OPEN ACCESS

Original research

Identifying employee, workplace and population characteristics associated with COVID-19 outbreaks in the workplace: a population-based study

Christopher E Overton ,[1,2] Rachel Abbey,[1] Tarrion Baird,[1,3] Rachel Christie,[1] Owen Daniel,[4] Julie Day,[1] Matthew Gittins ,[5] Owen Jones,[1] Robert Paton,[1] Maria Tang,[1] Tom Ward,[1] Jack Wilkinson ,[5] Camilla Woodrow-Hill,[1,6] Timothy Aldridge,[7] Yiqun Chen [8]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/oemed-2023-109032).

[1]UK Health Security Agency, London, UK
[2]Department of Mathematical Sciences, University of Liverpool, Liverpool, UK
[3]Department of Pathology, University of Cambridge, Cambridge, UK
[4]Office for National Statistics, Newport, UK
[5]Centre for Biostatistics, The University of Manchester, Manchester, UK
[6]University of Manchester, Manchester, UK
[7]Health and Safety Executive, Buxton, UK
[8]Science Division, Health and Safety Executive, Buxton, UK

Correspondence to
Dr Christopher E Overton, Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 7ZL, UK; c.overton@liverpool.ac.uk

Received 31 May 2023
Accepted 15 November 2023
Published Online First 8 January 2024

To cite: Overton CE, Abbey R, Baird T, *et al*. *Occup Environ Med* 2024;**81**:92–100.

ABSTRACT

Objectives To identify risk factors that contribute to outbreaks of COVID-19 in the workplace and quantify their effect on outbreak risk.

Methods We identified outbreaks of COVID-19 cases in the workplace and investigated the characteristics of the individuals, the workplaces, the areas they work and the mode of commute to work, through data linkages based on Middle Layer Super Output Areas in England between 20 June 2021 and 20 February 2022. We estimated population-level associations between potential risk factors and workplace outbreaks, adjusting for plausible confounders identified using a directed acyclic graph.

Results For most industries, increased physical proximity in the workplace was associated with increased risk of COVID-19 outbreaks, while increased vaccination was associated with reduced risk. Employee demographic risk factors varied across industry, but for the majority of industries, a higher proportion of black/African/Caribbean ethnicities and living in deprived areas, was associated with increased outbreak risk. A higher proportion of employees in the 60–64 age group was associated with reduced outbreak risk. There were significant associations between gender, work commute modes and staff contract type with outbreak risk, but these were highly variable across industries.

Conclusions This study has used novel national data linkages to identify potential risk factors of workplace COVID-19 outbreaks, including possible protective effects of vaccination and increased physical distance at work. The same methodological approach can be applied to wider occupational and environmental health research.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ There are many risk factors that affect severity of disease caused by SARS-CoV-2 infection for individual infections. However, despite frequent reports of COVID-19 outbreaks in the workplace, there is limited evidence on factors that contribute to the risk of SARS-CoV-2 transmission in the workplace, with existing studies focusing on high-resolution analysis of individual workplace outbreaks with small sample sizes.

WHAT THIS STUDY ADDS

⇒ This is the first study to use large national data linkages to consider the association between where workers lived, worked and modes of commute and the risk of COVID-19 outbreaks in the workplace. Increased levels of vaccination and reduced physical proximity are shown to reduce the risk of workplace outbreaks. Increased levels of deprivation in the areas where workers lived showed increased risk of workplace outbreaks.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The findings contribute to the evidence base around social distancing measures in workplaces and vaccination at reducing transmission/infection. By identifying other risk factors associated with workplace outbreaks, this can inform workplace health and safety messaging. The innovative data linkages will support future workplace health research.

INTRODUCTION

Throughout the COVID-19 pandemic, outbreaks of cases in workplaces have frequently been reported worldwide.[1–3] Some industries have seen elevated rates of outbreaks, including healthcare settings, where the risk of exposure to patients with COVID-19 is high, and essential services, where work is unable to be carried out from home. The impact of workplace outbreaks is twofold. First, workplace outbreaks contribute to the prevalence of COVID-19 in the community. This is evident in the role workplace closures played in reducing transmission.[4] The second impact is on the workforce. Workplace outbreaks result in potentially large proportions of the workforce being unable to work effectively,[5] which can have financial impacts on the company, the economy[6] and on workers themselves.[7]

The frequency and size of outbreaks vary by time and space, depending on the community prevalence of COVID-19. This is the main exposure contributing to outbreaks in the workplace, but there are other potential exposures that could amplify the

risk of outbreaks occurring for given levels of COVID-19 prevalence. Demographic factors that have been associated with risk of infection include age, gender, ethnicity and deprivation.[8] Workplace factors that are likely to influence outbreak risk include physical proximity of workers, physical environment, including ventilation in the workplace, work activities, contract stability of the workforce and characteristics of the local area.

Workplaces present opportunities for a range of infectious diseases to spread, including respiratory syncytial virus,[9] influenza and norovirus.[10] Infectious diseases can be introduced through providing services to infectious individuals or workers attending workplaces while infectious. However, to study the influence of risk factors on potential transmission requires disease surveillance data at levels of detail previously unavailable. For COVID-19, however, mass testing of both the community and the workforce has provided a proxy measure for prevalence of infection in the community and the rate of outbreaks in workplaces, providing a unique opportunity to address questions on the risk factors associated to workplace outbreaks of COVID-19.

To mitigate the workplace impact of future waves of COVID-19, and to ensure accurate and reliable health and safety messaging can be issued to workplaces, it is essential to understand the different risk factors associated with these outbreaks. By using a combination of employee, workplace and population characteristics, our objective was to identify risk factors associated with workplace outbreaks of COVID-19 in different industry sectors. The findings of the study aim to support future pandemic preparedness.

METHODS

Setting

This study examined variation in workplace COVID-19 outbreak rates by Middle Layer Super Output Area (MSOA—geographical areas with resident populations between 5000 and 15 000 persons[11]) in England according to workplace characteristics, the characteristics of the local population of the workers, and the neighbourhood characteristics around the workplace. We considered three time periods for the study: 20 June 2021–30 September 2021 (Delta variant dominant strain), 7 December 2021–20 February 2022 (Omicron variant dominant strain),[12] and 20 June 2021–20 February 2022 (ie, overall, encompassing both periods and the intervening period). The period between Delta and Omicron was not assessed independently due to a mixture of variants circulating. These study periods were considered since testing policy was relatively stable.

Data sources

We used population data to investigate potential risk factors for workplace outbreaks by sector. The data describe demographics of employees and residents in England; characteristics of workplaces in England and COVID-19 testing and vaccination data.

Demographic characteristics were obtained from the 2011 Census conducted by the ONS.[13] The variables extracted from the census were age, gender, ethnicity and modes of commute to work. We also used data on indices of multiple deprivation (IMD) from the Department for Levelling Up, Housing and Communities.[14]

Workplace COVID-19 clusters were identified from the UKHSA Contact Tracing and Advisory Service (CTAS) data.[15] These data collected information on individuals who tested positive for COVID-19 in England and were contacted by the contact tracing service. For individuals who tested positive (and attended their workplaces 3–7 days prior to their positive test), their places of work were coded using the unique property reference numbers. This enabled workplace clusters to be counted based on two or more cases of self-reported COVID-19 occurring at the same workplace building within a rolling 6-day period, using the method described in Abbey *et al*.[16] We use clusters as a proxy measure for workplace outbreaks in this analysis.

Mass testing data from UKHSA was used to measure community case-incidence over time.[17] During the study period, 20 June 2021–20 February 2022, testing policy was constant across England. Therefore, infection incidence (the number of new infections each day) is assumed to be proportional to case incidence (the number of positive tests each day) across this period, allowing case incidence to be used as a proxy for the force of infection in the community. The National Immunisation Management System data from UKHSA, which contains COVID-19 vaccination records,[18] provides data on all vaccination doses given in England, from which we calculated the proportion of working age individuals vaccinated.

The National Population Database (NPD)[19] workplace layer is derived from the Inter-Departmental Business Register[20] and includes estimates of the number and location of workplaces and workers in Great Britain by industry type using the Standard Industrial Classification (SIC). The number of workplaces in each Workplace Zone (a geographical area that is nested within an MSOA) was calculated by industry type, and subsequently aggregated to MSOA level. Workplaces with identified COVID-19 outbreaks were linked to the NPD workplace layer using business name, address and postcode information[16] so that an SIC code could be applied. Workplace information was aggregated into 34 COVID-19-related sector groups,[21] which were then aggregated into 11 broad industry sectors. Of the total number of workplace outbreaks detected, 42% were excluded since these could not be reliably linked to a specific industry sector when merging CTAS data with NPD data.

Other workplace data used in this study included data on physical proximity (mean proximity score) in the workplace at the two-digit SIC grouping level,[22] the proportion of workers on permanent contracts at the two-digit SIC grouping level,[23] and the classification of each workplace MSOA by mobility class,[24] which provides a refined definition of rural/urban classification.

Data preparation

Details of data preparation are contained in online supplemental material. The variables used in this study fall into five broad categories: workplace, COVID-19 case and outbreak rates, employee (characteristics of the employees, restricted to those who travel into workplaces), work commute modes and neighbourhood (characteristics of people who live near the workplace). Online supplemental table S1 lists the variables and groupings.

Statistical methods

Our outcome of interest was the workplace COVID-19 outbreak rate in the MSOA, which we defined as the number of active outbreaks in workplaces in an MSOA divided by the total number of workplaces in that MSOA. Since the number of active outbreaks is an integer valued count, these can be considered as samples from a counting process, such as a negative binomial distribution. Therefore, instead of modelling the outbreak rate directly, we modelled the time-varying number of active outbreaks in the workplace as the outcome variable of a negative binomial generalised linear model, and introduced the log transformed number of workplaces as an offset to transform this into a rate.

Table 1 Descriptive summaries of the data used in the study

Variable	Value by industry				
	Services	Education	Transport, distribution and warehousing	Manufacturing	Human health and social work
Workplace characteristics					
Total no of workplaces included	800 710	24 636	77 771	68 437	68 449
Mean physical proximity score in the workplace	58	–	59	54	–
Mean proportion of workers on permanent contracts	94%	–	96%	96%	–
Most common mobility class	Exurban	Exurban	Exurban	Suburban	Exurban
COVID-19 case and outbreak rates					
Mean no of active outbreaks	0.54	0.48	0.25	0.28	0.24
Mean log (employee_case_rate_lag_7 day)	−7.61	−7.70	−7.61	−7.60	−7.61
Mean log (outbreak_rate_other)	−16.02	−16.93	−13.16	−13.84	−15.07
Mean log (outbreak_rate_lag_7 day)	−19.24	−17.64	−21.11	−20.46	−21.04
Employee characteristics					
Mean proportion of employees with 2 or more vaccination doses	78.3%	79.1%	78.2%	78.6%	78.4%
Mean proportion of employees aged 18–29	27.4%	20.5%	23.0%	21.5%	21.8%
Mean proportion of employees aged 30–44	36.9%	36.3%	37.5%	36.0%	36.0%
Mean proportion of employees aged 45–59	30.1%	36.2%	33.3%	35.5%	35.4%
Mean proportion of employees aged 60–64	5.6%	7.0%	6.3%	7.1%	6.8%
Mean proportion of employees with an asian ethnicity	7.7%	6.1%	7.2%	5.4%	7.2%
Mean proportion of employees with a black/African/Caribbean ethnicity	3.4%	2.8%	2.8%	1.7%	3.4%
Mean proportion of employees with a mixed/multiple/other ethnicity	2.9%	2.2%	2.3%	1.6%	2.3%
Mean proportion of employees with a white ethnicity	86.0%	89.0%	87.7%	91.3%	87.1%
Mean proportion of employees using bus/metro/tram	17.4%	10.2%	10.6%	6.6%	11.6%
Mean proportion of employees using taxi/vehicle passenger	5.4%	5.7%	6.3%	7.2%	5.9%
Mean proportion of employees using other transport	0.4%	0.5%	0.4%	0.4%	0.5%
Mean proportion of employees using train	9.1%	4.4%	5.3%	2.4%	4.4%
Mean proportion of employees using single occupancy	67.6%	79.2%	77.4%	83.4%	77.6%
Mean proportion of employees who identify as female	48.0%	52.8%	36.8%	34.4%	54.4%
Mean proportion of employees who identify as male	52.0%	47.2%	63.2%	65.6%	45.7%
Most common IMD quintile among employees	3	3	3	3	3
Neighbourhood characteristics					
Mean proportion of neighbourhood residents with two or more vaccination doses	77.9%	78.4%	76.5%	77.7%	77.6%
Most common IMD quintile among neighbourhood residents	3	3	2	3	3

Services; education; transport, distribution and warehousing; manufacturing; human health and social work. Proportions are recorded as percentages. Dashes indicate variables for which insufficient data are available. Neighbourhood characteristics describe characteristics of individuals in the location of the workplace.
IMD, Indices of Multiple Deprivation.

The risk factors that we considered fall into three distinct groups: workplace, employee and work commute characteristics. These risk factors were selected by health and safety experts for examination in the current study. The full list of risk factors examined is shown in online supplemental table S1, alongside common confounder variables.

For each risk factor, we fitted three models. The first included only the risk factor (model 1: unadjusted). The second extended model 1 by including the neighbourhood variables (characteristics of individuals living near the workplace) and the log-transformed 7-day lagged: values of COVID rates among employees, outbreak rate in that industry and outbreak rate in other industries (model 2: minimal adjusted). The third extended model 2 by adding variables from other risk factor groups which are plausible causal antecedents of the risk factor of interest, that is, they do not occur later down the causal pathway (model 3: priority risk factor-fully adjusted). Our confounder sets were decided on by discussion among the author team, which included consideration of plausible directed acyclic graphs (DAGs) representing the relationships between the study variables (online supplemental figure S1). Acknowledging the complexity of the underlying process and the challenge of identifying a single DAG, we selected confounder sets which would be compatible with the principles described by VanderWeele.[25] Further details on the statistical models are provided in online supplemental material.

Note that models 1 and 2 included the same explanatory variables for all risk factors. Model 3 used different explanatory variables depending on which group of risk factors the factor of interest falls within. When measuring the effect of vaccination, we considered the periods where the Delta and Omicron variants were dominant in the population independently, since different COVID-19 variants have different levels of vaccine effectiveness.[26]

Each industry was modelled independently, because different testing policies across industry sectors lead to different case ascertainment rates. Therefore, outbreaks are not directly comparable across industry sectors. For education, we removed school holiday periods, excluding an extra week after the holiday period to account for delays from infection to cases being detected. This resulted in outbreaks between 18 July 2021 and 12 September 2021, 24/ October 2021 and 31 October 2021, 19 December 2021 and 9 January 2022 being excluded for education.

Table 2 Descriptive summaries of the data used in the study

Variable	Value by industry					
	Utilities	Mining and quarrying	Public service activities	Construction	Waste management and remediation	Agriculture, forestry and fishing
Workplace characteristics						
Total no of workplaces included	3624	108	7102	83 224	1202	13 744
Mean physical proximity score in the workplace	53	–	–	61	–	–
Mean proportion of workers on permanent contracts	98%	–	–	97%	–	–
Most common mobility class	Exurban	Exurban	Exurban	Exurban	Suburban	Suburban
COVID-19 case and outbreak rates						
Mean number of active outbreaks	0.14	0.11	0.27	0.11	0.11	0.09
Mean log(employee_case_rate_lag_7 day)	−7.60	−7.62	−7.61	−7.61	−7.60	−7.65
Mean log (outbreak_rate_other)	−11.25	−14.55	−12.32	−12.74	−11.90	−14.82
Mean log (outbreak_rate_lag_7 day)	−22.27	−22.70	−19.89	−23.23	−22.82	−23.50
Employee characteristics						
Mean proportion of employees with 2 or more vaccination doses	78.5%	80.8%	78.2%	78.7%	78.5%	81.2%
Mean proportion of employees aged 18–29	25.1%	21.3%	20.2%	22.2%	21.4%	19.4%
Mean proportion of employees aged 30–44	41.2%	36.2%	37.4%	36.6%	35.6%	31.8%
Mean proportion of employees aged 45–59	29.1%	34.5%	36.4%	34.3%	35.8%	39.0%
Mean proportion of employees aged 60–64	4.6%	8.1%	6.1%	6.9%	7.2%	9.8%
Mean proportion of employees with an asian ethnicity	8.2%	3.9%	5.3%	5.7%	4.0%	1.2%
Mean proportion of employees with a black/African/Caribbean ethnicity	3.6%	1.6%	2.8%	2.5%	1.8%	0.4%
Mean proportion of employees with a mixed/multiple/other ethnicity	3.0%	1.7%	1.9%	2.0%	1.5%	0.8%
Mean proportion of employees with a white ethnicity	85.2%	92.7%	90.0%	89.8%	92.8%	97.6%
Mean proportion of employees using bus/metro/tram	21.5%	11.5%	11.4%	9.7%	5.2%	2.3%
Mean proportion of employees using taxi/vehicle passenger	4.0%	5.4%	5.2%	6.1%	7.3%	7.4%
Mean proportion of employees using other transport	0.3%	0.6%	0.4%	0.5%	0.4%	0.8%
Mean proportion of employees using train	18.0%	8.2%	6.6%	5.1%	1.7%	0.8%
Mean proportion of employees using single occupancy	56.3%	74.3%	76.4%	78.7%	85.5%	88.7%
Mean proportion of employees who identify as female	39.4%	35.6%	50.5%	41.0%	30.8%	40.9%
Mean proportion of employees who identify as male	60.6%	64.4%	49.5%	59.0%	69.3%	59.1%
Most common IMD quintile among employees	3	3	3	3	3	3
Neighbourhood characteristics						
Mean proportion of neighbourhood residents with two or more vaccination doses	75.4%	82.9%	73.5%	77.3%	76.3%	84.5%
Most common IMD quintile among neighbourhood residents	2	3	2	3	2	3

Utilities; mining and quarrying; public service activities; construction; waste management and remediation; agriculture, forestry and fishing. Proportions are recorded as percentages. Dashes indicate variables for which insufficient data are available. Neighbourhood characteristics describe characteristics of individuals in the location of the workplace.
IMD, Indices of Multiple Deprivation.

RESULTS

Descriptive results

Across the study period, 1149 007 workplaces were included. The study period contains 36 weeks, and we have 6721 distinct MSOAs and 11 industry sectors. Altogether, there were 801 874 data points, where each data point corresponds to a unique combination of MSOA, week and industry. Note that not all permutations are included, since for some MSOAs not all industry sectors are present.

Tables 1 and 2 show a descriptive summary of the data used in this study by industry. The mean is based on the number of workplaces, rather than the number of employees. Figure 1 shows the outbreak rate at a national level over time.

Main results

Tables 3 and 4 show the results for the priority risk factor fully adjusted model. Online supplemental tables S2 and S3 show the results for the unadjusted and minimal adjusted models for all sectors, respectively. Table 3 contains results for the five industry sectors identified as high importance by HSE: education; human health and social work activities; services; transport, distribution and warehousing; and manufacturing, with the remaining six industry sectors in table 4. Comparing the unadjusted, minimal adjusted and the priority risk factor fully adjusted model, some of the results are substantially different, demonstrating that the identified causal structure has a large influence.

Workplace characteristics

Across all industries with sufficient data available, increasing physical proximity in the workplace was associated with increased risk of outbreaks, except for construction where the effect was not statistically significant (table 4). Increasing the proportion of workers on permanent contracts was associated with reduced risk in all industries except construction, where it was associated with an increase in risk (table 4). The relationship with mobility class varied across the industries but was generally found to be a significant risk factor.

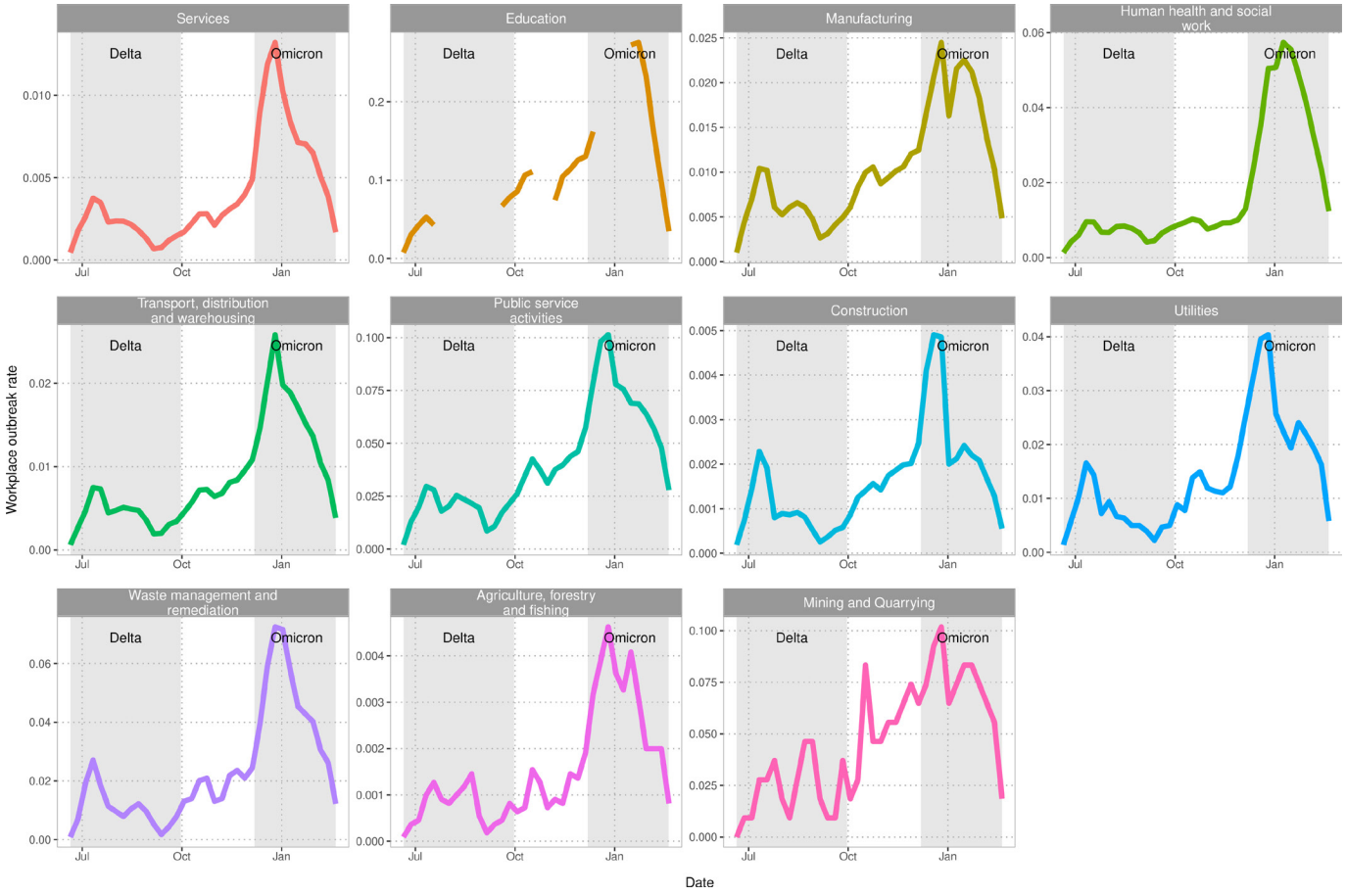

Figure 1 Workplace outbreak rates (total number of active outbreaks in the workplace divided by total number of workplaces) over time, broken down by industry sector. The shaded regions mark the Delta and Omicron periods. The full time series is the overall study period. Note that each panel has a different y-axis. Comparison between sectors should be avoided since each sector has a different testing policy.

Employee vaccination

Increasing the cumulative proportion of employees with two or more doses of vaccination, during the Delta and Omicron periods, was associated with reduced risk (or no significant effect) for all industries except education (table 3). In education, vaccination was associated with increased risk during Delta and no significant effect during Omicron.

Demographic characteristics

Across every industry, having higher proportions of individuals in the 60–64 age group, in comparison to the baseline group (30–44), was associated with substantially reduced risk of outbreaks (or no significant effect). In all industries, a higher proportion of individuals in either the 18–29 age group or the 44–59 age group, relative to the baseline group, was associated with an elevated risk. The presence of higher proportions of individuals from black/African/Caribbean ethnicities, relative to the baseline group (white ethnicity), were associated with increased outbreak risk (or no significant effect) in all industries. Mixed/multiple/other were associated with reduced risk, relative to the baseline group, in all industries except public service activities and agriculture, forestry and fishing (table 4). The trend across the remaining groups of ethnicities varied across industries. A higher proportion of female workers was associated with reduced risk (or no significant effect) in all industries except for education, transport, distribution and warehousing, and waste management and remediation. In all industries, increasing the IMD quintile (reducing deprivation) was associated with reduced risk of workplace outbreaks, except for mining and quarrying and agriculture, forestry and fishing, where the effect was not significant (table 4).

Modes of commute to work

Work commute modes generally had a significant association with the risk of workplace outbreaks across all industries, though the association varied substantially across industries. Taxi (or vehicle passenger) was associated with increased risk (or no significant effect) across all industries except education, utilities and public service activities (tables 3 and 4). Trains were associated with reduced risk (or no significant effect) in all industries. The results for bus, tram or metro use varied.

DISCUSSION

Key results

We set out to identify potential risk factors associated with workplace outbreaks of COVID-19 in different industry settings. We found increased physical proximity in the workplace was associated with increased risk of workplace outbreaks in most industries. Airborne transmission of COVID-19[27] will be highly dependent on physical proximity, which likely contributes to the reduced outbreak risk in industries with lower physical proximity. Other workplace characteristics, such as the proportion of workers on permanent contracts and the mobility class of the local MSOA, were found to have substantial variation across industries.

Table 3 HSE priority-specific fully adjusted model, percentage change in risk of workplace outbreak rate by industry

Variable	Percentage change in risk by industry (fully adjusted model)				
	Services	Education	Transport, distribution and warehousing	Manufacturing	Human health and social work
Workplace–proximity					
Physical proximity in the workplace	5.3 (5.1 to 5.5)	–	12 (12 to 13)	4.6 (4.2 to 5.1)	–
Workplace–permanence					
Proportion of workers on permanent contracts	−0.77 (−1.3 to −0.25)	–	−22 (−24 to −21)	−13 (−14 to −12)	–
Workplace–mobility class					
Mobility class–exurban	−17 (−19 to −16)	−9.5 (−11 to −7.7)	0.18 (−3.3 to 3.8)	−1.5 (−4.1 to 1.3)	−3.6 (−5.9 to −1.3)
Mobility class–metropolitan	−9.3 (−11 to −7.9)	−2.2 (−3.9 to −0.44)	−7.7 (−10 to −4.8)	−7.1 (−10 to −4.1)	−0.44 (−2.4 to 1.6)
Mobility class–rural	−11 (−14 to −8.4)	−44 (−46 to −42)	−6.3 (−11 to −1.1)	−0.16 (−4.3 to 4.2)	11 (6.7 to 16)
Mobility class–suburban	Baseline	Baseline	Baseline	Baseline	Baseline
Employee–vaccination					
Proportion of employees with two or more vaccination doses–Delta	−3.7 (−4 to −3.5)	1.8 (1.4 to 2.3)	−3.2 (−3.8 to −2.7)	−1.7 (−2.2 to −1.1)	−1.8 (−2.3 to −1.3)
Proportion of employees with two or more vaccination doses–Omicron	−3.2 (−3.5 to −2.9)	0.24 (−0.12 to 0.61)	−4.2 (−4.8 to −3.7)	−1.3 (−1.9 to −0.79)	−3.2 (−3.5 to −2.8)
Employee–age					
Proportion of employees aged 18–29	−0.83 (−1 to −0.62)	0.36 (0.081 to 0.64)	2.7 (2.2 to 3.2)	−0.78 (−1.3 to −0.3)	−1.1 (−1.5 to −0.77)
Proportion of employees aged 30–44	Baseline	Baseline	Baseline	Baseline	Baseline
Proportion of employees aged 45–59	3.4 (3.1 to 3.7)	2.5 (2.2 to 2.7)	5.1 (4.5 to 5.6)	3.7 (3.2 to 4.1)	2.2 (1.9 to 2.5)
Proportion of employees aged 60–64	−4.6 (−5.2 to −4.1)	−1.1 (−1.6 to −0.61)	−2.9 (−3.8 to −1.9)	−12 (−13 to −11)	−3.2 (−3.8 to −2.6)
Employee–ethnicity					
Proportion of employees with an asian ethnicity	0.046 (−0.065 to 0.16)	−0.25 (−0.37 to −0.12)	−1.2 (−1.4 to −0.94)	−0.48 (−0.7 to −0.26)	0.21 (0.058 to 0.37)
Proportion of employees with a black/African/ Caribbean ethnicity	1.6 (1.3 to 1.9)	0.49 (0.19 to 0.79)	2.7 (2.1 to 3.4)	1.8 (0.9 to 2.7)	2.4 (2.1 to 2.7)
Proportion of employees with a mixed/multiple/other ethnicity	−12 (−12 to −11)	−4.9 (−5.7 to −4.2)	−6.4 (−7.9 to −4.8)	−9.2 (−11 to −7.5)	−9.4 (−10 to −8.5)
Proportion of employees with a white ethnicity	Baseline	Baseline	Baseline	Baseline	Baseline
Employee–sex					
Proportion of employees who identify as female	−1.3 (−1.4 to −1.3)	1.4 (1.3 to 1.5)	1 (0.88 to 1.2)	−1.1 (−1.3 to −0.96)	0.31 (0.2 to 0.42)
Proportion of employees who identify as male	Baseline	Baseline	Baseline	Baseline	Baseline
Employee–IMD					
Employee IMD quintile	−15 (−16 to −14)	−5.2 (−6.4 to −4)	−14 (−16 to −12)	−10 (−12 to −8.3)	−17 (−18 to −16)
Work commute mode					
Proportion of employees using bus/metro/tram	−0.0093 (−0.13 to 0.12)	−1.4 (−1.6 to −1.3)	1.4 (1.1 to 1.6)	0.14 (−0.22 to 0.49)	0.049 (−0.12 to 0.22)
Proportion of employees using other transport	20 (18 to 22)	0.56 (−1.2 to 2.3)	8.7 (5.7 to 12)	2.6 (−1.4 to 6.7)	17 (14 to 19)
Proportion of employees using single occupancy	Baseline	Baseline	Baseline	Baseline	Baseline
Proportion of employees using taxi/vehicle passenger	5.3 (4.8 to 5.7)	−1.9 (−2.3 to −1.4)	4.1 (3.5 to 4.7)	1.6 (1.1 to 2.1)	4.3 (3.8 to 4.9)
Proportion of employees using train	−1.1 (−1.3 to −0.99)	−2.4 (−2.6 to −2.2)	−2.6 (−2.8 to −2.3)	−4.2 (−4.8 to −3.7)	−2.4 (−2.7 to −2.2)

Services; education; transport, distribution and warehousing; manufacturing; human health and social work. CIs are 90%. Dashes indicate variables for which insufficient data are available.
IMD, Indices of Multiple Deprivation.

In most industries, increased vaccination uptake was associated with reduced risk of workplace outbreaks. Despite primarily protecting against severe disease rather than infection, vaccination has been shown to reduce the length of the infectious period in infected individuals,[28] which could reduce the length of time workers are exposed to infectious individuals, thus reducing the chance of infection. During the Delta period, vaccination was associated with increased risk in Education. However, this is likely driven by the lack of vaccination in students, with children under 16 ineligible for vaccination during the Delta study period,[29] who are the main exposure for individuals working in education. Combined with the prevalence of Delta growing during the study period, the model cannot distinguish between the impact of increased vaccine uptake (which increases monotonically with time) and increased transmission risk.

Employee demographic risk factors varied across industry, but for most industries, a higher proportion of black ethnicities, 45–59 years, living in deprived areas (areas with reduced IMD), were associated with increased outbreak risk, and a higher proportion of 60–64 years in the workforce were associated with reduced outbreak risk in the majority of industries. The increased risk associated with a higher proportion of black ethnicities is consistent with community-based studies[30] and could be linked to increased risk of severe infection leading to higher ascertainment rates. The increased risk associated with more deprived areas has been observed in other studies,[31] potentially driven by larger household sizes or necessity to work while unwell. The reduced risk in the oldest age group may be related to increased public health messaging around COVID-19 risks for vulnerable individuals during the pandemic.[32]

Workplaces where a higher proportion of employees were expected to use Taxis (or travel as a passenger in a private car) and other forms of transport were found to have increased outbreak risk for many industries, with trains or single occupancy typically associated with the lowest risk. Some of the reduced risk associated with train travel could be due to increased awareness of

Table 4 HSE priority-specific fully adjusted model, percentage change in risk of workplace outbreak rate by industry

Variable	Percentage change in risk by industry (fully adjusted model)					
	Utilities	Mining and quarrying	Public service activities	Construction	Waste management and remediation	Agriculture, forestry and fishing
Workplace–proximity						
Physical proximity in the workplace	18 (14 to 22)	–	–	–0.3 (–2 to 1.4)	–	–
Workplace–permanence						
Proportion of workers on permanent contracts	–16 (–21 to –12)	–	–	1e+02 (60 to 1.6e+02)	–	–
Workplace–mobility class						
Mobility class–exurban	15 (0.27 to 32)	8.5 (–25 to 57)	–34 (–38 to –31)	–14 (–21 to –7.8)	10 (–5.7 to 29)	–39 (–49 to –27)
Mobility class–metropolitan	–52 (–58 to –47)	–82 (–91 to –64)	4.3 (0.29 to 8.4)	–20 (–25 to –15)	42 (23 to 64)	1e+02 (26 to 2.3e+02)
Mobility class–rural	–14 (–30 to 5.4)	24 (–16 to 83)	–20 (–26 to –13)	–20 (–29 to –9.3)	46 (19 to 79)	–69 (–75 to –63)
Mobility class–suburban	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Employee–vaccination						
Proportion of employees with two or more vaccination doses–Delta	–2.4 (–4.4 to –0.41)	6.8 (–7.2 to 23)	–3.2 (–4 to –2.3)	–1.5 (–2.9 to –0.21)	–2.2 (–5.3 to 0.98)	–7 (–12 to –2)
Proportion of employees with two or more vaccination doses–Omicron	–8.4 (–10 to –6.3)	5.2 (–3 to 14)	–6 (–6.7 to –5.2)	–1.5 (–3 to 0.011)	1.4 (–1.3 to 4.1)	–9.9 (–14 to –6.1)
Employee–age						
Proportion of employees aged 18–29	5.3 (4 to 6.6)	0.86 (–3.5 to 5.4)	1.1 (0.54 to 1.7)	1.3 (0.15 to 2.5)	5.6 (3.3 to 7.9)	3.5 (0.84 to 6.3)
Proportion of employees aged 30–44	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Proportion of employees aged 45–59	8.8 (7.4 to 10)	10 (4.8 to 17)	–0.16 (–0.78 to 0.45)	4 (2.8 to 5.1)	6.6 (4.4 to 8.9)	–2 (–5.1 to 1.2)
Proportion of employees aged 60–64	0.28 (–2.5 to 3.2)	–12 (–18 to –6.5)	–2.3 (–3.6 to –1)	–3.7 (–5.9 to –1.4)	–0.079 (–3.5 to 3.5)	–8.3 (–13 to –3.4)
Employee–ethnicity						
Proportion of employees with an Asian ethnicity	1.6 (0.41 to 2.7)	9.5 (3.6 to 16)	0.12 (–0.2 to 0.45)	1.2 (0.55 to 1.8)	1.4 (0.16 to 2.6)	2.1 (–0.37 to 4.6)
Proportion of employees with a black/African/Caribbean ethnicity	8.9 (6.8 to 11)	25 (–7 to 67)	0.47 (–0.069 to 1)	0.74 (–0.6 to 2.1)	3.5 (0.31 to 6.7)	–1.9 (–12 to 9.7)
Proportion of employees with a mixed/multiple/other ethnicity	–30 (–34 to –25)	–43 (–63 to –14)	6.4 (4.4 to 8.5)	–11 (–14 to –8)	–9.7 (–16 to –3)	55 (32 to 82)
Proportion of employees with a white ethnicity	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Employee–sex						
Proportion of employees who identify as female	–0.56 (–0.96 to –0.17)	–2.1 (–3.2 to –0.91)	–1.2 (–1.4 to –1)	–1 (–1.4 to –0.62)	1.3 (0.72 to 1.9)	0.81 (–0.54 to 2.2)
Proportion of employees who identify as male	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Employee–IMD						
Employee IMD quintile	–31 (–36 to –26)	23 (–6.8 to 61)	–14 (–17 to –11)	–19 (–23 to –14)	–11 (–20 to –1.1)	–3.5 (–15 to 10)
Work commute mode						
Proportion of employees using bus/metro/tram	–1.7 (–2.5 to –0.81)	2.7 (–6.2 to 12)	–2.2 (–2.5 to –1.8)	–1.1 (–1.7 to –0.49)	4.1 (1.8 to 6.3)	0.62 (–2.5 to 3.9)
Proportion of employees using other transport	13 (–2.1 to 29)	–32 (–49 to –8.1)	1.7 (–2.7 to 6.3)	3.8 (–6.4 to 15)	8.2 (–8.5 to 28)	–18 (–32 to –1.8)
Proportion of employees using single occupancy	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
Proportion of employees using taxi/vehicle passenger	–2.4 (–4.4 to –0.36)	–5.6 (–13 to 2.8)	–2.4 (–3.5 to –1.4)	3.1 (1.2 to 4.9)	–1.4 (–4.1 to 1.4)	4.1 (1.6 to 6.6)
Proportion of employees using train	–5.6 (–6.4 to –4.9)	–2.1 (–12 to 8.6)	–1.5 (–1.8 to –1.3)	0.29 (–0.23 to 0.82)	0.54 (–2 to 3.2)	–0.42 (–5.3 to 4.8)

Utilities; mining and quarrying; public service activities; construction; waste management and remediation; agriculture, forestry and fishing. CIs are 90%. Dashes indicate variables for which insufficient data are available.
IMD, Indices of Multiple Deprivation.

social distancing on trains (after strict social distancing measures had been in place[33]), or repeat infections over time on public transport leading a higher level of natural immunity at the time of the study.

Limitations

This analysis used multiple data sources linked by geographical unit (MSOA) with, on average, 2532 workers per unit. Ideally, analysis would be conducted at an individual or workplace level, but such data are not readily available at the national level. Therefore, the employee demographics had to be approximated, using the expected demographics of employees in the MSOA where people live and industry distribution in the MSOA where people work. If some workplaces have very atypical employee demographics, this would not be captured. Additionally, because the required variables were not yet available in the 2021 Census, the 2011 Census was the closest data that could be used to capture the demographic variables. If some locations have seen substantial changes since 2011, this may not accurately reflect the current conditions. The nature of the data creates the possibility

of ecological bias, whereby relationships observed at population-level differ from those at the level of individuals or workplaces.

Motivated by causal assumptions elucidated using DAGs, we have adjusted for key confounders. However, we have not considered causal relationships between variables in the same group (employee characteristics, workplace characteristics and modes of work commute). Moreover, the challenge of creating a DAG capable of capturing the complex network of factors relating to this question is high. We fitted a series of models for each candidate risk factor to investigate the impact of different confounder sets on the estimated relationships.

The outcome variable in our analysis, workplace outbreaks of COVID-19, uses clusters of COVID-19 cases in the workplace as a proxy. Some of these may be linked to false positive tests, and some may be community-acquired infections rather than workplace outbreaks. We identified workplace ascertainment rate as an important factor. However, this variable is unobserved, so we were unable to compare risk factors across industry sectors. We assumed that the ascertainment rate of workplace COVID-19 cases depends only on the type of workplace and not the location. There may be some variation due to local test availability which is not captured here.

Data on workplace physical proximity and the proportion of employees on permanent contracts (permanence) is only available at the two-digit SIC grouping level. Therefore, we assume that all workplaces in each two-digit SIC grouping have the same physical proximity and permanence. Therefore, the estimated effect size reflects the risk in the different sub-industries in each industry sector, rather than individual workplaces. For some industry sectors (agriculture, forestry and fishing, human health and social work, mining and quarrying, public service activities, and waste management and remediation), there are too few two-digit SIC groupings, so physical proximity and permanence are not considered for these industries.

We have only been able to assess risk factors where data were available. Many other risk factors will influence the risk of COVID-19 outbreaks in the workplace, including employee behaviour and workplace specific variables such as ventilation or humidity. Measuring such variables at a population level, as required by this study, is likely infeasible. Such analysis relies on site-specific outbreak investigations.[2 34]

Interpretation

This work identifies risk factors associated with risk of COVID-19 outbreaks in workplaces for different industry settings. This study was performed at a population level, so does not reflect individual level risk. The employee characteristics and work commute modes show how the expected characteristics of the local workforce are associated with the risk of workplace outbreaks, rather than the risk of being involved in an outbreak for individuals. The workplace characteristics reflect how the characteristics of sub-industries within the industry sector are associated with the risk of workplace outbreak occurrence. The mobility class characteristics reflect how the local area where the workplace was located is associated with the risk of workplace outbreaks.

Generalisability

Although this study focused on COVID-19, the methods used may be generalisable to other communicable and non-communicable diseases. The range of potential risk factors identified, and methods for linking these data at a population level, are likely to be relevant for many diseases. However, the influence of each risk factor will vary across different diseases.

CONCLUSION

Workplace outbreaks during the COVID-19 pandemic had substantial health and economic impacts. The novel national data linkages in this study have enabled the identification and investigation of workplace settings with increased risk of COVID-19 outbreaks in England. This innovative approach to investigate multiple and complex social, environmental and behavioural determinants associated with where people live and work, can support future pandemic preparedness and is applicable to wider occupational and environmental health research more generally.

Twitter Camilla Woodrow-Hill @CamillaWoodrow

Contributors Guarantor: CEO. Conceptualisation: CEO, RA, TW, OD, MG, JW, TA and YC. Literature review: RA, MG, JW, TA and YC. Data curation: CEO, RC, RP, JD, OJ, TB and CW-H. Formal analysis: CEO, RP and MT. Methodology: CEO, RP, OJ, MT, MG, JW and OD. Software: CEO, RP, MT, OJ, CW-H, TB, RC and JD. Visualisation: CEO. Writing–original draft preparation: CEO, OD, RA, JW, MG, TA and YC. Writing–review and editing: CEO, OD, RA, JW, MG, TA, YC and TW.

Funding This work was funded through the PROTECT (Partnership for Research in Occupational, Transport and Environmental COVID Transmission) COVID-19 National Core Study on Transmission and Environment, managed by the Health and Safety Executive on behalf of HM Government.

Disclaimer The contents of this paper, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect Health and Safety Executive policy or UK Health Security Agency or Office the National Statistics views.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study was conducted as part of the PROTECT National Core Study. Work was undertaken in line with national data regulations.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The data have not been made available due to the sensitivity of workplace-level data. An application for data access can be made to the UK Health Security Agency. Data requests can be made to the Office for Data Release (https://www.gov.uk/government/publications/accessing-ukhsa-protected-data/accessing-ukhsa-protected-data) and by contacting DataAccess@ukhsa.gov.uk. All requests to access data are reviewed by the Office for Data Release and are subject to strict confidentiality provisions in line with the requirements of: the common law duty of confidentiality, data protection legislation (including the General Data Protection Regulation), Caldicott principles, the Information Commissioner's statutory data sharing code of practice, and the national data opt-out programme. Access to the National Population Database can be requested by contacting: npd@hse.gov.uk.

ORCID iDs
Christopher E Overton http://orcid.org/0000-0002-8433-4010
Matthew Gittins http://orcid.org/0000-0002-9888-1197
Jack Wilkinson http://orcid.org/0000-0003-3513-4677
Yiqun Chen http://orcid.org/0000-0003-4218-084X

REFERENCES

1 Murti M, Achonu C, Smith BT, *et al*. COVID-19 workplace outbreaks by industry sector and their associated household transmission. *J Occup Environ Med* 2021;63:574–80.
2 Raja AI, van Veldhoven K, Ewuzie A, *et al*. Investigation of a SARS-CoV-2 outbreak at an automotive manufacturing site in England. *Int J Environ Res Public Health* 2022;19:6400.
3 Chen Y, Aldridge T, UK COVID-19 National Core Studies Consortium, *et al*. COVID-19 outbreak rates and infection attack rates associated with the workplace: a descriptive epidemiological study. *Epidemiology* [Preprint] 2021.
4 Talic S, Shah S, Wild H, *et al*. Effectiveness of public health measures in reducing the incidence of COVID-19, SARS-CoV-2 transmission, and COVID-19 mortality: systematic review and meta-analysis. *BMJ* 2021;375:e068302.
5 Rayner C, Campbell R. Long covid implications for the workplace. *Occup Med* 2021;71:121–3.
6 Keogh-Brown MR, Jensen HT, Edmunds WJ, *et al*. The impact of COVID-19, associated behaviours and policies on the UK economy: a computable general equilibrium model. *SSM Popul Health* 2020;12:100651.
7 Hamouche S. COVID-19 and employees' mental health: stressors, moderators and agenda for organizational actions. *Emerald Open Res* 2020;2:15.
8 Pijls BG, Jolani S, Atherley A, *et al*. Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. *BMJ Open* 2021;11:e044640.
9 Bardsley M, Morbey RA, Hughes HE, *et al*. Epidemiology of respiratory syncytial virus in children younger than 5 years in England during the COVID-19 pandemic, measured by laboratory, clinical, and syndromic surveillance: a retrospective observational study. *Lancet Infect Dis* 2023;23:56–66.
10 Hofmann FM, Olawumi E, Michaelis M, *et al*. Significance of Norovirus in occupational health: a review of published Norovirus outbreaks in central and northern Europe. *Int Arch Occup Environ Health* 2020;93:911–23.
11 Office for National Statistics. 2011 census geographie. Available: https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/2011censusgeographies [Accessed 25 Sep 2023].
12 Thakur V, Bhola S, Thakur P, *et al*. Waves and variants of SARS-CoV-2: understanding the causes and effect of the COVID-19 catastrophe. *Infection* 2022;50:309–25.
13 Office for National Statistics. 2011 census. Available: https://www.ons.gov.uk/census/2011census [Accessed 01 Mar 2023].
14 GOV.UK. English indices of deprivation 2019. Available: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 [Accessed 01 Mar 2023].
15 GOV.UK. Chapter 7: contact tracing and isolation. Available: https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk/chapter-7-contact-tracing-and-isolation [Accessed 01 Mar 2023].
16 Abbey R, Ricketts M, Lewis L, *et al*. The spatial clustering of contact tracing data to understand the risk of COVID-19 in workplace settings; 2023.
17 GOV.UK. COVID-19 testing data: methodology note [Withdrawn]. 2022. Available: https://www.gov.uk/government/publications/coronavirus-covid-19-testing-data-methodology/covid-19-testing-data-methodology-note
18 Statistics » COVID-19 Vaccinations. 2023 Available: https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/
19 The National population database: overview and developments. Available: https://npd-portal-healthandsafety.hub.arcgis.com/documents/192c0de323ad475f991e3e1305beccd2 [Accessed 01 Mar 2023].
20 Office for National Statistics. Inter-departmental business register (IDBR). Available: https://www.ons.gov.uk/aboutus/whatwedo/paidservices/interdepartmentalbusinesregisteridbr [Accessed 01 Mar 2023].
21 Aldridge T, Chen Y. Sector groupings for COVID-19 data analysis – COVID-OUT study. 2023. Available: https://osf.io/nzq7h/ [Accessed 06 Apr 2023].
22 Office for National Statistics. Industry by physical proximity to others in the UK 2019. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/adhocs/11838industrybyphysicalproximitytoothersintheuk2019 [Accessed 01 Mar 2023].
23 Office for National Statistics. Employees in temporary and permanent agency roles by demographic and employment variables, UK, April 2019 to March 2020. Available: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/adhocs/12545employeesintemporaryandpermanentagencyrolesbydemographicandemploymentvariablesukapril2019tomarch2020 [Accessed 01 Mar 2023].
24 Jahanshahi K, Jin Y. Identification and mapping of spatial variations in travel choices through combining structural equation modelling and latent class analysis: findings for great Britain. *Transportation* 2021;48:1329–59.
25 VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol* 2019;34:211–9.
26 Castelli JM, Rearte A, Olszevicki S, *et al*. Effectiveness of mRNA-1273, BNT162b2, and BBIBP-Corv vaccines against infection and mortality in children in Argentina, during predominance of Delta and Omicron COVID-19 variants: test negative, case-control study. *BMJ* 2022;379:e073070.
27 Coronavirus disease (COVID-19): how is it transmitted? Available: https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-how-is-it-transmitted [Accessed 12 Oct 2023].
28 Garcia-Knight M, Anglin K, Tassetto M, *et al*. Infectious viral shedding of SARS-CoV-2 Delta following vaccination: a longitudinal cohort study. *PLoS Pathog* 2022;18:e1010802.
29 NHS England. NHS rolls out COVID-19 jab to children aged 12 to 15. Available: https://www.england.nhs.uk/2021/09/nhs-rolls-out-covid-19-jab-to-children-aged-12-to-15/ [Accessed 25 Sep 2023].
30 Mathur R, Rentsch CT, Morton CE, *et al*. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the opensafely platform. *Lancet* 2021;397:1711–24.
31 Prats-Uribe A, Paredes R, Prieto-Alhambra D. Ethnicity, comorbidity, socioeconomic status, and their associations with COVID-19 infection in England: a cohort analysis of UK biobank data. *Epidemiology* [Preprint] 2020.
32 Coronavirus (COVID-19) – advice for workplaces. Available: https://www.hse.gov.uk/coronavirus/index.htm [Accessed 05 Oct 2023].
33 Britton P. New social distancing guidance for train passengers as services increase [Manchester Evening News]. 2020. Available: https://www.manchestereveningnews.co.uk/news/greater-manchester-news/trains-new-timetable-coronavirus-capacity-18267665 [Accessed 05 Oct 2023].
34 Chen Y, Beattie H, Simpson A, *et al*. A COVID-19 outbreak in a large meat-processing plant in England: transmission risk factors and controls. *Int J Environ Res Public Health* 2023;20:6806.