Check for updates

RESEARCH ARTICLE

REVISED Association of smoking status with hospitalisation for COVID-19 compared with other respiratory viruses a year previous: a case-control study at a single UK National Health Service trust [version 3; peer review: 2 approved]

David Simons[1-3], Olga Perski[4], Lion Shahab[4], Jamie Brown[4], Robin Bailey[2,3]

[1]Centre for Emerging, Endemic and Exotic Diseases, Royal Veterinary College, London, London, UK
[2]Hospital for Tropical Diseases, University College London Hospitals NHS Trust, London, UK
[3]London School of Hygiene and Tropical Medicine, London, UK
[4]Department of Behavioural Science and Health, University College London, London, UK

v3 **First published:** 24 Aug 2021, **10**:846
https://doi.org/10.12688/f1000research.55502.1
Second version: 15 Nov 2021, **10**:846
https://doi.org/10.12688/f1000research.55502.2
Latest published: 17 Jan 2022, **10**:846
https://doi.org/10.12688/f1000research.55502.3

Abstract

Background: It is unclear whether smoking increases the risk of COVID-19 hospitalisation. We first examined the association of smoking status with hospitalisation for COVID-19 compared with hospitalisation for other respiratory viral infections a year previous. Second, we examined the concordance between smoking status recorded on the electronic health record (EHR) and the contemporaneous medical notes.

Methods: This case-control study enrolled adult patients (446 cases and 211 controls) at a single National Health Service trust in London, UK. The outcome variable was type of hospitalisation (COVID-19 vs. another respiratory virus a year previous). The exposure variable was smoking status (never/former/current smoker). Logistic regression analyses adjusted for age, sex, socioeconomic position and comorbidities were performed. The study protocol and analyses were pre-registered in April 2020 on the Open Science Framework.

Results: Current smokers had lower odds of being hospitalised with COVID-19 compared with other respiratory viruses a year previous (OR_{adj}=0.55, 95% CI=0.31-0.96, p=.04). There was no significant association among former smokers (OR_{adj}=1.08, 95% CI=0.72-1.65, p =.70). Smoking status recorded on the EHR (compared with the contemporaneous medical notes) was incorrectly recorded for 168 (79.6%) controls ($\chi^2(3)$=256.5, p=<0.001) and 60 cases (13.5%) ($\chi^2(3)$=34.2, p=<0.001).

Conclusions: In a single UK hospital trust, current smokers had reduced odds of being hospitalised with COVID-19 compared with

Open Peer Review

Reviewer Status ✔ ✔

	Invited Reviewers	
	1	**2**
version 3 (revision) 17 Jan 2022		✔ report
version 2 (revision) 15 Nov 2021	✔ report	? report
version 1 24 Aug 2021	✘ report	✘ report

1. **Maria Rosaria Galanti**, Karolinska Institutet, Stockholm, Sweden

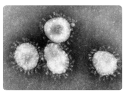
2. **Jonathan M. Samet**, Colorado School of Public Health, Aurora, USA

Any reports and responses or comments on the article can be found at the end of the article.

other respiratory viruses a year previous, although it is unclear whether this association is causal. Targeted post-discharge recording of smoking status may account for the greater EHR-medical notes concordance observed in cases compared with controls.

Keywords
tobacco, smoking, respiratory infections, COVID-19, case-control study, hospitalisation

This article is included in the Emerging Diseases and Outbreaks gateway.



This article is included in the Coronavirus collection.

Corresponding author: Olga Perski (olga.perski@ucl.ac.uk)

Author roles: Simons D: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Perski O**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Shahab L**: Conceptualization, Methodology, Writing – Review & Editing; **Brown J**: Conceptualization, Methodology, Writing – Review & Editing; **Bailey R**: Data Curation, Supervision, Writing – Review & Editing

Competing interests: DS, OP and RB have no conflicts of interest to declare. LS has received a research grant and honoraria for a talk and travel expenses from manufacturers of smoking cessation medications (Pfizer and Johnson & Johnson). JB has received unrestricted research funding to study smoking cessation from companies who manufacture smoking cessation medications. All authors declare no financial links with tobacco companies or e-cigarette manufacturers or their representatives.

Grant information: DS is supported by a PhD studentship from the UK Biotechnology and Biological Sciences Research Council [BB/M009513/1]. OP receives salary support from Cancer Research UK [C1417/A22962]. JB, LS, & OP are members of SPECTRUM, a UK Prevention Research Partnership Consortium [MR/S037519/1]. UKPRP is an initiative funded by the UK Research and Innovation Councils, the Department of Health and Social Care (England) and the UK devolved administrations, and leading health research charities.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2022 Simons D *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Simons D, Perski O, Shahab L *et al.* **Association of smoking status with hospitalisation for COVID-19 compared with other respiratory viruses a year previous: a case-control study at a single UK National Health Service trust [version 3; peer review: 2 approved]** F1000Research 2022, **10**:846 https://doi.org/10.12688/f1000research.55502.3

First published: 24 Aug 2021, **10**:846 https://doi.org/10.12688/f1000research.55502.1

REVISED **Amendments from Version 2**

This version of the article contains a clarification about exposure classification in instances where discrepancies were observed between the EHR and the contemporaneous medical notes (Methods section), in addition to a reference to the US Surgeon General's 2014 report (Discussion section).

Any further responses from the reviewers can be found at the end of the article

Introduction

COVID-19 is a respiratory disease caused by the SARS-CoV-2 virus. There are in excess of 118 million confirmed COVID-19 cases globally, with over 2.6 million deaths reported (Johns Hopkins Coronavirus Resource Center, 2021). Large age and sex differences in case severity and mortality have been observed (Guan *et al*., 2020), with hypertension, diabetes and obesity identified as important risk factors (Fang *et al*., 2020). There are *a priori* reasons to believe that current smokers are at increased risk of contracting COVID-19 and experiencing greater disease severity once infected. SARS-CoV-2 enters epithelial cells through the ACE-2 receptor (Hoffmann *et al*., 2020). Evidence suggests that gene expression and subsequent ACE-2 receptor levels are elevated in the airway and oral epithelium of current smokers (Brake *et al*., 2020; Cai, 2020), potentially making smokers vulnerable to contracting SARS-CoV-2. Other studies, however, show that smoking downregulates the ACE-2 receptor (Oakes *et al*., 2018). In addition, smoking involves repeated hand-to-mouth movements, which may mean that smokers are more likely to contract respiratory viruses such as SARS-CoV-2 (Simons *et al*., 2020a). Early data from the ongoing pandemic have not provided clear evidence for an association of smoking status with COVID-19 outcomes, with a living review and unadjusted Bayesian meta-analysis of over 60 studies indicating that current smokers, compared with those who had never smoked, may be at reduced risk of SARS-CoV-2 infection, while former smokers are at increased risk of hospitalisation, disease severity and in-hospital mortality compared with those who had never smoked (Simons *et al*., 2021a).

Most studies to date have been limited by the lack of appropriate controls, poor recording of smoking status and insufficient adjustment for relevant covariates. Many studies relied on routine electronic health records (EHRs) to obtain data on demographic characteristics, comorbidities and smoking status. This is problematic, as previous research suggests that data on smoking status obtained via EHRs tend to be incomplete or inaccurate, with implausible longitudinal changes observed (Polubriaginof *et al*., 2018). As hospitalised populations differ by age and sex from the general population (Secondary Care Analytical Team, 2020), comparisons of current and former smoking prevalence in hospitalised and non-hospitalised populations are likely biased. There is therefore a need for alternative study designs with relevant comparator groups and adjustment for covariates to better understand the association of smoking status with COVID-19 disease outcomes.

However, the selection of an appropriate comparator group is not straightforward. Ideally, controls should represent the underlying population from which cases emerged, both geographically and demographically (Grimes and Schulz, 2005). In the context of COVID-19 hospitalisation, disease severity and death, we therefore reasoned *a priori* that historical controls – i.e. patients hospitalised at the same trust with another respiratory viral infection (e.g. influenza) a year previous – would act as a useful comparator, as they represent a geographically matched population at risk of severe disease from a circulating respiratory virus with a similar route of transmission (i.e. respiratory droplets and aerosols) and detection (i.e. laboratory-confirmed infection prior to or upon hospitalisation) (McCarthy and Giesecke, 1999). In addition, risk factors for hospitalisation with other respiratory viruses are similar to those for hospitalisation with COVID-19 (e.g. older age, comorbidities) (Falsey *et al*., 2014; Peralta *et al*., 2010).

In the present case-control study, we therefore first aimed to examine the association of smoking status with hospitalisation for COVID-19 compared with hospitalisation for other respiratory viral infections (e.g. influenza, respiratory syncytial virus) a year previous at a single UK hospital trust. Second, we aimed to examine whether there is a discordance between smoking status recorded on the summary EHR and within the contemporaneous medical notes. As current smoking in April 2020 (when our study protocol was registered) was *a priori* expected to be associated with an increased risk of COVID-19 hospitalisation (Alqahtani *et al*., 2020; Simons *et al*., 2020a), and with the association expected to be of a similar magnitude to that observed for other respiratory viruses, we opted for a non-inferiority design to test the hypothesis that the proportion of current smokers in patients hospitalised with COVID-19 is similar to that in patients hospitalised with other respiratory viral infections a year previous.

Methods

Ethics statement

This study was approved by the UCL/UCLH Joint Research Office Research Strategy Group and UCLH Data Access Committee. Approval to conduct research limited to pseudonymised patient data was provided by the NHS Health

Research Authority (IRAS_282704). The requirement for informed consent was waived by the NHS Health Research Authority due to the observational nature of the study.

Study design

This was an observational case-control study with historical controls, performed at a single National Health Service (NHS) hospital trust (comprising two hospital sites) in London, UK. The study protocol and analysis plan were pre-registered on the Open Science Framework in April 2020 (Simons *et al.*, 2020a). The pre-registered protocol stipulated a non-inferiority design (i.e. a one-tailed statistical test) to maximise statistical power to detect a significantly lower proportion of current smokers (i.e. <10%) among patients hospitalised with COVID-19 compared with patients hospitalised with another respiratory viral infection a year previous (i.e. 20%). The protocol was amended after data collection but prior to statistical analysis in September 2020 to implement a traditional case-control design (i.e. a two-tailed statistical test), as a delay in study approval meant that the number of eligible cases and controls exceeded our expectations; providing sufficient power for a two-tailed test. We had also planned to compare current smoking in cases with age- and sex-matched London prevalence, with data obtained from the representative Annual Population Survey. However, following an external review on an earlier manuscript draft, we decided against presenting data from this comparison due to smoking rates in hospitalised populations typically being greater than in the general population (Benowitz *et al.*, 2009).

A sample size calculation, updated after data collection but prior to data analysis, indicated that 363 cases and 109 controls would provide 80% power to detect a 10% difference in current smoking prevalence in cases compared with controls (e.g. 10% in cases and 20% in controls) with alpha set to 5%. We included all cases from 1st March 2020 to the 26th August 2020 (the date on which data were obtained) and all controls from the 1st January 2019 and the 31st December 2019.

Eligibility criteria

Inclusion criteria

Cases

1. Consecutive patients admitted to an adult hospital ward (i.e. 18+ years) between 1st March 2020 and 26th August 2020 (the date on which data were obtained).

2. Diagnosis of COVID-19 on or within five days of hospital admission, identified via associated International Classification of Diseases version 10 (ICD-10) codes (World Health Organisation, 2019). This temporal boundary was set to prevent inclusion of patents with nosocomial (hospital-acquired) infection and allowed for a delay of three days in requesting a COVID-19 test and two days for receiving and reporting the results on the EHR. The median incubation time for COVID-19 is estimated at 5.1 days (95% CI = 4.5-5.8) (Lauer *et al.*, 2020). We sought to exclude individuals with nosocomial COVID-19 infection as they are a different population (e.g. older, more frail) compared with those infected in the community and subsequently requiring hospitalisation.

Controls

1. Consecutive patients admitted to an adult hospital ward (i.e. 18+ years) between 1st January 2019 and 31st December 2019.

2. Diagnosis of a viral respiratory infection (e.g. influenza, parainfluenza) on or within 5 days of admission, identified via ICD-10 codes.

Exclusion criteria

1. No record of smoking status on the summary EHR or within the medical notes.

2. A primary diagnosis of infectious exacerbation of chronic obstructive pulmonary disease (COPD) due to the strong causal association of COPD with current and former smoking.

Measures

Data on demographic and smoking characteristics were collected from the summary EHR or the medical notes. In the UK, the summary EHR is produced at the point of an individual's first interaction with a specific NHS hospital trust. Further

information is added to the summary EHR following subsequent interactions with the hospital trust. The medical notes include contemporaneous clinical notes, General Practitioner referral letters and outpatient clinic letters, and are updated more frequently than the summary EHR.

Outcome variable

The outcome of interest was the type of hospital admission (i.e. with COVID-19 vs. other respiratory viral infections a year previous).

Exposure variable

Smoking status (i.e. current, former, never) was obtained from the summary EHR or the medical notes. A number of cases were recorded as 'non-smokers' without distinguishing between 'former smokers' and 'never smokers'. For the primary analysis, patients categorised as a 'non-smoker' were treated as 'never smokers'. Where possible, information on use of smokeless tobacco, waterpipe and/or alternative nicotine products (e.g. e-cigarettes) was extracted. We searched within the contemporaneous medical records for free-text entries of smoking status. The most recent record of smoking status, obtained from either the summary EHR or the contemporaneous medical notes, was extracted. Where there was discordance between the summary EHR and the contemporaneous medical notes, smoking status was classified based on the contemporaneous medical notes (as this clinician obtained information was assumed to be more recent and therefore accurate). Where available, data on pack-year history of smoking (i.e. the number of packs of cigarettes smoked per day multiplied by the number of years of smoking, with a pack equal to 20 cigarettes) were extracted.

Covariates

Covariates included age, sex, ethnicity, socioeconomic position (SEP; with post codes linked by the research team to the Index of Multiple Deprivation (IMD) (Department for Communities and Local Government, 2019)) and comorbidities (classified by organ system, including cardiac, metabolic and respiratory diseases). Medical conditions not expected to be strongly associated with COVID-19 hospitalisation were not considered in the analyses (e.g. sciatica and fibromyalgia; see *Extended data*). Age was treated as a continuous variable in the primary analysis, with banded age groups (i.e. 18-29 years, 30-44 years, 45-59 years, 60-74 years, 75-89 years and > 90 years) used in exploratory analyses. The IMD was categorised as quintiles to reduce the impact of sparse data.

Data analysis

All analyses were conducted in R version 4.0.2. (R Core Team, 2020). Descriptive statistics for cases and controls are reported. To explore differences between cases and controls, Pearson's Chi-square tests, Cochran-Armitage tests for trend and ANOVAs were used, as appropriate.

To examine the association of former and current smoking with hospitalisation for COVID-19 compared with hospitalisation for other respiratory viral infections, unadjusted and two different adjusted generalised linear models with a binomial distribution and logit link function were performed. The first model adjusted for age, sex and SEP, with a second model adjusting for age, sex, SEP and comorbidities. We report odds ratios (ORs), 95% confidence intervals (CIs) and *p*-values. Two sensitivity analyses were subsequently performed. First, those recorded as 'non-smokers' were removed from the analysis. Second, those excluded from the analytic sample due to missing data on smoking status (see section above on 'Exclusion criteria') were included and coded as i) 'never smokers' and then as ii) 'current smokers' to assess the robustness of the associations.

To examine the concordance between smoking status recorded on the summary EHR and within the contemporaneous medical notes, Pearson's Chi-squared tests were performed for the entire sample, and then separately for cases and controls.

Results

A total of 610 potential cases and 514 potential controls were identified. A total of 446 cases and 211 controls were included in the analytic sample (see Figure 1). In total, 13 potential controls and 60 potential cases were excluded due to not having a record of documented smoking status. This was likely due to patients having no prior contact with the NHS foundation trust. Notably, 37 (62%) potential cases that were excluded because of missing smoking status did not survive to hospital discharge, with no in-hospital mortality in potential controls, which suggests that data may be missing due to increased mortality in cases.

Compared with controls, cases were more likely to be male (55% vs. 35.9%) and older (64.9 years vs 62.5 years) (see Table 1). Approximately 10% of cases and controls had missing data for ethnicity. Compared with cases, controls were more likely to be admitted from more deprived areas (IMD quintiles 1 and 2) (41.8% vs. 32.9%, $p < 0.001$). Cases were more likely than controls to have pre-existing metabolic (30.3% vs 13.3%) and cardiac comorbidities (53.4% vs 30.3%).

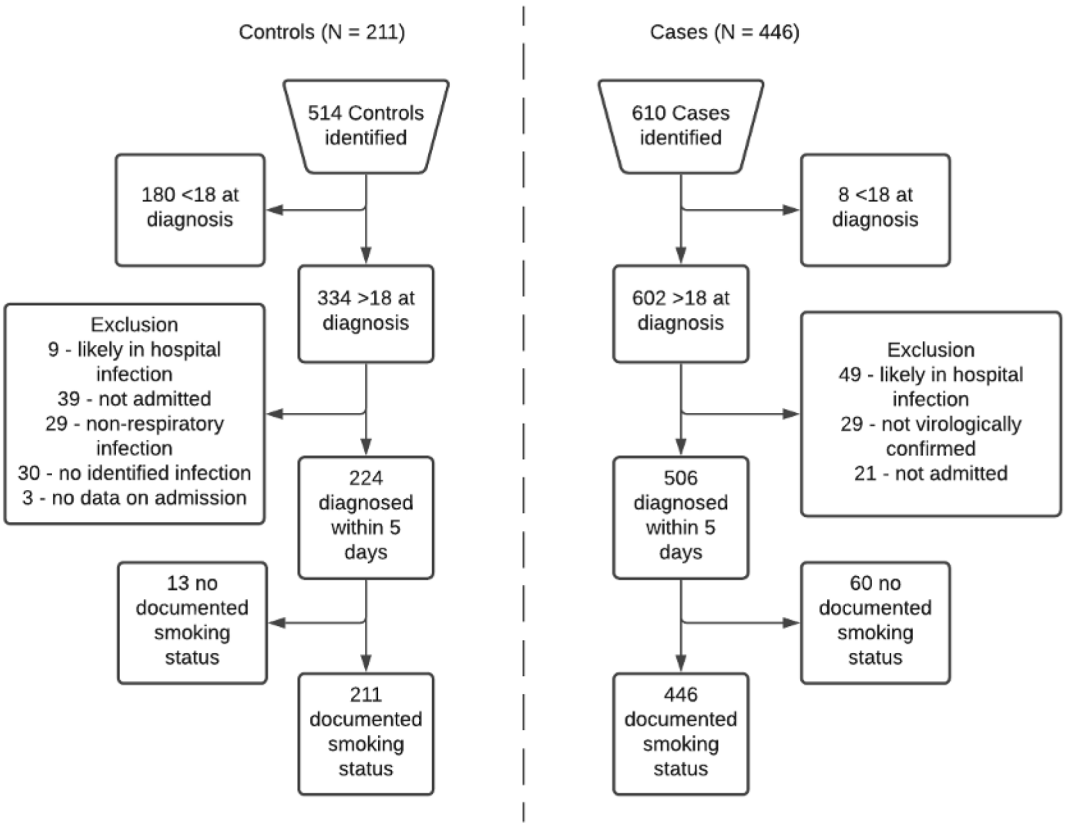


Figure 1. Eligibility flow diagram for controls (left hand side) and cases (right hand side).

A significantly larger proportion of cases compared with controls did not survive to discharge (28.7% vs. 4.3%). Among 128 cases not surviving to discharge, 53 (41.4%) were never smokers, 63 (49.2%) were former smokers and 12 (9.4%) were current smokers. For patients who survived to discharge, the median length of hospital stay for cases and controls was 9 (IQR = 4-18) and 4 (IQR = 2-9) days, respectively (see Table 1).

Cases and controls were predominantly admitted from North central and North East central London (see *Extended data*, Figure S1). The number of cases admitted from peripheral locations was greater than in controls and represents transfer of inpatients from other hospitals and diversion of patients that would otherwise have attended local hospitals due to bed pressures. The Chi-square test for trend found inconclusive evidence for any difference in SEP between cases and controls, $\chi^2(3) = 8.93$, $p = 0.06$ (see *Extended data*).

Association of smoking status with type of hospitalisation

The prevalence of former smoking was higher in cases compared with controls (38.6% vs. 31.8%). Current smoking prevalence was lower in cases compared with controls (9.4% vs. 17.1%). A single patient from the case sample was recorded as a dual cigarette and e-cigarette user. Two patients, one from each sample, were recorded as dual cigarette and shisha/waterpipe users. Pack-year history of smoking was only recorded for 40% of patients with a smoking history (see Table 1).

In the univariable analysis, current smokers had reduced odds of being hospitalised with COVID-19 compared with other respiratory viruses a year previous (OR = 0.52, 95% CI = 0.31-0.86, $p = 0.01$). The odds for former smokers were equivocal (OR = 1.16, 95% CI = 0.81-1.68, $p = 0.43$).

In the multivariable analysis adjusted for sex, age and SEP, current smokers had reduced odds of being hospitalised with COVID-19 compared with other respiratory viruses a year previous (OR = 0.48, 95% CI = 0.28-0.83, $p < 0.01$). There was no significant association among former smokers (OR = 0.90, 95% CI = 0.61-1.34, $p = 0.61$). Results were not materially altered when also adjusting for relevant comorbidities (current smokers, OR = 0.55, 95% CI = 0.31-0.96, $p = 0.04$; former smokers, OR = 1.08, 95% CI = 0.72-1.61, $p = 0.70$).

Table 1. Demographic and smoking characteristics of cases and controls.

	Controls (N = 211)	Cases (N = 446)	Total (N = 657)	*p* value
Female sex	116 (55.0%)	160 (35.9%)	276 (42.0%)	<0.001
Age in years (continuous)				0.019
- Median (IQR)	62.5 (48.5-75.8)	64.9 (52.4-76.2)	64.58 (51.3-76.1)	
Age in years (banded)				0.007
- 18-29	18 (8.5%)	11 (2.5%)	29 (4.4%)	
- 30-44	22 (10.4%)	37 (8.3%)	59 (9.0%)	
- 45-59	46 (21.8%)	113 (25.3%)	159 (24.2%)	
- 60-74	67 (31.8%)	162 (36.3%)	229 (34.9%)	
- 75+	58 (27.5%)	123 (27.6%)	181 (27.5%)	
Ethnicity				0.243
- Black British, Black African or Black Caribbean	19 (9.0%)	63 (14.1%)	82 (12.5%)	
- White British or White Other	116 (55.0%)	220 (49.3%)	336 (51.1%)	
- Not stated	26 (12.3%)	43 (9.6%)	69 (10.5%)	
- Asian British or Asian	28 (13.3%)	68 (15.2%)	96 (14.6%)	
- Other	22 (10.4%)	52 (11.7%)	74 (11.3%)	
IMD quintile				0.06
- Missing[a]	5 (2.4%)	5 (1.1%)	10 (1.5%)	
- 1 - Most deprived	21 (10.0%)	26 (5.8%)	47 (7.2%)	
- 2	67 (31.8%)	121 (27.1%)	188 (28.6%)	
- 3	42 (19.9%)	122 (27.4%)	164 (25.0%)	
- 4	38 (18.0%)	98 (22.0%)	136 (20.7%)	
- 5 - Least deprived	38 (18.0%)	74 (16.6%)	112 (17.0%)	
Most recent record of smoking status (summary EHR or medical notes)				0.012
- Never smoker	108 (51.2%)	232 (52.0%)	340 (51.8%)	
- Former smoker	67 (31.8%)	172 (38.6%)	239 (36.4%)	
- Current smoker	36 (17.1%)	42 (9.4%)	78 (11.9%)	
Record of pack-year smoking history in former and current smokers	36 (35%)	91 (42.5%)	127 (40%)	
- Pack-year smoking history in current smokers, median (IQR)	30 (22.5-45)	30 (6-45)	30 (16.8-45.5)	
- Pack-year smoking history in former smokers, median (IQR)	20 (10-40)	25 (10-40)	25 (10-40)	
Recorded smoking status on summary EHR				<0.001
- Never smoker	23 (10.9%)	233 (52.2%)	256 (39.0%)	
- Former smoker	22 (10.4%)	152 (34.1%)	174 (26.5%)	
- Current smoker	7 (3.3%)	30 (6.7%)	37 (5.6%)	
- Unknown	159 (75.4%)	31 (7.0%)	190 (28.9%)	
Survival to discharge	202 (95.7%)	318 (71.3%)	520 (79.1%)	<0.001

Table 1. *Continued*

	Controls (N = 211)	Cases (N = 446)	Total (N = 657)	*p* value
Length of hospital admission for survivors (days)				<0.001
- Median (IQR)	4 (2-9)	9 (4-18)	7 (3-14)	
Cancer (current or past)	66 (31.3%)	90 (20.2%)	156 (23.7%)	0.002
Auto-immune disease (present)	12 (5.7%)	16 (3.6%)	28 (4.3%)	0.213
Metabolic disease (present)	28 (13.3%)	135 (30.3%)	163 (24.8%)	<0.001
Haematological disease (present)	3 (1.4%)	4 (0.9%)	7 (1.1%)	0.541
Cardiac disease (present)	64 (30.3%)	238 (53.4%)	302 (46.0%)	<0.001
Neurological disease (present)	24 (11.4%)	54 (12.1%)	78 (11.9%)	0.786
Respiratory disease (present)	54 (25.6%)	91 (20.4%)	145 (22.1%)	0.134
Renal disease (present)	12 (5.7%)	37 (8.3%)	49 (7.5%)	0.235
HIV (present)	2 (0.9%)	7 (1.6%)	9 (1.4%)	0.522
No relevant comorbidities	43 (20.4%)	85 (19.1%)	128 (19.5%)	0.690

[a]IMD is not available for individuals with home addresses outside of England.

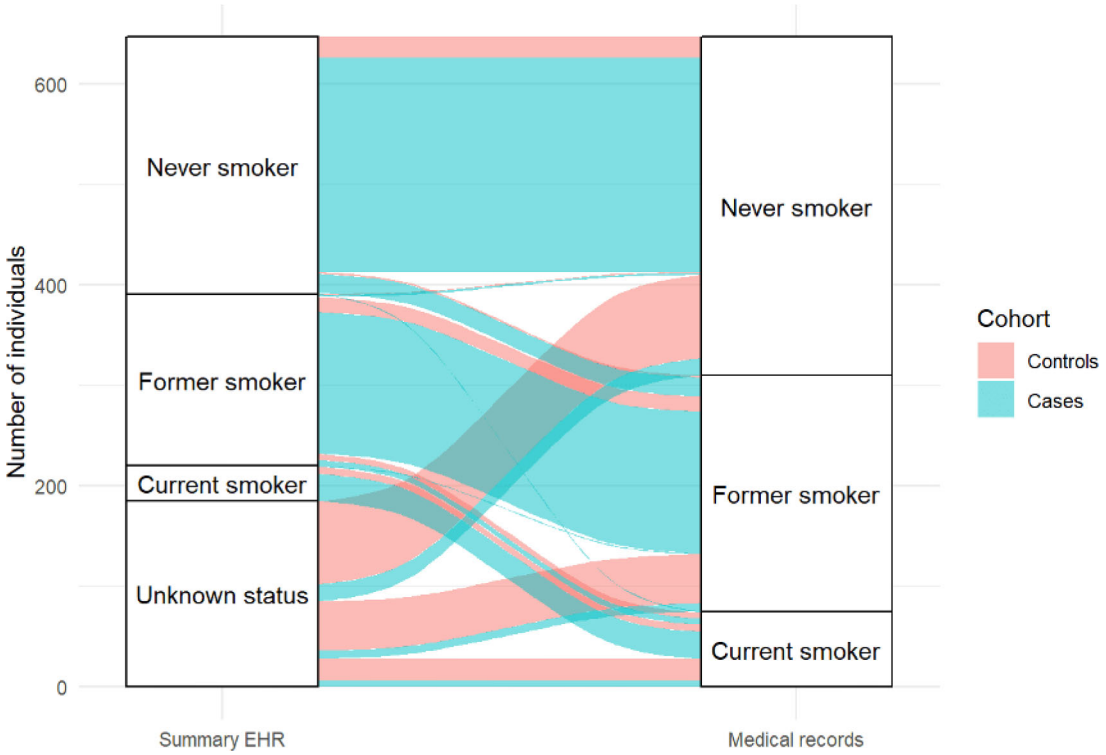


Figure 2. Concordance between smoking status recorded on the summary EHR and the medical notes for controls (red) and cases (blue).

Sensitivity analyses

First, in a sensitivity analysis with patients recorded as 'non-smokers' excluded from the sample (leaving 398 cases and 159 controls), current smokers had reduced odds of being hospitalised with COVID-19 compared with other respiratory viruses a year previous (OR = 0.41, 95% CI = 0.22-0.74, $p = 0.03$). There was no significant association among former smokers (OR = 0.78, 95% C.I. = 0.49-1.23, $p = 0.28$).

Second, in a sensitivity analysis with those with missing data on smoking status (n = 73) treated as 'never smokers' (resulting in 506 cases and 224 controls), patients hospitalised with COVID-19 had reduced odds of being a current

smoker compared with those admitted with other respiratory viruses (OR = 0.51, 95% CI = 0.30-0.88, $p = 0.01$). Next, when those with missing data on smoking status were treated as 'current smokers', there was no significant association between current smoking and hospitalisation with COVID-19 (OR = 0.94, 95% CI = 0.61-1.46, $p = 0.80$).

Third, in a sensitivity analysis removing SEP from the multivariable model, retaining sex and age only, current smokers had reduced odds of being hospitalised with COVID-19 compared with other respiratory viruses a year previous (OR = 0.51, 95% CI = 0.31-0.86, $p < 0.01$). There was no significant association among former smokers (OR = 0.95, 95% CI = 0.65-1.40, $p = 0.80$).

Concordance of smoking status recorded on the summary EHR and the medical notes

Controls were more likely to have no record of smoking status on the summary EHR compared with cases (75.4% vs. 7%) (see Figure 2). However, smoking status could be ascertained from the contemporaneous medical notes for all included cases and controls. Smoking status on the summary EHR (including 'unknown' status) was incorrectly recorded for 168 (79.6%) controls and 60 cases (13.5%) ($\chi^2(3) = 226.7$, $p = < 0.001$). In cases, six current smokers were misclassified as former smokers, one current smoker as a never smoker and six current smokers had no record of smoking status on the summary EHR. In controls, six current smokers were misclassified as former smokers and 23 current smokers had no record of smoking status on the summary EHR. There was greater discordance between smoking status recorded on the summary EHR and within the contemporaneous medical notes in controls ($\chi^2(3) = 256.5$, $p = < 0.001$) than in cases ($\chi^2(3) = 34.2$, $p = < 0.001$).

Discussion

This observational case-control study with patients admitted to a single UK hospital trust found a lower proportion of current smokers in cases hospitalised with COVID-19 during the first phase of the pandemic compared with controls hospitalised with other respiratory viral infections a year previous. Further, we found that smoking status is typically poorly recorded in the summary EHR. This was more prominent in controls than cases – a difference that is likely explained by the observation that COVID-19 patients were followed up by the respiratory medicine team after discharge, as part of a COVID-19 follow-up clinic where they specifically asked about smoking status (Mandal *et al.*, 2020). The observed discrepancy between smoking status recorded on summary EHRs and the contemporaneous medical notes is a concern, particularly for studies relying solely on EHRs as the source of information on smoking status.

A living review and unadjusted Bayesian meta-analysis of observational studies conducted during the COVID-19 pandemic, up-to-date as of July 2021, has found that current smokers appear to be at reduced risk of COVID-19 infection, but that there is as inconclusive evidence of an increased risk of more severe disease among smokers who are infected. In addition, former smokers appear to be at increased risk of severe disease and mortality from COVID-19 (Simons *et al.*, 2020b, 2021a).

Strengths and limitations

To our knowledge, this is one of few studies specifically designed to examine the association between smoking status and hospitalisation with COVID-19. It was further strengthened by an assessment of the quality of data on smoking status gleaned from summary EHRs.

However, this study has several important limitations, the majority of which pertain to the selection of the controls. First, current smoking is expected *a priori* to be associated with hospitalisation for non-COVID-19 respiratory viruses (Stämpfli & Anderson, 2009). As concluded in the US Surgeon General's 2014 report, smoking and infection with acute respiratory illnesses (including pneumonia) appear causally related. However, evidence as to the association between smoking and increased disease severity (once infected) is currently lacking. Ideally, hospital-based case-control studies should avoid selecting a control disease which is associated with the exposure of interest (i.e. smoking status) (Vandenbroucke & Pearce, 2012). However, to our knowledge, there is no other control disease with a similar route of acquisition and mechanism for hospitalisation/severe disease that is not *a priori* also associated with smoking status. The greater smoking prevalence in controls compared with the general population from which the cases emerged (Vandenbroucke & Pearce, 2012) therefore likely contributes to the significantly reduced odds of current smoking in our cases.

Second, the risk profile for controls likely differs from cases in that there is prior immunity to other respiratory viruses (e.g. influenza, respiratory syncytial virus), with no prior immunity in the population to SARS-CoV-2.

Third, we selected the controls on the basis of sharing a similar route of transmission and risk factors for hospitalisation as cases. However, at the time of writing (March 2021), we now suspect that COVID-19 differs from other respiratory viruses in several ways. For example, COVID-19 gains cell entry via the ACE-2 receptor (Hoffmann *et al.*, 2020), with

unknown receptor binding in flu (Killingley & Nguyen-Van-Tam, 2013) and appears to display less fomite and physical contact transmission than flu (Ben-Shmuel *et al*., 2020). In addition, emerging evidence suggests that COVID-19 has a significantly different pathological process compared with other respiratory viruses. For example, mortality rates from COVID-19 differ widely from those due to epidemic influenza (Office for National Statistics, 2020a). Although we currently do not know the importance of these factors, taken together, emerging observations may mean that direct comparison of risk profiles in cases and controls is limited.

Fourth, while no known behavioural restrictions were implemented during the control period, London was under lockdown restrictions from March to July 2020, which likely impacted the risk of viral exposure in cases (Davies *et al*., 2020). This may further have impacted the different risk profiles of controls and cases beyond the adjustments made in this analysis for sex, age and SEP.

Fifth, hospital admission routines were severely disrupted in many countries during the early stages of the COVID-19 pandemic (e.g., patient transfer between hospitals, altered admission criteria). This may have influenced patient selection and/or the recording of smoking status. With regards to patient selection, albeit possible, we are unaware of evidence from UK hospitals of patient characteristics correlated with smoking (e.g., age, SEP) determining whether a patient would be admitted, rather than their clinical situation on presentation. For example, smoking status has not been included in a widely used hospital-based algorithm for predicting clinical deterioration from COVID-19 (i.e., the ISARIC-4C deterioration score; Gupta *et al.*, 2021). With regards to the recording of smoking status, to the best of our knowledge, the admission routines (including the medical history taking) within the selected hospital trust remained stable during the pandemic. However, patients admitted with COVID-19 may generally have been more unwell at the point of admission compared with controls, which may have led to less accurate documentation of smoking status. This potential bias was somewhat reduced by those surviving their COVID-19 illness being followed up through a specialised respiratory medicine clinic implemented during the pandemic, with smoking status ascertained during follow-up calls/visits. However, the finding that COVID-19 patients with unknown (compared with those with recorded) smoking status had greater in-hospital mortality means that residual bias cannot be ruled out. A sensitivity analysis with all missing assumed to be current smokers made the negative association of smoking status and COVID-19 go away. However, it should be noted that this is a very strong assumption, unlikely to hold true.

Sixth, the selection of historical controls may mean that there are non-trivial differences in smoking status between controls and cases due to a declining trend in London smoking prevalence (Office for National Statistics, 2020b). However, a single year was used for the selection of controls, and there was no important change in smoking prevalence among all (Jackson *et al.*, 2021), so we expect any potential impact of pandemic-related trends in smoking prevalence on the results to be minimal. We considered using a contemporaneous control (i.e. patients hospitalised with other respiratory viral infections in 2020), which would have mitigated against this potential bias. However, due to factors such as reduced national and international travel, physical distancing, increased hand hygiene and potential viral dominance by COVID-19, exposure to and hospitalisation with other respiratory viruses has been substantially reduced in 2020 (GOV.UK, 2020), which would have limited the sample size for controls.

Seventh, given the broad and currently poorly understood COVID-19 pathology following SARS-CoV-2 infection, it was not possible to differentiate patients presenting to hospital due to COVID-19 or those admitted for other reasons. This may have introduced additional bias in that individuals may have been admitted to hospital with co-incident SARS-CoV-2 infection that was not associated with their current health need. We tried to mitigate this by only including cases if they had a positive COVID-19 test result less than 5 days from their date of hospital admission.

Eighth, a history of current or past cancer was high in both groups at greater than 20% and was significantly greater in controls compared with cases. This reflects a bias in the population that regularly interacts with the selected NHS hospital trust, which is a specialist cancer referral centre. We visualised the geographic regions where patients were admitted from to examine any systemic differences between cases and controls, and caution that the differing catchment areas of the two samples may have led to important differences in the underlying populations. In addition, during the peak of the first wave of the pandemic in the UK (i.e. March-April 2020), many cases were transferred across hospital sites due to bed pressures (Dunhill, 2020).

Finally, we have argued elsewhere that studies that aim to elucidate the potential causal relationship between smoking status and COVID-19 should first clarify the mechanism(s) through which smoking and/or nicotine use is expected to influence COVID-19 infection and/or disease outcomes contingent on infection and plan their studies accordingly (Perski *et al.*, 2021). Considering these reflections, and following constructive peer review, we acknowledge that the present study through its selection of the control sample does not distinguish between two potential mechanisms: comparable

exposure to SARS-CoV-2 virus with differential infection in current compared with never smokers, and comparable SARS-CoV-2 infection with differential hospitalisation in current compared with never smokers. Future research should aim to isolate these potential mechanisms by design, thus getting closer to unbiased estimates of the association of smoking/nicotine use with COVID-19 outcomes. However, we note that – given the time-varying dynamics of respiratory viruses – it would be incredibly challenging to design a representative study to compare rates of infection in those with similar exposure using available datasets. Similarly, as data linkage of representative infection surveys with hospital records is also limited at present, designing a study to examine the second potential pathway (i.e., examining hospitalisation rates in a representative sample of those infected) would also be very challenging.

Despite these limitations, alternative designs were impracticable or would have had different limitations. In the future, the current study can be considered alongside findings across multiple such alternative methodological approaches, each with different sources of bias, to triangulate on the extent to which associations between smoking and COVID-19 are causal.

Implications for policy and practice

COVID-19 will continue to place a large burden on healthcare services in the UK and internationally over the coming months and years. To mitigate against this, multiple non-pharmacological interventions are being implemented to reduce the intensity of demand on acute and intensive services. Irrespective of any direct link between smoking and COVID-19 disease outcomes, smoking is a significant cause for healthcare demand globally. We have argued elsewhere for the need to ramp up smoking cessation support to reduce the current and future burden on healthcare and social services (Simons *et al*., 2020a).

Avenues for future research

The selection of appropriate controls in hospital-based case-control studies is very challenging for a novel respiratory virus such as COVID-19 (which means we converged on a hybrid approach, combining elements from hospital-based case series and case-control designs with historical controls). We recommend the use of representative population-studies with data from multiple sites and with purposeful acquisition of smoking status, to better understand the role of smoking as a potential risk or protective factor for COVID-19 hospitalisation and disease severity.

Conclusion

In a single hospital trust in the UK, patients hospitalised with COVID-19 had lower odds of appearing as smokers compared with patients with other respiratory viruses a year previous, although we caution against interpreting this as a causal association. Smoking status was poorly recorded, with high observed discordance between smoking status recorded on the summary EHR and the contemporaneous medical notes.

Data availability

Underlying data

Due to the sensitive nature of the data, we do not have ethical approval to release the individual-level data underpinning the analyses. Anonymised and de-identified individual-level data are available upon request from the corresponding author to bona fide researchers and following approval from the Biomedical Research Centre Clinical and Research Informatics Unit at University College London Hospital NHS foundation trust.

Extended data

OSF: Association of smoking status with hospitalisation for COVID-19 compared with other respiratory viruses a year previous: a case-control study at a single UK National Health Service trust: Protocol. https://doi.org/10.17605/OSF.IO/84VYD (Simons *et al*., 2021b).

This project contains the following extended data:

- Date file 1. (*Extended data*, PDF format)

- Data file 2. (Protocol, docx format)

Data are available under the terms of the Creative Commons Attribution 4.0 International (CC-By Attribution 4.0 International).

References

Alqahtani JS, Oyelade T, Aldhahir AM, *et al.*: **Prevalence, Severity and Mortality associated with COPD and Smoking in patients with COVID-19: A Rapid Systematic Review and Meta-Analysis.** *PLoS ONE.* 2020. 15.
PubMed Abstract | Publisher Full Text | Free Full Text

Benowitz NL, Schultz KE, Haller CA, *et al.*: **Prevalence of smoking assessed biochemically in an urban public hospital: a rationale for routine cotinine screening.** *Am. J. Epidemiol.* 2009; **170**: 885–891.
PubMed Abstract | Publisher Full Text | Free Full Text

Ben-Shmuel A, Brosh-Nissimov T, Glinert I, *et al.*: **Detection and infectivity potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) environmental contamination in isolation units and quarantine facilities.** *Clin. Microbiol. Infect.* 2020; **26**: 1658–1662.
PubMed Abstract | Publisher Full Text | Free Full Text

Brake SJ, Barnsley K, Lu W, *et al.*: **Smoking Upregulates Angiotensin-Converting Enzyme-2 Receptor: A Potential Adhesion Site for Novel Coronavirus SARS-CoV-2 (Covid-19).** *J Clin Med.* 2020; **9**: 841.
PubMed Abstract | Publisher Full Text | Free Full Text

Brown J: **The Smoking Toolkit Study, 2020.** 2020.
Reference Source

Cai G: **Bulk and Single-Cell Transcriptomics Identify Tobacco-Use Disparity in Lung Gene Expression of ACE2, the Receptor of 2019-nCov.** 2020.
Publisher Full Text | Reference Source | Reference Source

Davies NG, Kucharski AJ, Eggo RM, *et al.*: **Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study.** *Lancet Public Health.* 2020; **5**: e375–e385.
Publisher Full Text

Department for Communities and Local Government: **The English Indices of Deprivation 2019 31.** 2019.

Falsey AR, McElhaney JE, Beran J, *et al.*: **Respiratory Syncytial Virus and Other Respiratory Viral Infections in Older Adults With Moderate to Severe Influenza-like Illness.** *J. Infect. Dis.* 2014; **209**: 1873–1881.
Publisher Full Text

Fang L, Karakiulakis G, Roth M: **Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection?** *Lancet Respir Med.* 2020. **0**.
Publisher Full Text

GOV.UK: **Weekly national flu and COVID-19 surveillance reports published [WWW Document].** GOV.UK. 2020. (accessed 3.11.21).
Reference Source

Grimes DA, Schulz KF: **Compared to what? Finding controls for case-control studies.** *Lancet.* 2005; **365**: 1429–1433.
Publisher Full Text

Guan W, Liang W, Zhao Y, *et al.*: **Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis.** *Eur. Respir. J.* 2020; 55.
PubMed Abstract | Publisher Full Text | Free Full Text

Gupta RK, Harrison EM, Ho A, *et al.*: **Development and validation of the ISARIC 4C Deterioration model for adults hospitalised with COVID-19: a prospective cohort study.** *Lancet Respir. Med.* 2021; **9**4: P349–359.
Publisher Full Text

Hoffmann M, Kleine-Weber H, Schroeder S, *et al.*: **SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor.** *Cell.* 2020; **181**: 271-280.e8.
Publisher Full Text

Jackson SE, Garnett C, Shahab L, *et al.*: **Association of the COVID-19 lockdown with smoking, drinking and attempts to quit in England: an analysis of 2019–20 data.** *Addiction.* 2021; **116**: 1233–1244.
Publisher Full Text

Johns Hopkins Coronavirus Resource Center: **COVID-19 Map [WWW Document].** *Johns Hopkins Coronavirus Resour. Cent.* 2021. (accessed 3.11.21).
Reference Source

Killingley B, Nguyen-Van-Tam J: **Routes of influenza transmission.** *Influenza Other Respir. Viruses.* 2013; **7**(Suppl 2): 42–51.
PubMed Abstract | Publisher Full Text | Free Full Text

Lauer SA, Grantz KH, Bi Q, *et al.*: **The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application.** *Ann. Intern. Med.* 2020; **172**: 577–582.
PubMed Abstract | Publisher Full Text | Free Full Text

Mandal S, Barnett J, Brill SE, *et al.*: **'Long-COVID': a cross-sectional study of persisting symptoms, biomarker and imaging abnormalities following hospitalisation for COVID-19.** *Thorax.* 2020.
PubMed Abstract | Publisher Full Text | Free Full Text

McCarthy N, Giesecke J: **Case-case comparisons to study causation of common infectious diseases.** *Int. J. Epidemiol.* 1999; **28**: 764–768.
PubMed Abstract | Publisher Full Text

Oakes JM, Fuchs RM, Gardner JD, *et al.*: **Nicotine and the renin-angiotensin system.** *Am J Physiol - RegulIntegr Comp Physiol.* 2018; **315**: 895–906.
PubMed Abstract | Publisher Full Text | Free Full Text

Office for National Statistics: **Deaths due to coronavirus (COVID-19) compared with deaths from influenza and pneumonia, England and Wales [WWW Document].** 2020a. (accessed 10.29.20).
Reference Source

Office for National Statistics: **Adult smoking habits in the UK: 2019, Annual Population Survey.**Office for National Statistics; 2020b.

Peralta PS-O, Cortes-García M, Vicente-Herrero M, *et al.*: **Risk factors for disease severity among hospitalised patients with 2009 pandemic influenza A (H1N1) in Spain, April – December 2009.** *Eurosurveillance.* 2010; **15**: 19667.
PubMed Abstract | Publisher Full Text

Perski O, Simons D, Shahab L, *et al.*: **Smoking, nicotine and COVID-19: Triangulation of methods and pre-registration are required for robust causal inference.** *Nicotine Tob. Res.* 2021; ntab214.
Publisher Full Text

Polubriaginof F, Salmasian H, Albert DA, *et al.*: **Challenges with Collecting Smoking Status in Electronic Health Records.** *AMIA. Annu. Symp. Proc.* 2018; **2017**: 1392–1400.
PubMed Abstract | Free Full Text

R Core Team: **R: A Language and Environment for Statistical Computing.** *R Foundation for Statistical Computing, Vienna, Austria.* 2020.

Secondary Care Analytical Team: **Hospital Admitted Patient Care Activity 2019-20.** *Hospital Admitted Patient Care Activity.* 2020.

Simons D, Perski O, Brown J: **Covid-19: The role of smoking cessation during respiratory virus epidemics.** *The BMJ.* 2020a.

Simons D, Shahab L, Brown J, *et al.*: **The association of smoking status with SARS-CoV-2 infection, hospitalization and mortality from COVID-19: a living rapid evidence review with Bayesian meta-analyses (version 7).** *Addiction.* 2020b; **116**(6): 1319–1368.
PubMed Abstract | Publisher Full Text | Free Full Text

Simons D, Shahab L, Brown J, *et al.*: **The association of smoking status with SARS-CoV-2 infection, hospitalisation and mortality from COVID-19: A living rapid evidence review with Bayesian meta-analyses (version 12).** *Qeios.* 2021a.
Publisher Full Text

Simons D, Perski O, Brown J, *et al.*: *Protocol.* 2021b.
Publisher Full Text

Stämpfli MR, Anderson GP: **How cigarette smoke skews immune responses to promote infection, lung disease and cancer.** *Nat. Rev. Immunol.* 2009; **9**: 377–384.
PubMed Abstract | Publisher Full Text

Vandenbroucke JP, Pearce N: **Case–control studies: basic concepts.** *Int. J. Epidemiol.* 2012; **41**: 1480–1489.
PubMed Abstract | Publisher Full Text

World Health Organisation: **ICD-10 [WWW Document].** 2019. (accessed 10.28.20).
Reference Source

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 3

Reviewer Report 27 January 2022

https://doi.org/10.5256/f1000research.120034.r120026

© **2022 Samet J.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Jonathan M. Samet**
Colorado School of Public Health, Aurora, CO, USA

My two remaining concerns have been satisfactorily addressed and the manuscript is suitable for indexing.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Pulmonary disease, epidemiology, tobacco control

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 2

Reviewer Report 25 November 2021

https://doi.org/10.5256/f1000research.79264.r100232

© **2021 Galanti M.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Maria Rosaria Galanti**
Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

The authors did a considerable effort to comply with the review, and the article is now acceptable for indexing.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiology, tobacco use

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 16 November 2021

https://doi.org/10.5256/f1000research.79264.r100233

© 2021 Samet J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Jonathan M. Samet
Colorado School of Public Health, Aurora, CO, USA

1. With regard to the exposure classification, in the instances when the notes and the HER were in disagreement, how was exposure classified? There does not seem to be a specification as to how this situation was handled.

2. Smoking has been causally associated with respiratory infections, including influenza (see the 2014 report of the US Surgeon General). Consider the implications for interpretation of the findings.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Pulmonary disease, epidemiology, tobacco control

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 05 Jan 2022

Olga Perski, University College London, London, UK

Thank you for your helpful comments.

With regard to the exposure classification, in the instances when the notes and the HER were in disagreement, how was exposure classified? There does not seem to be a specification as to how this situation was handled.

We have now clarified this in the Methods section (within the 'Exposure variable' subsection):

"The most recent record of smoking status, obtained from either the summary EHR or the contemporaneous medical notes, was extracted. Where there was discordance between the summary EHR and the contemporaneous medical notes, smoking status was classified based on the contemporaneous medical notes (as this clinician obtained information was assumed to be more recent and therefore accurate)."

Smoking has been causally associated with respiratory infections, including influenza (see the 2014 report of the US Surgeon General). Consider the implications for interpretation of the findings.

Thank you – we now refer to the US Surgeon General's report in the Discussion:

"As concluded in the US Surgeon General's 2014 report, smoking and infection with acute respiratory illnesses (including pneumonia) appear causally related. However, evidence as to the association between smoking and increased disease severity (once infected) is currently lacking."

Competing Interests: None.

Version 1

Reviewer Report 08 October 2021

https://doi.org/10.5256/f1000research.59083.r94125

© 2021 Samet J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Jonathan M. Samet
Colorado School of Public Health, Aurora, CO, USA

General Comments:

This paper describes the findings of a hospital-based case-control study on cigarette smoking on risk for hospitalization from COVID-19. The paper is lacking a clear explication of the underlying hypothesis, which is critical to determining what is the appropriate control group. Smoking might influence the risk of becoming infected with SARS-CoV-2; and/or the risk of more severe disease given infection, which could lead to increased risk for hospitalization and death. The effects of smoking on airway receptors could plausibly effect risk for infection and modify its severity, as could the effects of smoking on the lungs and its causation of non-communicable diseases.

The authors write extensively about the problems of control group selection but, lacking a clearly focused hypothesis, their discussion is general. If the question is whether smoking affects the risk

of hospitalization, given infection, then the appropriate control group comprises non-hospitalized, SARS-CoV-2 infection who were not hospitalized. If the question related to risk for infection, than the appropriate control group has comparable exposure but without becoming infected.

The authors turn to a problematic historical control group of individuals hospitalized with other respiratory infections during a pre-pandemic time period. How this control group relates to the potential hypotheses is unclear. Additionally, there may have been pandemic-related trends in smoking that would not be taken into account. Additionally, smoking increases risk for respiratory infections generally (see 2014 report of the US Surgeon General[1]), further complicating interpretation of the results.

Why is SEP included in the model? It is a powerful predictor of smoking and of co-morbidities, but is it a likely confounder and why?

Were the cases all diagnosed with COVID-19? What about those admitted for other reasons, but testing positive for SARS-CoV-2 on admission?

The discussion does not adequately integrate this new study into existing literature, including ongoing systematic reviews.

References

1. National Center for Chronic Disease Prevention and Health Promotion US Office on Smoking and Health: The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. 2014. PubMed Abstract

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Pulmonary disease, epidemiology, tobacco control

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 08 Nov 2021

Olga Perski, University College London, London, UK

This paper describes the findings of a hospital-based case-control study on cigarette smoking on risk for hospitalization from COVID-19. The paper is lacking a clear explication of the underlying hypothesis, which is critical to determining what is the appropriate control group. Smoking might influence the risk of becoming infected with SARS-CoV-2; and/or the risk of more severe disease given infection, which could lead to increased risk for hospitalization and death. The effects of smoking on airway receptors could plausibly effect risk for infection and modify its severity, as could the effects of smoking on the lungs and its causation of non-communicable diseases.

The authors write extensively about the problems of control group selection but, lacking a clearly focused hypothesis, their discussion is general. If the question is whether smoking affects the risk of hospitalization, given infection, then the appropriate control group comprises non-hospitalized, SARS-CoV-2 infection who were not hospitalized. If the question related to risk for infection, than the appropriate control group has comparable exposure but without becoming infected.
The authors turn to a problematic historical control group of individuals hospitalized with other respiratory infections during a pre-pandemic time period. How this control group relates to the potential hypotheses is unclear.

Thank you for the constructive feedback on our paper, which we believe have helped strengthen it.

We agree that the selection of controls ultimately mixes two potential mechanisms through which smoking status may influence COVID-19 outcomes (i.e., comparable exposure leading to differential infection rates and/or comparable infection rates leading to differential hospitalisation rates). As elaborated on in the discussion, and now clarified in the introduction on p.4-5, we selected the controls on the basis of sharing a similar route of transmission (i.e., influenza) and risk factors for hospitalisation as cases. Since planning this study in early 2020, the research community has learnt more about COVID-19 infection and disease mechanisms, and we have ourselves taken more care to map out potential causal pathways that need to be considered when planning studies to examine the relationship of smoking status and COVID-19 infection/disease to avoid confounding (see Perski et al., 2021). We now discuss these considerations and the issues with the selection of the controls on p.13-14:

“Finally, we have argued elsewhere that studies that aim to elucidate the potential causal relationship between smoking status and COVID-19 should first clarify the mechanism(s) through which smoking and/or nicotine use is expected to influence COVID-19 infection and/or disease outcomes contingent on infection and plan their studies accordingly (Perski

et al., 2021). Considering these reflections, and following constructive peer review, we acknowledge that the present study through its selection of the control sample does not distinguish between two potential mechanisms: comparable exposure to SARS-CoV-2 virus with differential infection in current compared with never smokers, and comparable SARS-CoV-2 infection with differential hospitalisation in current compared with never smokers. Future research should aim to isolate these potential mechanisms by design, thus getting closer to unbiased estimates of the association of smoking/nicotine use with COVID-19 outcomes. However, we note that – given the time-varying dynamics of respiratory viruses – it would be incredibly challenging to design a representative study to compare rates of infection in those with similar exposure using available datasets. Similarly, as data linkage of representative infection surveys with hospital records is also limited at present, designing a study to examine the second potential pathway (i.e., examining hospitalisation rates in a representative sample of those infected) would also be very challenging."

Additionally, there may have been pandemic-related trends in smoking that would not be taken into account.

With regards to pandemic-related trends in smoking, we have now added the following to the discussion on p.14:

"However, a single year was used for the selection of controls, and there was no significant change in smoking prevalence among all adults in England between immediately before and early in the pandemic (Apr-Jul 2020, Jackson et al., 2020), so we expect any potential impact of pandemic-related trends in smoking prevalence on the results to be minimal."

Additionally, smoking increases the risk for respiratory infections generally (see 2014 report of the US Surgeon General1), further complicating interpretation of the results.

As the controls came from a similar population (i.e., individuals at risk of infection and hospitalisation with a circulating respiratory disease), we believe this was a strength in their selection. We have therefore not seen it necessary to discuss this beyond what is already mentioned about the selection of the control sample.

Why is SEP included in the model? It is a powerful predictor of smoking and of co-morbidities, but is it a likely confounder, and why?

SEP was included as a potential confounder due to being a strong predictor of both smoking and COVID-19 hospitalisation. In response to a similar comment from Reviewer 1, we have now included a sensitivity analysis with SEP removed (reported on p.10), with results remaining similar.

Were the cases all diagnosed with COVID-19? What about those admitted for other reasons, but testing positive for SARS-CoV-2 on admission?

Thank you for highlighting – all patients included in the case series had PCR confirmed infection with COVID-19. Given the broad and currently poorly understood COVID-19 pathology following infection, it was not possible to exclude patients with incident positive

tests based on their presenting symptoms to ensure the exclusion of those testing positive upon admission to hospital who were admitted for reasons other than COVID-19. This is an important potential source of bias and we have now included the following in the limitations on p.13:

"Seventh, given the broad and currently poorly understood COVID-19 pathology following SARS-CoV-2 infection, it was not possible to differentiate patients presenting to hospital due to COVID-19 or those admitted for other reasons. This may have introduced additional bias in that individuals may have been admitted to hospital with co-incident SARS-CoV-2 infection that was not associated with their current health need. We tried to mitigate this by only including cases if they had a positive COVID-19 test result less than 5 days from their date of hospital admission."

The discussion does not adequately integrate this new study into existing literature, including ongoing systematic reviews.

We have now included a paragraph in the discussion on p.11 which highlights how our findings compare with those from an ongoing systematic review:

"A living review and unadjusted Bayesian meta-analyses of observational studies conducted during the COVID-19 pandemic, up-to-date as of July 2021, has found that current smokers appear to be at reduced risk of COVID-19 infection, but that there is inconclusive evidence of an increased risk of more severe disease among smokers who are infected. In addition, former smokers appear to be at increased risk of severe disease and mortality from COVID-19 (Simons et al., 2020; Simons et al., 2021)."

Competing Interests: N/A

Reviewer Report 13 September 2021

https://doi.org/10.5256/f1000research.59083.r92725

© 2021 Galanti M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Maria Rosaria Galanti
Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

This was a case-control study aiming at shedding light on the potential causal association between tobacco smoking and hospitalization because of COVID-19. In the light of the controversial associations between smoking and risk for infection with SARS-COV-2 on the one side, clinical disease and its severity on the other side, the contribution has a great interest. The manuscript is clearly and soundly written, and the results are thoughtfully discussed within a series of sensitivity analyses.

However, the several weaknesses in the design and assessment of exposure, with corresponding risk for bias (to be sure, fully acknowledged by the authors) did not allow the attainment of the study's goal. Besides the weaknesses already addressed by the authors, the following points should be considered, bearing in mind that in a case-control study the concern is on characteristics being different or differentially assessed in the case and in the referent series and possibly associated with the exposure:

1. Can we assume that the procedure for hospital admissions for COVID-19 was the same as for other respiratory infections one year earlier (see page 6, lines 3-4)? To my knowledge, hospital admission routines during the pandemic were profoundly disrupted in many countries, with patient transfer between hospitals, different priorities of admission, etc. Even if not explicitly, some of the patient characteristics correlated with smoking may have determined different patient selection and/or ascertainment of smoking.

2. Could the assessment of smoking at the point of hospitalization among COVID-19 patients be biased by the pandemic or by the disease? Starting from March 2020, much attention was drawn on smoking, causing adverse outcomes in COVID-19 hospitalized patients, a position even endorsed by the WHO later in Spring 2020 (https://www.who.int/news-room/commentaries/detail/smoking-and-covid-19). Thereafter, some smokers (especially if already suffering from smoking-related diseases) may have quit before hospital admission or may have concealed their smoking to the healthcare staff. Also, the accuracy of the staff assessing smoking behavior may have varied depending on the severity of the disease. This risk is supported by the notation that the COVID-19 patients with unknown smoking status (more numerous than among the referent cases) had higher mortality than those with a full assessment of smoking. Following the premises in the introduction (logical expectation of higher risk among smokers), it would be sensible to assume that among these “unknown smoking” patients the proportion of smokers would be higher. When tested in a sensitivity analysis, this assumption made the inverse association smoking-COVID-19 hospital admission weaken/disappear (page 8).

3. The adjustment for SEP was done on the ecologic level. This may have introduced bias (over-adjustment for opportunity of infection, including due to high smoking prevalence?) if the rate of infection/hospital admission with SARS-COV-2 was differently clustered in socio-economically disadvantaged areas compared to affluent areas, at odds with the more “widespread” diffusion of other respiratory infections.

Some minor points:

1. I would avoid the notation “case cohort” and “control cohort” (page 6), which is misleading, using rather “case series” or “case sample”.

2. I would rephrase the conclusions as “...patients hospitalized with COVID-19 had lower odds of appearing as smokers compared with patients... ”, which is more respectful of what was actually analyzed.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiology, tobacco use

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 08 Nov 2021

Olga Perski, University College London, London, UK

This was a case-control study aiming at shedding light on the potential causal association between tobacco smoking and hospitalization because of COVID-19. In the light of the controversial associations between smoking and risk for infection with SARS-COV-2 on the one side, clinical disease and its severity on the other side, the contribution has a great interest. The manuscript is clearly and soundly written, and the results are thoughtfully discussed within a series of sensitivity analyses.

Thank you for the constructive feedback on our manuscript, which we believe has helped strengthen it.

However, the several weaknesses in the design and assessment of exposure, with the corresponding risk for bias (to be sure, fully acknowledged by the authors) did not allow the attainment of the study's goal. Besides the weaknesses already addressed by the authors, the following points should be considered, bearing in mind that in a case-control study the concern is on characteristics being different or differentially assessed in the case and in the referent series and possibly associated with the exposure:

Can we assume that the procedure for hospital admissions for COVID-19 was the same as for other respiratory infections one year earlier (see page 6, lines 3-4)? To my knowledge, hospital admission routines during the pandemic were profoundly disrupted in many countries, with patient transfer between hospitals, different priorities of admission, etc. Even if not explicitly, some of the patient characteristics correlated with smoking may have

determined different patient selection and/or ascertainment of smoking.

Could the assessment of smoking at the point of hospitalization among COVID-19 patients be biased by the pandemic or by the disease? Starting from March 2020, much attention was drawn on smoking, causing adverse outcomes in COVID-19 hospitalized patients, a position even endorsed by the WHO later in Spring 2020 (https://www.who.int/news-room/commentaries/detail/smoking-and-covid-19). Thereafter, some smokers (especially if already suffering from smoking-related diseases) may have quit before hospital admission or may have concealed their smoking to the healthcare staff. Also, the accuracy of the staff assessing smoking behavior may have varied depending on the severity of the disease. This risk is supported by the notation that the COVID-19 patients with unknown smoking status (more numerous than among the referent cases) had higher mortality than those with a full assessment of smoking. Following the premises in the introduction (logical expectation of higher risk among smokers), it would be sensible to assume that among these "unknown smoking" patients the proportion of smokers would be higher. When tested in a sensitivity analysis, this assumption made the inverse association smoking-COVID-19 hospital admission weaken/disappear (page 8).

Thank you - we have now reflected on these important points in the limitations section on p.12:

"Fifth, hospital admission routines were severely disrupted in many countries during the early stages of the COVID-19 pandemic (e.g., patient transfer between hospitals, altered admission criteria). This may have influenced patient selection and/or the recording of smoking status. With regards to patient selection, albeit possible, we are unaware of evidence from UK hospitals of patient characteristics correlated with smoking (e.g., age, SEP) determining whether a patient would be admitted, rather than their clinical situation on presentation. For example, smoking status has not been included in a widely used hospital-based algorithm for predicting clinical deterioration from COVID-19 (i.e., the ISARIC-4C deterioration score; Gupta et al., 2021). With regards to the recording of smoking status, to the best of our knowledge, the admission routines (including the medical history taking) within the selected hospital trust remained stable during the pandemic. However, patients admitted with COVID-19 may generally have been more unwell at the point of admission compared with controls, which may have led to less accurate documentation of smoking status. This potential bias was somewhat reduced by those surviving their COVID-19 illness being followed up through a specialised respiratory medicine clinic implemented during the pandemic, with smoking status ascertained during follow-up calls/visits. However, the finding that COVID-19 patients with unknown (compared with those with recorded) smoking status had greater in-hospital mortality means that residual bias cannot be ruled out. A sensitivity analysis with all missing assumed to be current smokers made the negative association of smoking status and COVID-19 hospital admission go away. However, it should be noted that this is a very strong assumption, unlikely to hold true."

The adjustment for SEP was done on the ecologic level. This may have introduced bias (over-adjustment for the opportunity of infection, including due to high smoking prevalence?) if the rate of infection/hospital admission with SARS-COV-2 was differently clustered in socio-economically disadvantaged areas compared to affluent areas, at odds

with the more "widespread" diffusion of other respiratory infections.

To investigate the role of SEP as a potential confounder in our analysis, we have now conducted an additional sensitivity analysis removing SEP from the multivariable model (reported on p.10). In the sensitivity analysis with sex and age only, current smokers had reduced odds of being hospitalised with COVID-19 compared with other respiratory viruses a year previous (OR = 0.51, 95% CI = 0.31-0.86, p = 0.01). There was no significant association among former smokers (OR = 0.95, 95% CI = 0.65-1.40, p = 0.80). These were similar to the analyses including SEP (OR = 0.48, 95% CI = 0.28-0.83, p < 0.01 and OR = 0.90, 95% CI = 0.61-1.34, p = 0.61, respectively).

Some minor points:

I would avoid the notation "case-cohort" and "control cohort" (page 6), which is misleading, using rather "case series" or "case sample".

Thank you for pointing out this difference – we have now changed to "case sample" throughout.

I would rephrase the conclusions as "...patients hospitalized with COVID-19 had lower odds of appearing as smokers compared with patients... ", which is more respectful of what was actually analyzed.

This has now been amended accordingly.

Competing Interests: N/A

Comments on this article

Version 1

Reader Comment 27 Aug 2021

Mohammad Ali, Uttara Adhunik Medical College and Hospital, Dhaka, Bangladesh

This study's results amazed me. Other studies found a strong association between smoking and hospitalization from COVID-19. Interestingly, we found an association between current smoking status and COVID-19 vaccine hesitancy. Find the study here https://pubmed.ncbi.nlm.nih.gov/34429316/.

Good luck.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research