

RESEARCH

Open Access



# Differential Risk of SARS-CoV-2 Infection by Occupation: Evidence from the Virus Watch prospective cohort study in England and Wales

Sarah Beale<sup>1,2\*</sup>, Susan Hoskins<sup>2</sup>, Thomas Byrne<sup>1</sup>, Wing Lam Erica Fong<sup>1</sup>, Ellen Fragaszy<sup>1,3</sup>, Cyril Geismar<sup>1,2</sup>, Jana Kovar<sup>2</sup>, Annalan M. D. Navaratnam<sup>1,2</sup>, Vincent Nguyen<sup>1,2</sup>, Parth Patel<sup>1</sup>, Alexei Yavlinsky<sup>1</sup>, Anne M. Johnson<sup>4</sup>, Martie Van Tongeren<sup>5</sup>, Robert W. Aldridge<sup>1</sup> and Andrew Hayward<sup>2</sup> on behalf of the Virus Watch Collaborative

## Abstract

**Background** Workers across different occupations vary in their risk of SARS-CoV-2 infection, but the direct contribution of occupation to this relationship is unclear. This study aimed to investigate how infection risk differed across occupational groups in England and Wales up to April 2022, after adjustment for potential confounding and stratification by pandemic phase.

**Methods** Data from 15,190 employed/self-employed participants in the Virus Watch prospective cohort study were used to generate risk ratios for virologically- or serologically-confirmed SARS-CoV-2 infection using robust Poisson regression, adjusting for socio-demographic and health-related factors and non-work public activities. We calculated attributable fractions (AF) amongst the exposed for belonging to each occupational group based on adjusted risk ratios (aRR).

**Results** Increased risk was seen in nurses (aRR = 1.44, 1.25–1.65; AF = 30%, 20–39%), doctors (aRR = 1.33, 1.08–1.65; AF = 25%, 7–39%), carers (1.45, 1.19–1.76; AF = 31%, 16–43%), primary school teachers (aRR = 1.67, 1.42–1.96; AF = 40%, 30–49%), secondary school teachers (aRR = 1.48, 1.26–1.72; AF = 32%, 21–42%), and teaching support occupations (aRR = 1.42, 1.23–1.64; AF = 29%, 18–39%) compared to office-based professional occupations. Differential risk was apparent in the earlier phases (Feb 2020–May 2021) and attenuated later (June–October 2021) for most groups, although teachers and teaching support workers demonstrated persistently elevated risk across waves.

**Conclusions** Occupational differences in SARS-CoV-2 infection risk vary over time and are robust to adjustment for socio-demographic, health-related, and non-workplace activity-related potential confounders. Direct investigation into workplace factors underlying elevated risk and how these change over time is needed to inform occupational health interventions.

**Keywords** SARS-CoV-2, Infection, Pandemic, Occupational health

\*Correspondence:

Sarah Beale

[sarah.beale.19@ucl.ac.uk](mailto:sarah.beale.19@ucl.ac.uk)

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Notable occupational inequalities in infection risk have emerged during the Coronavirus Disease 2019 (COVID-19) pandemic. Research and surveillance data across various global regions have repeatedly indicated elevated risk of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in workers in various essential and/or public-facing industries, such as health and social care, transportation, education, and cleaning and service occupations [1–5] compared to other workers or the adult population. Occupational differences in the ability to work from home, the frequency and intensity of workplace exposure to other people, environmental features of the workspace, and the implementation of infection control procedures plausibly contribute to differential risk of infection and transmission at work [6–8]. However, occupation is intimately linked with other socio-demographic factors such as deprivation, household size, activities outside the workplace and health status, that can compound to influence infection risk [9, 10]. Establishing the contribution of work-related exposure to occupational inequalities in infection risk consequently depends on careful consideration of other non-occupational factors.

Few estimates of the effect of occupation on SARS-CoV-2 infection risk or outcomes have comprehensively accounted for sociodemographic confounding beyond age and sex.

Age, sex, geographic factors, education, living conditions, and pre-pandemic health were estimated to account for 70–80% of the effect of occupation on COVID-19 mortality in the UK in 2020 [11]. Health-care, care, and some service and transport occupations (among men) and elementary cleaning and plant workers (among women) demonstrated elevated mortality compared to all other occupations, but the strength of these estimates was greatly attenuated by adjustment. While these findings indicate the importance of comprehensive adjustment, mortality data are strongly affected by clinical risk factors and the impact of work-related factors on differential infection risk cannot therefore be inferred from these findings.

Data from Germany [12] (February – September 2020) and Sweden [13] (January 2020 – February 2021) indicated elevated risk of infection amongst essential workers – including health, care, and service workers – compared to non-essential workers across the respective study periods, after adjustment for a range of socio-demographic factors. However, occupational differences in risk may vary by global region and comparative investigation for the UK is limited. Probability of antigen test positivity differed little across occupations after adjustment for age, sex, region, ethnicity, household composition, deprivation, ability to work from home, use of face coverings at

work, and ability to socially distance at work, based on the UK Office for National Statistics (ONS) Coronavirus Infection Survey [14] between early September– early January 2021. However, the inclusion of work-related potential mediators in this analysis precludes disaggregating the impact of occupational and non-occupational factors.

Differential risk across occupations is also plausibly influenced by time, due to changes in public health interventions and restrictions—including sectoral closures, social distancing, and infection control in the workplace—as well as fluctuating levels of community transmission across the pandemic and changes in immunity due to infection or vaccination. Preliminary evidence from the UK and Norway suggests that occupational differences in infection risk vary across time, with health [3, 3, 15–17] and social care workers [15, 16] and transport workers [3] demonstrating elevated infection risk during the first pandemic wave and other public-facing occupations including education [15, 16], manufacturing [15, 16] and food service as well as transport workers [3] demonstrating elevated risk in the second wave. More recent data including the period of relaxation of pandemic restrictions in the UK are lacking, as are estimates over time comprehensively adjusted for non-occupational factors.

Using data from a prospective community cohort study in England and Wales (Virus Watch) [18], this study aimed to extend current understanding of the direct effect of occupation on SARS-CoV-2 infection risk over time. Specific objectives were: (1) to estimate the relative risk of SARS-CoV-2 infection by occupation across the pandemic, adjusting for socio-demographic and health-related factors and non-work public activities; (2) to investigate whether occupational infection risk differed across pandemic waves; and (3) to estimate the attributable fraction amongst the exposed for different occupations overall and by pandemic wave.

## Methods

### Participants

Participants in the current study ( $n = 15,190$ ) were an adult sub-cohort of the Virus Watch longitudinal cohort study ( $n = 58,692$  as of 12/02/2022 when cohort recruitment was completed). Participants were included in the present study if they were (1)  $\geq 16$  years, (2) in employment or self-employment and reported their occupation upon study registration, and (3) completed at least one monthly survey between November 2020 and March 2022 concerning their activities across a recent week. Further detail of the full Virus Watch cohort study, including inclusion criteria for the full cohort, can be obtained from the study protocol [18].

## Exposure

Occupation was derived based on free-text responses to the Virus Watch baseline survey (94% of classified responses) or a Virus Watch monthly survey conducted in February 2022 (6% of classified responses); the baseline survey was used as a preferential source, with the monthly survey used only if participants' occupation was missing at baseline. Following the protocol recommended by the UK Office for National Statistics (ONS) [19], we performed semi-automatic coding using Cascot Version 5.6.3 [20] to assign participants UK Standard Occupational Classification (SOC) 2020 codes [19]. Occupations were then classified into the following groups, which aimed to broadly reflect workplace environments while retaining, as far as possible, ONS-defined occupational skill groupings: administrative and secretarial occupations; healthcare occupations; indoor trade and process/plant occupations; leisure and personal service occupations; managers, directors, and senior officials; outdoor trade occupations; sales and customer service occupations; social care and community protective services; teaching education and childcare occupations; transport and mobile machine operatives; and other professional and associate occupations (broadly office-based professional and associate professional occupations).

Where possible, we also extracted more specific occupational groupings based on three-digit SOC groups for occupations within the essential worker classification [21] and classified by the investigators as public facing/frontline roles. These more detailed occupational groups were included where group sizes exceeded  $n = 100$  and some SOC groupings were split or combined together to reflect working environment/role, to yield the following included groupings: nurses, doctors, warehouse and process/plant occupations, food preparation and hospitality occupations, teachers (primary school), teachers (secondary school), teachers (higher education), teaching assistants and support occupations, carers, social work and welfare occupations, cleaners, and salespeople/cashiers/shopkeepers.

For further methodological details of exposure classification and UK SOC 2020 codes within each category, please see 'Occupational Classification' in the [Supplementary Materials](#).

## Outcomes

The outcome of interest was binary SARS-CoV-2 infection status (yes/no ever infected) based on any clinical evidence of infection (positive lateral flow (LFT), polymerase chain reaction (PCR), anti-nucleocapsid antibody serological test, or anti-spike antibody serological test in absence of vaccination. Susceptibility to reinfections

was not the focus of this paper; consequently, outcome data correspond to participants' first infection and follow-up was ended after first infection. Please see 'Clinical Outcomes' in the [Supplementary Material](#) for further information about clinical data in Virus Watch and how infection status was derived.

Where possible, we attributed results to the following phases of the pandemic based on test date: Wave 1 and 2 (February 2020 to May 2021) characterised by stringent public health restrictions and the dominance of the SARS-CoV-2 wild type and subsequently Alpha variant in the UK; Wave 3 (June 2021 to November 2021) characterised by relaxation of restrictions and the dominance of the Delta variant; or Wave 4 (December 2021 to April 2022) characterised by further relaxations of restrictions and the dominance of the Omicron variant. Test results were only available until 1 April 2022 due to the termination of the national testing programme in England affecting self-reported testing data and the termination of monthly serological testing in Virus Watch. Waves 1 and 2 were amalgamated into a single phase as it was not possible to attribute specific waves to serology tests conducted during Wave 2, and as mass population testing was largely introduced after the first pandemic wave in England and Wales. Both Waves 1 and 2 included periods of stringent public health restrictions, whilst Waves 3 and 4 occurred during the relaxation of public health measures in included regions, with a brief reintroduction of some limited restrictions in December 2021 and January 2022 due to the emergence of the Omicron variant. Some infections could not be attributed to a particular period as they were based on seropositivity without a prior seronegative result.

## Covariates

Where appropriate (see Statistical Analyses), models were adjusted for the following socio-demographic and health-related covariates: age (< 30, 30–39, 40–49, 50–59, 60 + years), sex at birth, binary vulnerability status (defined as any condition on the UK NHS/government list of clinically vulnerable conditions [22], obesity, and/or having received an NHS shielding letter), minority ethnicity (White British vs other), geographic region (ONS national region), deprivation based on English or Welsh Indices of Multiple Deprivation Quintile derived from postcode, annual household income (£0–24,900, £25,000–£49,999, £50,000–£75,000, and £75,000 +) and household size (excluding participant).

Models were adjusted for non-work public activities based on monthly surveys where participants reported the median number of days that they undertook the following activities across each survey week: using transport (using a bus, underground or overground train/

tram, taxi, or sharing a car with a non-household member), visiting essential shops, and leisure and social activities (attending the theatre, cinema, concert or sports event; eating in a restaurant, cafe or canteen; going to a bar, pub or club; going to a party; or non-essential shops or personal care services). Responses from November 2020 and February–April 2021 were allocated to Waves 1 and 2, with the second wave used to extrapolate to both early phases of the pandemic. Responses from May 2021–October 2021 were allocated to Wave 3, and from November 2021 – March 2022 to Wave 4. Monthly surveys were conducted towards the end of each month, so surveys conducted on the boundary months between pandemic waves were allocated to the subsequent wave.

### Statistical Analyses

To assess the influence of occupation on SARS-CoV-2 infection risk, we performed Poisson regression with robust standard errors, an established method to estimate risk ratios for binary outcomes [23]. Separate models were conducted for the full pandemic and by wave, with the reference category set as (1) ‘Other Professional and Associate occupations’ and (2) the full working population of Virus Watch excluding the occupational group under consideration. Other Professional and Associate occupations were selected as a reference group following similar criteria to previous studies of occupation and COVID-19 [4, 11], as this was the largest occupational group in Virus Watch, had a low absolute infection risk (see Supplementary Table S2) and was a non-frontline group with low prior estimates of exposure-relevant workplace factors [6, 8].

We identified potential confounders based on a purpose-developed directed acyclic graph (DAG—see ‘Directed Acyclic Graphs’ in Supplementary Materials), with models presented unadjusted and fully adjusted for the following potential confounders: age, sex, ethnicity, region, deprivation and household size, vulnerability status, and non-work public activities. While adjustment for non-work public activities was not required for minimally-sufficient adjustment according to our DAG, this was included in the final adjustment set. Evidence around the relationship between occupation and non-work public activities is lacking and any association is likely to involve complex inter-relationships with socio-economic position and other demographic factors; however, the potential impact of non-work public activities that may occur adjacent to work (e.g., transport use) has been previously discussed in the context of occupational infection risk [24] and we consequently included this in our final model given longitudinal availability of these data in Virus Watch. Additionally, deprivation and socio-economic position can be challenging to quantify and

adjustment for non-work public activities may mitigate against residual confounding via these pathways. Vaccination status was not directly included in the models due to being a determinant of vaccination status according to UK policy, and therefore on the causal pathway between occupation and infection risk. Additionally, other factors determining vaccination according to UK policy (i.e., age, health status) were included in the models and our DAG did not indicate any open confounding pathways (Supplementary Figs. 1a and b). No evidence of multicollinearity emerged based on variance inflation factors for any model. We performed a sensitivity analysis limited to participants who had undergone serological testing ( $n = 9114$ ) to address potential differential access and testing behaviour for virological/antigen testing across occupations; it was only possible to perform this analysis on broad occupational groups across the full study period, and not for specific occupations or by wave due to limited statistical power (see ‘Clinical Outcomes’ in the Supplementary Materials).

Based on the fully-adjusted models, we calculated attributable fractions for the exposed subpopulations (AFs) using the *punaf* programme in Stata Version 16 [25]. Attributable fractions range from -8 to 1, with negative values indicating a protective effect and positive values indicating a harmful effect [25]; while negative values are often transformed to express cases prevented in the unexposed group, we did not transform estimates in order to facilitate comparison by leaving all estimates with the same denominator.

Missing data were limited for all included sociodemographic variables (0–6%) and complete cases were included in the final analyses. We conducted a missing data sensitivity analysis by applying multivariate imputation by chained equations (*mice* package in R Version 4.0.3 [26]) with 5 datasets with 50 iterations per dataset to socio-demographic variables and re-testing models.

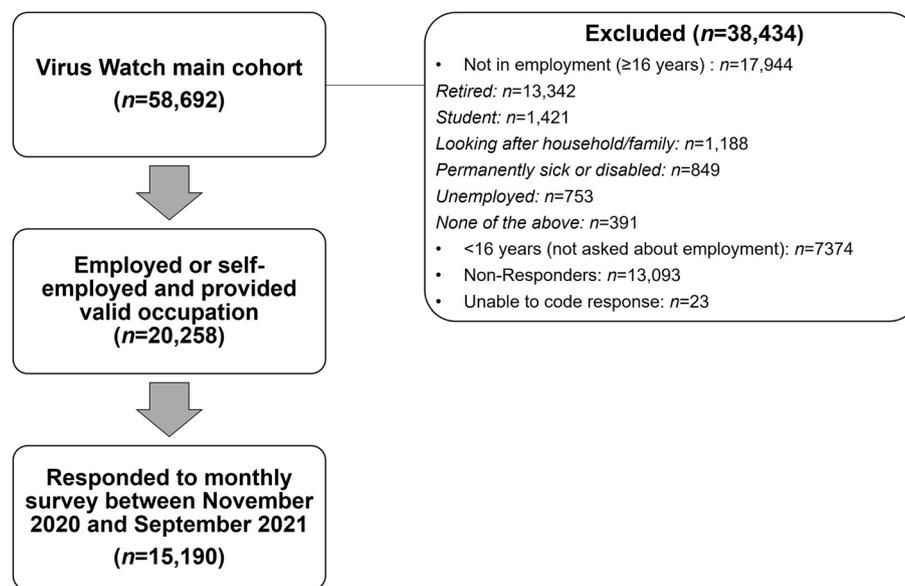
### Results

Participant selection is presented in Fig. 1, with demographic features of included participants ( $n = 15,190$ ) reported in Table 1.

#### Occupational group and infection risk

Absolute risk of infection by occupational risk is reported in Supplementary Table S2, and ranged from 26% in outdoor tradespeople to 42% in teaching, education and childcare workers across the full study period.

Risk ratios comparing each occupational group with ‘Other Professional and Associate’ occupations are illustrated in Fig. 2 for the full study period and stratified by wave. Across the full study period, healthcare (adjusted risk ratio (aRR) = 1.29, 1.18–1.40; attributable fraction



**Fig. 1** Flow diagram of participant eligibility

(AF) = 22%, 15–29%), leisure and personal service (aRR = 1.15, 1.03–1.29; AF = , 13%, 3–22%), social care and community protective service (aRR = 1.24, 1.12–1.38; AF = 19%, 11–27%), and teaching, education and childcare occupations (aRR = 1.34, 1.24–1.44; AF = 25%, 19–30%) demonstrated elevated infection risk compared to Other Professional and Associate Occupations (Fig. 2; see Supplementary Table S3 for adjusted AFs). When stratified by wave, between-occupational differences in infection risk were most prominent in Waves 1 and 2, and were attenuated in Waves 3 and 4 (Fig. 2). Most of the above groups demonstrated elevated risk in Waves 1 and 2 – along with indoor trades and process/plant occupations (aRR = 1.44, 1.18–1.76; AF = 31%, 15–43%) and sales and customer service occupations (aRR = 1.29, 1.02–1.94; AF = 22%, 2–39%). The only groups with elevated risk in later waves were teaching, education, and childcare workers – who had elevated risk across all waves – and healthcare workers who had elevated risk in Wave 4 (Fig. 2). Similar results were obtained in sensitivity analyses including only participants who underwent serological testing (Supplementary Fig. 2), and including imputed sociodemographic data (Supplementary Fig. 3a). Across all models, adjustment for sociodemographic, health-related and non-workplace activities had limited effects (Fig. 2).

Similar between-occupational trends were obtained when comparing each occupational group to the rest of the working population (Fig. 3 and Supplementary Table S2), with lower risk ratios and attributable fractions than those compared to Other Professional and Associate

occupations. Similar results were also observed in related sensitivity analyses comparing each occupation to the rest of the working population based on serological testing only (Supplementary Fig. 2) and including imputed sociodemographic data (Supplementary Fig. 3b).

### Specific Frontline Occupations

Absolute risk of infection for specific frontline occupations is reported in Supplementary Table S4; primary school teachers demonstrated the highest absolute risk (53%) across the full study period.

Risk ratios comparing frontline occupations with ‘Other Professional and Associate’ occupations are illustrated in Fig. 4 for the full study period and by wave. Across the full study period, the following groups demonstrated elevated risk: nurses (aRR = 1.44, 1.25–1.65; AF = 30%, 20–39%); doctors (aRR = 1.33, 1.08–1.65; AF = 25%, 7–39%); carers (1.45, 1.19–1.76; AF = 31%, 16–43%); primary school teachers (aRR = 1.67, 1.42–1.96; AF = 40%, 30–49%); secondary school teachers (aRR = 1.48, 1.26–1.72; AF = 32%, 21–42%); and teaching support occupations (aRR = 1.42, 1.23–1.64; AF = 29%, 18–39%) (Fig. 4; attributable fractions in Supplementary Table S5). When stratified by wave, all of these occupational groups demonstrated elevated risk during Waves 1 and 2 along with cleaners (aRR = 1.60, 1.01–2.51; AF = 37%, 1–60%); warehouse and process/plant workers (aRR = 1.93, 1.41–2.65; AF = 48%, 29–62%); and food preparation and hospitality workers (aRR = 1.82, 1.17–2.83; AF = 45%, 15–65%). The only occupations who demonstrated elevated risk in subsequent waves were carers (in

**Table 1** Characteristics of study participants

	<i>N</i> = 15,190 <sup>1</sup>
<b>Occupation</b>	
Administrative & secretarial	1,942 (13%)
Healthcare	1,272 (8.4%)
Indoor trades, process & plant	1,044 (6.9%)
Leisure & personal service	731 (4.8%)
Managers, directors & senior officials	1,249 (8.2%)
Other professional & associate	4,972 (33%)
Outdoor trades	372 (2.4%)
Sales & customer service	770 (5.1%)
Social care & community protective services	827 (5.4%)
Teaching, education & childcare	1,671 (11%)
Transport & mobile machine	340 (2.2%)
<b>Age</b>	
< 30	1,164 (7.7%)
30–39	2,244 (15%)
40–49	3,127 (21%)
50–59	4,524 (30%)
60 +	4,131 (27%)
<b>Sex</b>	
Female	8,479 (55.8%)
Male	6,686 (44%)
Unknown/Other <sup>2</sup>	25 (0.2%)
<b>Ethnicity</b>	
White British	12,574 (84%)
White Other	1,371 (9.1%)
South Asian	476 (3.2%)
Other Asian	142 (0.9%)
Black	133 (0.9%)
Mixed/Multiple Ethnicity	244 (1.6%)
Other Ethnicity	80 (0.5%)
Unknown <sup>2</sup>	170 (1.1%)
<b>Chronic Condition and/or Obesity</b>	
<b>Index of multiple deprivation quintile</b>	
1	1,493 (9.9%)
2	2,562 (17%)
3	3,085 (21%)
4	3,776 (25%)
5	4,103 (27%)
Unknown <sup>2</sup>	171 (1.1%)
<b>Household income</b>	
£0–£24,999	2,277 (17%)
£25,000–£49,999	4,414 (32%)
£50,000–£74,999	3,303 (24%)
£75,000 +	3,692 (27%)
Unknown <sup>2</sup>	1,504 (9.9%)
<b>Household size</b>	
1 person	3,272 (22%)
2 people	6,976 (46%)
3 people	2,329 (15%)

**Table 1** (continued)

	<i>N</i> = 15,190 <sup>1</sup>
4 people	2,020 (13%)
5 people	493 (3.2%)
6 people	100 (0.7%)
<b>Region</b>	
East Midlands	1,371 (9.0%)
East of England	2,922 (19%)
London	2,502 (16%)
North East	659 (4.3%)
North West	1,607 (11%)
South East	2,910 (19%)
South West	1,095 (7.2%)
Wales	390 (2.6%)
West Midlands	817 (5.4%)
Yorkshire and The Humber	746 (4.9%)
Unknown <sup>2</sup>	171 (1.1%)

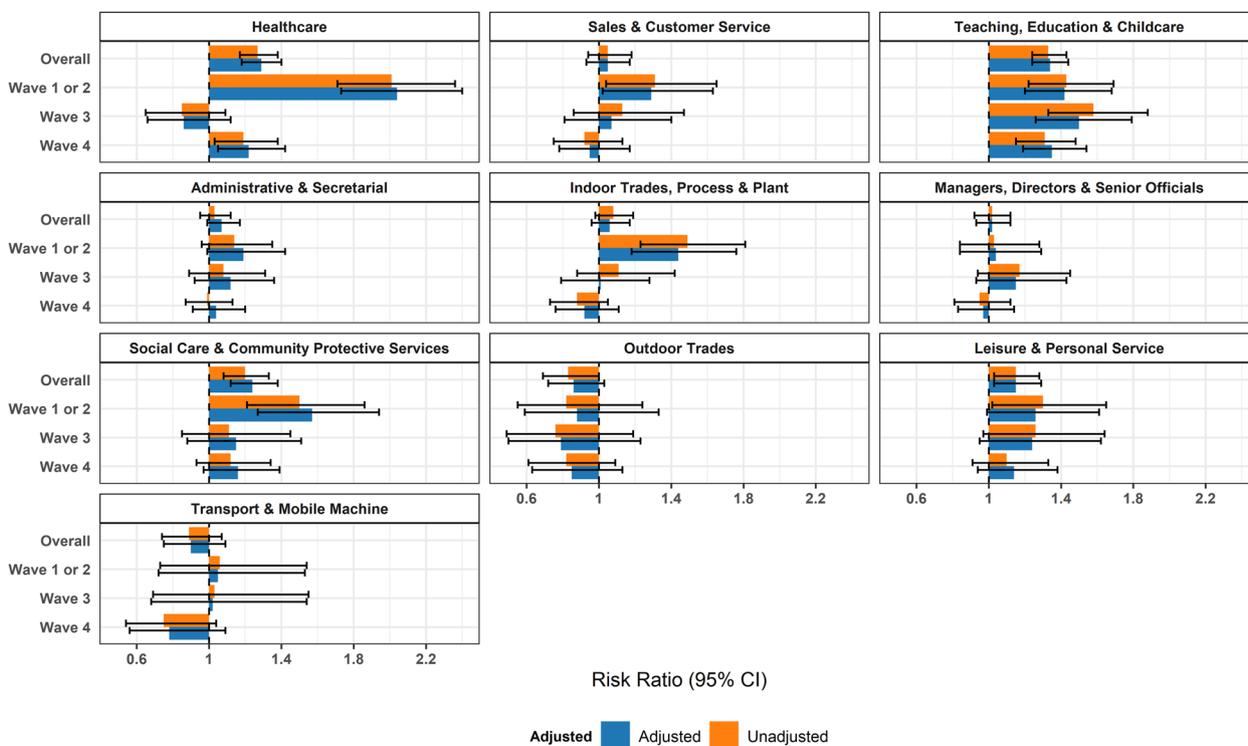
<sup>1</sup> n (%) of available; <sup>2</sup> n (%) of total

Wave 3), and primary and secondary school teachers and teaching support workers across all waves. Similar results were obtained when comparing each frontline occupation to the rest of the working population (Fig. 5 and Supplementary Table S4; imputed data sensitivity analysis in Supplementary Fig. 4b), with attenuated risk ratios and attributable fractions than when using Other Professional and Associate occupations as a comparator. Results were also consistent in sensitivity analyses using imputed sociodemographic data (Supplementary Fig. 4a).

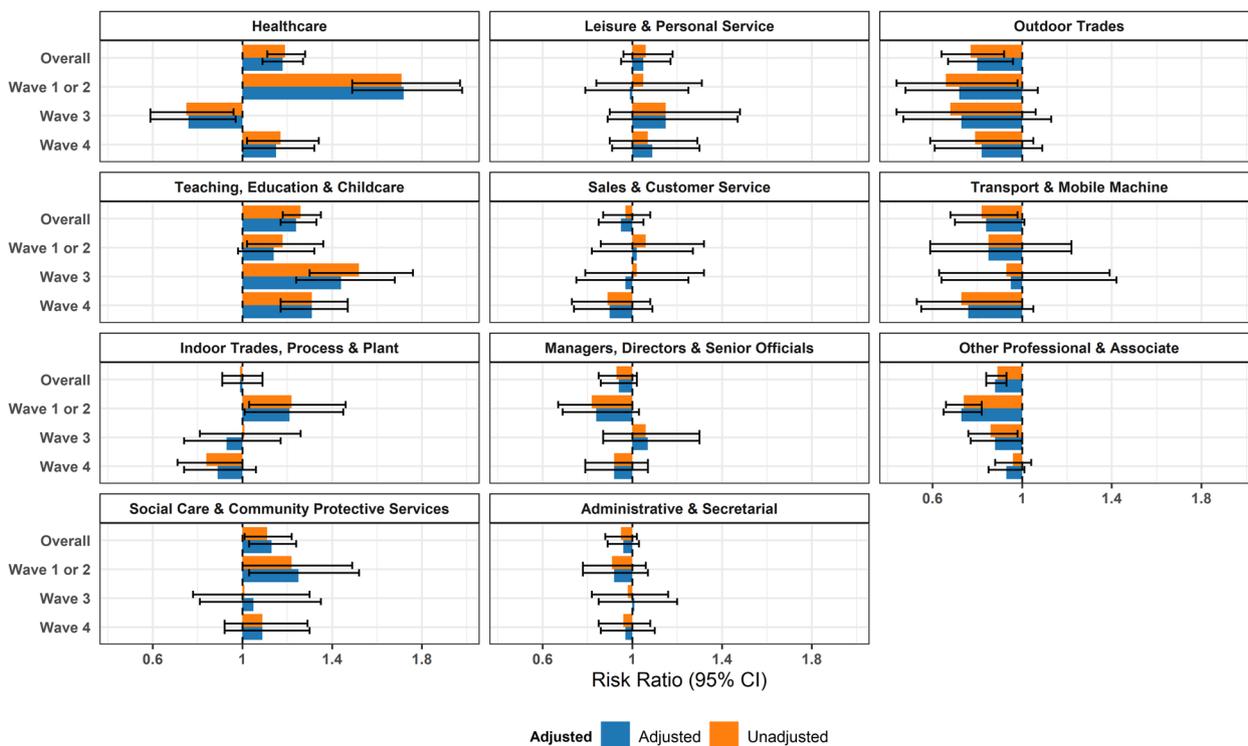
## Discussion

### Key Findings and Interpretation

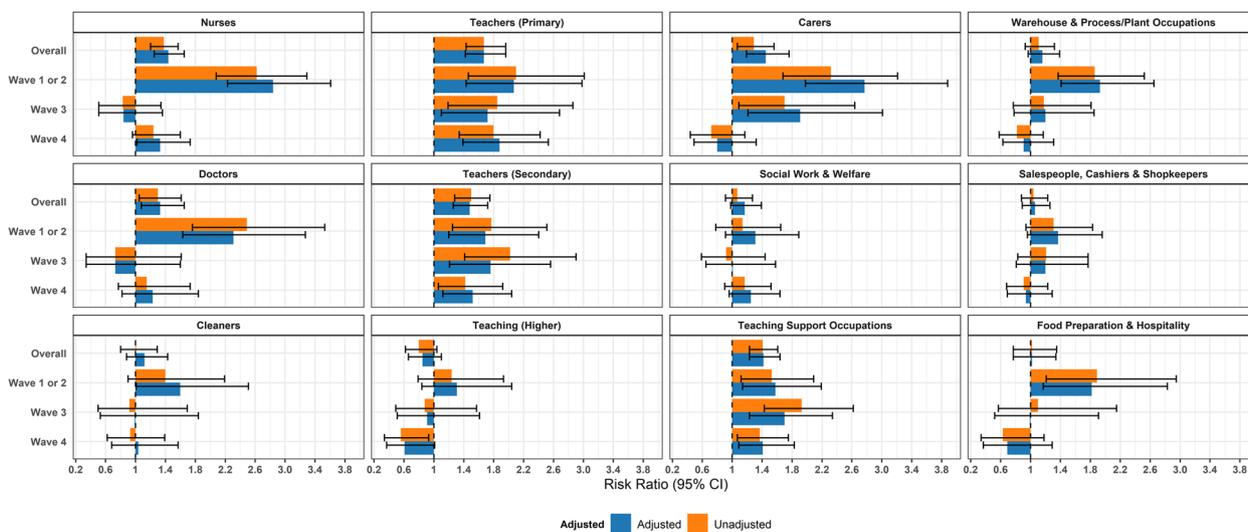
This study found persistent occupational differences in SARS-CoV-2 infection risk after comprehensive adjustment for non-work-related confounding, including socio-demographic and health-related factors and non-work social activities. Compared to Other Professional and Associate occupations—the largest occupational group in the sample with the lowest infection risk—workers in healthcare, teaching, education and childcare, social care and community protective services, and leisure and personal service occupations demonstrated elevated overall infection risk. In these groups, belonging to their occupation compared to the less risky group accounted for between 13% (for leisure and personal service workers) to 25% (for teaching, education and childcare workers) of their infection risk. Most at-risk occupations demonstrated elevated risk in the earlier pandemic phase which was later attenuated for most groups, with the exception of teaching, education and childcare occupations



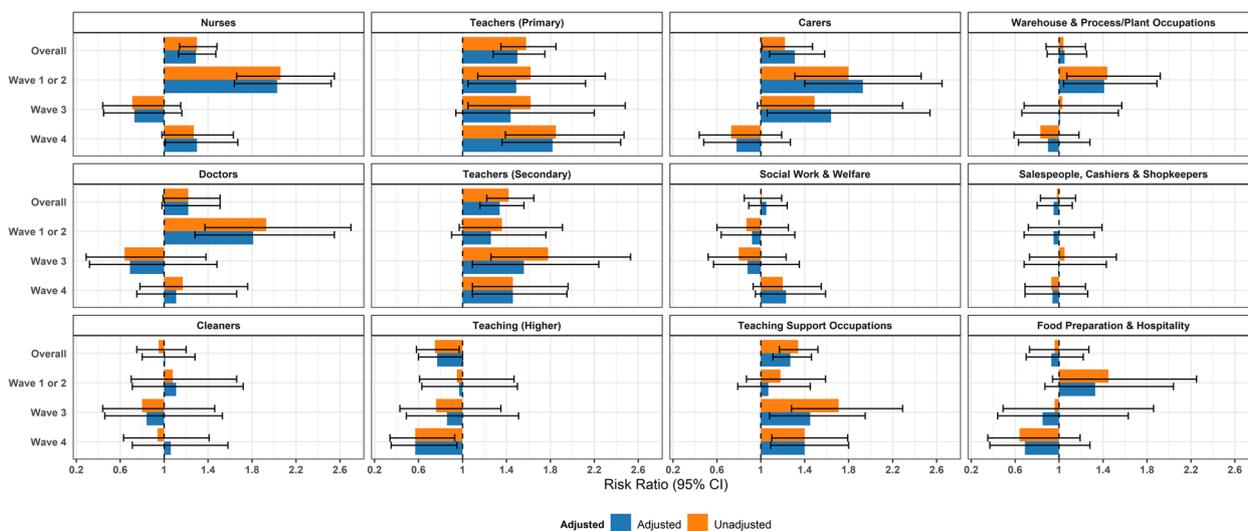
**Fig. 2** Risk Ratios by Occupational Group (versus Other Professional and Associate): Unadjusted and adjusted for age, sex, ethnicity, region, deprivation, household size, vulnerability status, and non-work public activities



**Fig. 3** Risk Ratios by Occupational Group (versus Working Population): Unadjusted and adjusted for age, sex, ethnicity, region, deprivation, household size, vulnerability status, and non-work public activities



**Fig. 4** Risk Ratios for Frontline Occupations (versus Other Professional and Associate): Unadjusted and adjusted for age, sex, ethnicity, region, deprivation, household size, vulnerability status, and non-work public activities



**Fig. 5** Risk Ratios for Frontline Occupations (versus Working Population): Unadjusted and adjusted for age, sex, ethnicity, region, deprivation, household size, vulnerability status, and non-work public activities

for whom risk remained elevated in Waves 3 and 4, and healthcare workers who also had elevated risk in Wave 4.

Where sample size was sufficient, we also investigated infection risk for specific frontline occupational groups. Occupational and temporal patterns mirrored those described above, with differences most apparent in earlier pandemic phases and persistent across all phases for teachers and teaching support workers. Substantial attributable fractions were identified for at-risk workers, with occupation accounting for between 25% (in doctors) to 40% (in primary school teachers) of workers’ risk

of infection. Findings may have been impacted by lack of power to detect modest effects in some groups.

Elevated infection risk in occupational groups with limited ability to work from home and those involving exposure to patients and/or the public echoes findings from the previous studies with more limited adjustment for potential confounding and from other global regions [1–5, 12, 13]. Across all analyses in the current study, adjustment for sociodemographic and health-related factors and non-work activities had limited impact on estimates. This result differs markedly to prior analysis

of occupational differences in COVID-19 mortality [11], in which adjustment for socio-demographic and health-related factors substantially reduced the effect of occupation. Occupation plausibly shapes SARS-CoV-2 exposure—and consequently infection risk—by influencing workers' ability to work from home, practise social distancing at work, work in well-ventilated environments, and access appropriate personal protective equipment. The specific mechanisms and relative contribution of different mitigating factors are likely to differ considerably by occupation, and are an important area for future research. Conversely, clinical factors that influence risk of severe morbidity and mortality once infected may differ across occupations, however the direct effect of occupation itself on severity of infection is likely to be more limited.

Changing patterns of differential infection risk by pandemic phase are likely to be multifactorial. Immunity-related factors that reduce the population of susceptible workers within a given occupation are likely to be important, and include prior infection in early phases of the pandemic, prioritization of some occupational groups (i.e. health and care workers [11, 27]) for vaccination, and potential differences in the speed and overall uptake of vaccination between occupations [28]. The removal of remaining public health restrictions in Wave 3 may also have reduced differential risk by increasing overall contact rates and networks, and probability of transmission outside of work due to the increasing range of potential venues for exposure at a time of persistently high community infection rates and reduced mitigations. Resurgent risk in healthcare workers, particularly nurses, in the fourth wave may reflect the impact of relaxed restrictions on some healthcare workers with intensive patient contact as well as the impact of the immune-evasive Omicron variant. Relatedly, persistently elevated risk in teaching and childcare occupations may reflect high-intensity workplace exposure in combination with high levels of infection in children [29, 30]. Direct investigation into potential mediators of this phase effect was beyond the scope of this study, and is warranted to better understand the processes shaping occupational infection risk. Relatedly, investigation into effective mitigation for the ongoing elevated infection risk in teachers is recommended both to address occupational inequalities and to reduce disruption in education settings.

### Strengths and Limitations

Strengths of this study include the large and diverse cohort that enabled investigation of infection risk from multiple study-derived and linked sources including both symptomatic testing and serology over multiple pandemic phases. Detailed information around participants'

demographic characteristics and activities over time allowed adjustment for a comprehensive series of potential confounders, including non-work-related public activities, informed by a directed acyclic graph.

However, the study has several important limitations. The Virus Watch cohort is demographically diverse but not representative of the UK population. The study population comprised larger numbers of workers in professional and administrative occupations relative to trades and leisure and personal service occupations. Power to detect effects may consequently have been reduced in the latter groups. Occupation was classified using officially-recommended, semi-automatic methods [19, 20]; consequently, any automatically-classified occupations that were marked as not having high certainty by the classification software were manually checked by a trained investigator to avoid misclassification [20]. Some misclassification may have remained due to undetected random error, and may have attenuated the magnitude of effects.

Potential confounders, such as deprivation, are challenging to measure and residual confounding cannot be excluded. Non-work public activities were inferred from self-reported activities across a given survey week, and may not have been an accurate reflection of participants' activity patterns across the entire relevant time period. Furthermore, social and leisure activities may have included work for some occupational groups (e.g. leisure and personal service occupations) but could not be disaggregated; however, the limited effect of adjustment in these models indicates that this was unlikely to be a major source of bias. Occupation was measured in broad categories, and only some specific occupations could be investigated due to small subsample sizes. Relatedly, the number of infections within a given pandemic phase was small for some frontline subsamples. Overall estimates of risk by occupational sector may be driven by particularly risky roles with considerable exposure [9], and further investigation into specific occupations is recommended. Additionally, inclusion of multiple test types to indicate SARS-CoV-2 positivity allowed for potential detection of asymptomatic or previously untested cases through serology, and detection of early cases through linkage. However, issues impacting the uptake and usage of each test type, including differential access to some tests in given phases of the pandemic, self-selection bias, and compliance with testing instructions may have affected estimates and are difficult to delineate. Notably, swab testing uptake may be influenced by differential testing behaviour between occupations. For example, health care workers undertake regular occupational testing which may lead to an overestimation of their relative risk of infection. However, a sensitivity analysis constrained to those

participants who underwent serological testing was not subject to such testing behaviour bias and demonstrated similar results to the main analyses.

## Conclusions

Despite these limitations, the present study indicates differential infection risk across occupational groups in England and Wales, with patterns of differential risk appearing to vary across pandemic waves. These findings illustrate the importance of work as a source of infection risk during the COVID-19 pandemic, with substantial fractions of infections attributable to occupation in at-risk groups. Occupations with persistently elevated risk (i.e., teachers) should be an ongoing target for interventions such as improved ventilation in schools, while understanding processes that shape differential risk in earlier phases of the pandemic is relevant for future outbreaks of respiratory infections. Investigation into the mechanisms underlying differential risk overall and over time, as suggested by this study, could inform evidence-based public health interventions in the workplace.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12995-023-00371-9>.

**Additional file 1: Supplementary Table 1.** UK Standard Occupational Classification 2020 (SOC-2020) Codes within Virus Watch Occupational Categories. Abbreviations: n.e.c. = not elsewhere classified; \* Limited to three most prevalent occupations per category to prevent declarative disclosure and due to large number of occupations across sample ( $n=402$ ). **Supplementary Figures 1a and 1b.** Directed Acyclic Graphs for Occupation – Infection Risk. Note: Supplementary Figure 1a and 1b illustrate how the current adjustment set remains sufficient under assumptions of health status influencing occupation or occupation influencing health status. Black arrows denote paths controlled by adjustment, while green arrows denote causal paths. **Supplementary Table 2.** Absolute Infection Risk by Occupational Group. \*Proportion of occupational category total. **Supplementary Table 3.** Attributable Fractions (Exposed) by Occupational Group. **Supplementary Figure 2.** Risk Ratios by Occupational Group – Serologically-Confirmed Infections vs No Evidence of Infection. **a.** Risk Ratios versus Other Professional and Associate Occupations. **b.** Risk Ratios versus Working Population. **Supplementary Figure 3a and 3b.** Risk Ratios by Occupational Group – Missing Data Sensitivity Analyses with Imputation. **a.** Risk Ratios versus Other Professional and Associate Occupations. **b.** Risk Ratios versus Working Population. **Supplementary Table 4.** Absolute Infection Risk for Frontline Occupations. \*Proportion of occupational category total. **Supplementary Table 5.** Attributable Fractions (Exposed) for Frontline Occupations. **Supplementary Figure 4a and 4b.** Risk Ratios for Frontline Occupations – Missing Data Sensitivity Analyses with Imputation. **a.** Risk Ratios versus Other Professional and Associate Occupations. **b.** Risk Ratios versus Working Population

## Acknowledgements

We acknowledge the following members of the Virus Watch study collaborative who provided theoretical and methodological input to the main Virus Watch cohort study from which this analysis was drawn: Susan Michie, Pia Hardelid, Linda Wijlaars, Eleni Nastouli, Moira Spyer, Ben Killingley, Ingemar Cox, Rachel A McKendry, Tao Cheng, Yunzhe Liu, Jo Gibbs, Richard Gilson, Alison Rodger.

## Authors' contributions

SB (conceptualisation, methodology, data curation, formal analysis, writing—original draft); SH (methodology, writing -review and editing); TB (investigation, writing -review and editing); WLEF (investigation, data curation, writing -review and editing); EF (supervision, methodology, writing -review and editing); CG (data curation, writing -review and editing); JK (project management, writing -review and editing); AN (data curation; writing -review and editing); VN (software, data curation, writing -review and editing); PP (investigation, writing -review and editing); AY (software, investigation, writing -review and editing); AJ (conceptualisation, methodology, writing -review and editing); MVT (conceptualisation, methodology, writing -review and editing); RA (conceptualisation, methodology, supervision, writing -review and editing); AH (conceptualisation, methodology, supervision, writing -review and editing). The author(s) read and approved the final manuscript.

## Funding

This work was supported by funding from the PROTECT COVID-19 National Core Study on transmission and environment, managed by the Health and Safety Executive on behalf of HM Government. The Virus Watch study is supported by the MRC Grant Ref: MC\_PC\_19070 awarded to UCL on 30 March 2020 and MRC Grant Ref: MR/V028375/1 awarded on 17 August 2020. The study also received \$15,000 of Facebook advertising credit to support a pilot social media recruitment campaign on 18th August 2020. This study was also supported by the Wellcome Trust through a Wellcome Clinical Research Career Development Fellowship to RA [206602]. SB and TB are supported by an MRC doctoral studentship (MR/N013867/1). The funders had no role in study design, data collection, analysis and interpretation, in the writing of this report, or in the decision to submit the paper for publication.

## Availability of data and materials

We aim to share aggregate data from this project on our website and via a "Findings so far" section on our website—<https://ucl-virus-watch.net/>. We also share some individual record level data on the Office of National Statistics Secure Research Service. In sharing the data we will work within the principles set out in the UKRI Guidance on best practice in the management of research data. Access to use of the data whilst research is being conducted will be managed by the Chief Investigators (ACH and RWA) in accordance with the principles set out in the UKRI guidance on best practice in the management of research data. We will put analysis code on publicly available repositories to enable their reuse.

## Declarations

### Ethics approval and consent to participate

Virus Watch was approved by the Hampstead NHS Health Research Authority Ethics Committee: 20/HRA/2320, and conformed to the ethical standards set out in the Declaration of Helsinki. All participants provided informed consent for all aspects of the study.

### Consent for Publication

Not applicable.

### Competing interests

AH serves on the UK New and Emerging Respiratory Virus Threats Advisory Group. AJ and AH are members of the COVID-19 transmission sub-group of the Scientific Advisory Group for Emergencies (SAGE). AJ is Chair of the UK Strategic Coordination of Health of the Public Research board and is a member of COVID National Core studies oversight group. The other authors declare no competing interests.

### Author details

<sup>1</sup>Centre for Public Health Data Science, Institute of Health Informatics, University College London, London NW1 2DA, UK. <sup>2</sup>Institute of Epidemiology and Health Care, University College London, London WC1E 7HB, UK. <sup>3</sup>Department of Infectious Disease Epidemiology, London, School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. <sup>4</sup>Institute for Global Health, University College London, London WC1N 1EH, UK. <sup>5</sup>Division of Population Health, Health Services Research & Primary Care, University of Manchester, Manchester M13 9NT, UK.

Received: 4 January 2023 Accepted: 21 March 2023  
Published online: 03 April 2023

## References

- Phuong Do D, Reanne F. U.S. frontline workers and COVID-19 inequities. *Prev Med.* 2021;153:106833.
- Fogh K, Strange JE, Scharff BF, et al. Testing Denmark: A Danish nationwide surveillance study of COVID-19. *medRxiv.* 2021;9(3):e0133021.
- Magnusson K, Nygård K, Methi F, Vold L, Telle K. Occupational risk of COVID-19 in the first versus second epidemic wave in Norway, 2020. *Eurosurveillance.* 2021;26:2001875.
- de Gier B, de Oliveira Bressane Lima P, van Gaalen RD, et al. Occupation- and age-associated risk of test positivity, the Netherlands, June to October 2020. *Eurosurveillance* 2020; 25: 2001884.
- Biarnés-Martínez M, Fábregas M, Coma E, et al. SARS-CoV-2 infection in occupational settings in Catalonia. *Rev Esp Salud Publica* 2022;96:e202205040.
- Zhang M. Estimation of differential occupational risk of COVID-19 by comparing risk factors with case data by occupational group. *Am J Ind Med.* 2021;64:39–47.
- Cabinet Office. COVID-19 and occupation: position paper 48. <https://www.gov.uk/government/publications/covid-19-and-occupation-iac-position-paper-48/covid-19-and-occupation-position-paper-48> (Accessed 23 Jul 2022).
- Office for National Statistics. Which occupations have the highest potential exposure to the coronavirus (COVID-19)? 2020; <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/whichoccupationshavethehighestpotentialexposuretothecoronaviruscovid19/2020-05-11> (Accessed 23 Jul, 2022).
- Cevik M, Baral SD. Networks of SARS-CoV-2 transmission. *Science.* 2021;373:162–3.
- Pearce N, Rhodes S, Stocking K, et al. Occupational differences in COVID-19 incidence, severity, and mortality in the United Kingdom: Available data and framework for analyses. *Wellcome Open Res.* 2021;6:102. <https://doi.org/10.12688/wellcomeopenres.16729.1>.
- Nafilyan V, Pawelek P, Ayoubkhani D, et al. Occupation and COVID-19 mortality in England: a national linked data study of 14.3 million adults. *medRxiv.* 2021;79(7):433–41.
- Reuter M, Rigó M, Formazin M, et al. Occupation and SARS-CoV-2 infection risk among 108 960 workers during the first pandemic wave in Germany. *Scand J Work Environ Health* 2022; published online June 7. DOI:<https://doi.org/10.5271/sjweh.4037>.
- Nwaru CA, Santosa A, Franzén S, Nyberg F. Occupation and COVID-19 diagnosis, hospitalisation and ICU admission among foreign-born and Swedish-born employees: a register-based study. *J Epidemiol Community Health.* 2022;76:440–7.
- Sutherland E, Willis, Z, & Fordham E. Coronavirus (COVID-19) Infection Survey - Office for National Statistics. 2021; published online Feb 22. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsinthecommunityinengland/characteristicsofpeopletestingpositiveforcovid19inengland22february2021> (Accessed 23 Jul, 2022).
- Health and Safety Executive. Technical summary of data - Coronavirus (COVID-19) disease reports made by employers 10 April 2020. <https://www.hse.gov.uk/statistics/coronavirus/april-to-july-2020-technical-summary-of-data.htm> (Accessed 23 Jul 2022).
- Agius R. Disease and death from work: RIDDOR and covid-19. *Occup Med DOI:*<https://doi.org/10.1093/occmed/kqaa155>.
- Riley S, Eales O, Walters CE, et al. REACT-1 round 7 interim report: fall in prevalence of swab-positivity in England during national lockdown. *medRxiv.* 2020;2020:2020-11. <https://doi.org/10.1101/2020.11.30.20239806>.
- Hayward A, Fragaszy E, Kovar J, et al. Risk factors, symptom reporting, healthcare-seeking behaviour and adherence to public health guidance: protocol for Virus Watch, a prospective community cohort study. *BMJ Open.* 2021;11(6):e048042. <https://doi.org/10.1136/bmjopen-2020-048042>.
- Office for National Statistics. SOC 2020. <https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassifications/soc2020> (accessed Jul 23, 2022).
- Cascot: Computer Assisted Structured Coding Tool. <https://warwick.ac.uk/fac/soc/ier/software/cascot/> (accessed Jul 23, 2022).
- UK Health Security Agency. Essential workers prioritised for COVID-19 testing; 2021; published online Feb 25. <https://www.gov.uk/guidance/essential-workers-prioritised-for-covid-19-testing> (Accessed 23 Jul 2022).
- UK National Health Service. Who is at high risk from coronavirus (COVID-19). <https://www.nhs.uk/conditions/coronavirus-covid-19/people-at-higher-risk/who-is-at-high-risk-from-coronavirus/> (Accessed 23 Jul 2022).
- Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol.* 2004;159(7):702–6. <https://doi.org/10.1093/aje/kwh090>.
- European Centre for Disease Prevention and Control. COVID-19 clusters and outbreaks in occupational settings in the EU/EEA and the UK. <https://www.ecdc.europa.eu/en/publications-data/covid-19-clusters-and-outbreaks-occupational-settings-eueea-and-uk> (Accessed 10 Feb 2023).
- Newton R. PUNAF: Stata module to compute population attributable fractions for cohort studies. *Statistical Software Components* 2015; published online Sept 4. <https://ideas.repec.org/c/boc/bocode/s457193.html> (Accessed 23 Jul 2022).
- Van Buuren, S. Package 'mice': Multivariate Imputation by Chained Equations. <https://cran.r-project.org/web/packages/mice/mice.pdf> (Accessed 23 Jul 2022).
- Public Health England. COVID-19 vaccination first phase priority groups. 2020; published online Dec 7. <https://www.gov.uk/government/publications/covid-19-vaccination-care-home-and-healthcare-settings-posters/covid-19-vaccination-first-phase-priority-groups> (Accessed 23 Jul 2022).
- Nafilyan V, Dolby T, Finning K, et al. Differences in COVID-19 vaccination coverage by occupation in England: a national linked data study. *medRxiv.* 2021;2021.11.10.21266124. <https://doi.org/10.1101/2021.11.10.21266124>.
- Oeser C, Whitaker H, Linley E, Borrow R, Tonge S, Brown CS, Gower C, Warrener L, Brown KE, Ramsay M, Amirthalingam G. Large increases in SARS-CoV-2 seropositivity in children in England: Effects of the delta wave and vaccination. *J Infect.* 2022;84(3):418–67.
- Office for National Statistics. COVID-19 Schools Infection Survey, England. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/covid19schoolsinfectedsurveyengland/pupilantibodiesandvaccinesentimentmarch2022> (Accessed 22 Sep 2022).

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

