroups and influence clinical practice at patient, provider, and systems levels.

Limitations

We acknowledge several limitations in the current study. The study population only included individuals who were assessed for SARS-CoV-2 infection by a physician and were tested for SARS-CoV-2. Thus, results only reflect risk factors for testing positive for SARS-CoV-2. There are limitations in the diagnostic tests used and the sensitivity and specificity of the tests themselves. For example, test sensitivity increases with viral load, so the SARS-CoV-2 positive patients may have been more likely to include patients who were more severely ill. On the other hand, a false negative test result indicating that a person has not been infected when the person actually does have the infection poses a challenge to COVID-19 diagnosis.[37] However, it may result in an underestimate of the association between COVID-19 and risk factors. In our sample, many patients had several simultaneous or repeated tests (up to 20 tests for SARS-CoV-2) which could overcome an individual test's limited sensitivity.

Although barriers to documentation of smoking status for patients tested for SARS-CoV-2 in vulnerable populations need to be better understood, the completeness, sensitivity, and positive predictive values of smoking history documentation in the EHR have improved considerably in recent years, indicating that the EHR is an acceptable source for identifying persons with a history of smoking for research and clinical purposes similar to our study.[12] However, the measurement of some smoking behaviors such as duration of smoking, age at which patient first smoked, type of tobacco smoked, level of smoking, and passive smoking, are limited by the available data and may

be underreported in the EHR.[38,39] More comprehensive assessments of secondhand smoke exposure and biological confirmation are required to better understand potential biases in reported exposure and the conditions under which secondhand smoke may affect COVID-19 transmission, infection, symptoms, and health outcomes. Acknowledging this limitation, similar studies like ours on the health effects of passive smoking among COVID-19 patients may become another reason for clinicians to be more committed to improved passive smoking documentation and patient education.

Lastly, we are studying a single healthcare organization in northern and central California. However, studying a single system with a shared infrastructure controls for access barriers and variations in testing supplies and protocols, and allows us to instead focus on variation across patients who had a test for SARS-CoV-2 and factors that could lead to COVID-19 prevention and control.

Conclusions

Our findings suggest that children, men, patients identifying as racial ethnic minorities, patients without documentation of smoking status, and patients with mild major comorbidities are at increased risk for testing positive for SARS-CoV-2. Federal, state, and local government should dedicate the resources necessary to support these particularly vulnerable populations during COVID-19 pandemic comprehensively. This should include, but not be limited to, protecting nonsmokers from exposure to secondhand smoke. Findings from this study would be useful in identifying the critical individual characteristics that should be considered in tailoring approaches to identify trusted sources of information and create effective messaging to disseminate COVID-19 prevention and control information. Additional research is also needed to improve documentation of smoking history, pack-years, vaping, and secondhand exposure especially among patients tested for SARS-CoV-2. Future studies should investigate the risk of severe outcomes among active and passive smokers, the role that exposure to tobacco smoke constitutes among nonsmokers, the role of comorbidities in COVID-19 disease course, and health disparities experienced by disadvantaged groups.

Acknowledgments

The authors would like to acknowledge Rene Everline, Senior Information Analyst, and Amandeep Mann, Statistical Analyst, at the Palo Alto Medical Foundation Research Institute, Center for Health Systems Research, Sutter Health, for their data extraction work.

ORCID iDs

Jiang Li https://orcid.org/0000-0003-4183-1006
Georg E Matt https://orcid.org/0000-0001-5604-4609

REFERENCES

1. Johns Hopkins University. *Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at* Johns Hopkins University (JHU). Johns Hopkins Coronavirus Resource Center. https://coronavirus.jhu.edu/map.html. Accessed July 12, 2022.
2. Liu W, Tao ZW, Wang L, et al. Analysis of factors associated with disease outcomes in hospitalized patients with 2019 novel coronavirus disease. *Chin Med J (Engl)*. 2020;133(9):1032-1038. doi:10.1097/CM9.0000000000000775.
3. Boundy E, Bowen V, Chow N, et al. CDC COVID-19 Response Team. Severe outcomes among patients with coronavirus disease 2019 (COVID-19)—United States, February 12–March 16, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(12):343-346. doi:10.15585/mmwr.mm6912e2.
4. Guan WJ, Liang WH, Zhao Y, et al. Comorbidity and its impact on 1590 patients with COVID-19 in China: A nationwide analysis. *Eur Respir J*. 2020;55(5):2000547. doi:10.1183/13993003.00547-2020.
5. Baskaran V, Murray RL, Hunter A, Lim WS, McKeever TM. Effect of tobacco smoking on the risk of developing community acquired pneumonia: A systematic review and meta-analysis. *PloS One*. 2019;14(7):e0220204. doi:10.1371/journal.pone.0220204.
6. Arbex MA, Santos UdP, Martins LC, et al. Air pollution and the respiratory system. *J Bras Pneumol Publicacao Soc Bras Pneumol E Tisilogia*. 2012;38(5):643-655. doi:10.1590/s1806-37132012000500015.
7. Jaakkola MS, Jaakkola JJK. Effects of environmental tobacco smoke on the respiratory health of adults. *Scand J Work Environ Health*. 2002;28(suppl 2):52-70. http://www.jstor.org/stable/40967255. Accessed July 12, 2022.
8. Patanavanich R, Glantz SA. Smoking is associated with COVID-19 progression: A meta-analysis. *Nicotine Tob Res*. 2020;22(9):1653-1656. doi:10.1093/ntr/ntaa082.
9. Vardavas CI, Nikitara K. COVID-19 and smoking: A systematic review of the evidence. *Tob Induc Dis*. 2020;18:20. doi:10.18332/tid/119324.
10. Farsalinos K, Barbouni A, Niaura R. Systematic review of the prevalence of current smoking among hospitalized COVID-19 patients in China: Could nicotine be a therapeutic option? *Intern Emerg Med*. 2020;15:845-852. doi:10.1007/s11739-020-02355-7.
11. Shastri MD, Shukla SD, Chong WC, et al. Smoking and COVID-19: What we know so far. *Respir Med*. 2021;176:106237. doi:10.1016/j.rmed.2020.106237.
12. Chen LH, Quinn V, Xu L, et al. The accuracy and trends of smoking history documentation in electronic medical records in a large managed care organization. *Subst Use Misuse*. 2013;48(9):731-742. doi:10.3109/10826084.2013.787095.
13. Azar KMJ, Lockhart SH, Shen Z, et al. Persistence of disparities among racially/ethnically marginalized groups in the coronavirus disease 2019 pandemic regardless of statewide shelter-in-place policies: An analysis from Northern California. *Am J Epidemiol*. 2021;190(11):2300-2313. doi:10.1093/aje/kwab191.
14. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis*. 1987;40(5):373-383. doi:10.1016/0021-9681(87)90171-8.
15. de Lusignan S, Dorward J, Correa A, et al. Risk factors for SARS-CoV-2 among patients in the Oxford royal college of general practitioners research and surveillance centre primary care network: A cross-sectional study. *Lancet Infect Dis*. 2020;20(9):1034-1042. doi:10.1016/S1473-3099(20)30371-6.
16. Iversen K, Bundgaard H, Hasselbalch RB, et al. Risk of COVID-19 in health-care workers in Denmark: An observational cohort study. *Lancet Infect Dis*. 2020;20(12):1401-1408. doi:10.1016/S1473-3099(20)30589-2.
17. Kashyap VK, Dhasmana A, Massey A, et al. Smoking and COVID-19: Adding fuel to the flame. *Int J Mol Sci*. 2020;21(18):6581. doi:10.3390/ijms21186581.
18. Mahabee-Gittens EM, Merianos AL, Matt GE. Letter to the editor regarding: "An imperative need for research on the role of environmental factors in transmission of novel coronavirus (COVID-19)"—Secondhand and thirdhand smoke as potential sources of COVID-19. *Environ Sci Technol*. 2020;54(9):5309-5310. doi:10.1021/acs.est.0c02041.
19. U.S. Department of Health and Human Services. *The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014. http://www.ncbi.nlm.nih.gov/books/NBK179276/. Accessed July 12, 2022.
20. Drope J, Liber AC, Cahn Z, et al. Who's still smoking? Disparities in adult cigarette smoking prevalence in the United States. *CA Cancer J Clin*. 2018;68(2):106-115. doi:10.3322/caac.21444.
21. Hu SS, Homa DM, Wang T, et al. State-specific patterns of cigarette smoking, smokeless tobacco use, and e-cigarette use among adults—United States, 2016. *Prev Chronic Dis*. 2019;16:E17. doi:10.5888/pcd16.180362.
22. Roberts ME, Doogan NJ, Kurti AN, et al. Rural tobacco use across the United States: How rural and urban areas differ, broken down by census regions and divisions. *Health Place*. 2016;39:153-159. doi:10.1016/j.healthplace.2016.04.001.
23. Alqahtani AS, Wiley KE, Mushta SM, et al. Association between Australian Hajj Pilgrims' awareness of MERS-CoV, and their compliance with preventive measures and exposure to camels. *J Travel Med*. 2016;23(5):taw046. doi:10.1093/jtm/taw046.

24. Choi JS, Kim JS. Factors influencing preventive behavior against middle east respiratory syndrome-coronavirus among nursing students in South Korea. *Nurse Educ Today*. 2016;40:168-172. doi:10.1016/j.nedt.2016.03.006.
25. Kim JS, Choi JS. Middle east respiratory syndrome-related knowledge, preventive behaviours and risk perception among nursing students during outbreak. *J Clin Nurs*. 2016;25(17-18):2542-2549. doi:10.1111/jocn.13295.
26. Baker P, White A, Morgan R. Men's health: COVID-19 pandemic highlights need for overdue policy action. *Lancet Lond Engl*. 2020;395(10241):1886-1888. doi:10.1016/S0140-6736(20)31303-9.
27. Betron M, Gottert A, Pulerwitz J, Shattuck D, Stevanovic-Fenn N. Men and COVID-19: Adding a gender lens. *Glob Public Health*. 2020;15(7):1090-1092. doi:10.1080/17441692.2020.1769702.
28. American Academy of Pediatrics. *Children and COVID-19: State-Level Data Report*. http://www.aap.org/en/pages/2019-novel-coronavirus-covid-19-infections/children-and-covid-19-state-level-data-report/. Accessed February 4, 2022.
29. FDA. *FDA Authorizes Pfizer-BioNTech COVID-19 Vaccine for Emergency Use in Children 5 through 11 Years of Age*. U.S. Food and Drug Administration. https://www.fda.gov/news-events/press-announcements/fda-authorizes-pfizer-biontech-covid-19-vaccine-emergency-use-children-5-through-11-years-age. Published October 29, 2021. Accessed February 4, 2022.
30. Stokes EK, Zambrano LD, Anderson KN, et al. Coronavirus disease 2019 case surveillance-United States, January 22-May 30, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(24):759-765. doi:10.15585/mmwr.mm6924e2.
31. Moore JT, Ricaldi JN, Rose CE, et al. Disparities in incidence of COVID-19 among underrepresented racial/ethnic groups in counties identified as hotspots during June 5–18, 2020 — 22 States, February–June 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(33):1122-1126. doi:10.15585/mmwr.mm6933e1.
32. Bowleg L. We're not all in this together: On COVID-19, intersectionality, and structural inequality. *Am J Public Health*. 2020;110(7):917-917. doi:10.2105/AJPH.2020.305766.
33. Wilder JM. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. *Clin Infect Dis*. 2021;72(4):703-706. doi:10.1093/cid/ciaa959.
34. Hooper MW, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. *JAMA*. 2020;323(24):2466-2467. doi:10.1001/jama.2020.8598.
35. Jones J, Sullivan PS, Sanchez TH, et al. Similarities and differences in COVID-19 awareness, concern, and symptoms by race and ethnicity in the United States: Cross-sectional survey. *J Med Internet Res*. 2020;22(7):e20001. doi:10.2196/20001.
36. Yancy CW. COVID-19 and African Americans. *JAMA*. 2020;323(19):1891-1892. doi:10.1001/jama.2020.6548.
37. Woloshin S, Patel N, Kesselheim AS. False negative tests for SARS-CoV-2 infection—Challenges and implications. *N Engl J Med*. 2020;383(6):e38. doi:10.1056/NEJMp2015897.
38. Mahabee-Gittens EM, Merianos AL, Gordon JS, et al. Electronic health record classification of tobacco smoke exposure and cotinine levels in hospitalized pediatric patients. *Hosp Pediatr*. 2019;9(9):659-664. doi:10.1542/hpeds.2018-0247.
39. Middleton C, Bruns D. Improving screening and education for secondhand smoke exposure in primary care settings. *AJN Am J Nurs*. 2019;119(8):51-58. doi:10.1097/01.NAJ.0000577456.92598.a5.