Original research

Occupation and COVID-19 mortality in England: a national linked data study of 14.3 million adults

Vahe Nafilyan ,[1,2] Piotr Pawelek,[3] Daniel Ayoubkhani ,[3] Sarah Rhodes ,[4] Lucy Pembrey,[5] Melissa Matz,[2,6] Michel Coleman,[6] Claudia Allemani,[6] Ben Windsor-Shellard,[1] Martie van Tongeren ,[7] Neil Pearce [8]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/oemed-2021-107818).

For numbered affiliations see end of article.

Correspondence to
Dr Vahe Nafilyan, Office for National Statistics, Newport, Newport, UK; vahe.nafilyan@ons.gov.uk

Received 5 July 2021
Accepted 9 December 2021

ABSTRACT

Objectives To estimate occupational differences in COVID-19 mortality and test whether these are confounded by factors such as regional differences, ethnicity and education or due to non-workplace factors, such as deprivation or prepandemic health.

Methods Using a cohort study of over 14 million people aged 40–64 years living in England, we analysed occupational differences in death involving COVID-19, assessed between 24 January 2020 and 28 December 2020.

We estimated age-standardised mortality rates (ASMRs) per 100 000 person-years at risk stratified by sex and occupation. We estimated the effect of occupation on COVID-19 mortality using Cox proportional hazard models adjusted for confounding factors. We further adjusted for non-workplace factors and interpreted the residual effects of occupation as being due to workplace exposures to SARS-CoV-2.

Results In men, the ASMRs were highest among those working as taxi and cab drivers or chauffeurs at 119.7 deaths per 100 000 (95% CI 98.0 to 141.4), followed by other elementary occupations at 106.5 (84.5 to 132.4) and care workers and home carers at 99.2 (74.5 to 129.4). Adjusting for confounding factors strongly attenuated the HRs for many occupations, but many remained at elevated risk. Adjusting for living conditions reduced further the HRs, and many occupations were no longer at excess risk. For most occupations, confounding factors and mediators other than workplace exposure to SARS-CoV-2 explained 70%–80% of the excess age-adjusted occupational differences.

Conclusions Working conditions play a role in COVID-19 mortality, particularly in occupations involving contact with patients or the public. However, there is also a substantial contribution from non-workplace factors.

Key messages

What is already known about this subject?

- Several studies have reported important occupational differences in the risk of SARS-CoV-2 infection and death, but there have been relatively few systematic comparisons of death rates in different occupations using population-level data. In addition, the mechanisms driving these differences are not well understood.

What are the new findings?

- There are large age-adjusted differences in COVID-19 mortality across occupations in England. Our results suggest that these differences are unlikely to be solely due to workplace exposures to SARS-CoV-2 but are largely due to confounding factors, such as geography, ethnicity or education, and factors other than workplace exposures, such as living conditions. However, people working in occupations that involve contacts with patients (eg, health and social care workers) or the public (eg, bus and taxi drivers, retail workers) remained at elevated risk of death involving COVID-19, even after accounting for other factors.

How might this impact on policy or clinical practice in the foreseeable future?

- While preventive measures are needed to reduce workplace exposures, reducing exposures outside the workplace is also crucial to reduce inequalities in the adverse consequences of COVID-19.

Check for updates

To cite: Nafilyan V, Pawelek P, Ayoubkhani D, *et al*. *Occup Environ Med* Epub ahead of print: [please include Day Month Year]. doi:10.1136/oemed-2021-107818

INTRODUCTION

The COVID-19 pandemic has been particularly severe in the UK. While most deaths occur among elderly adults,[1] many deaths have also occurred among those of working age, particularly among essential workers, such as healthcare workers, transport workers and those working in food retail and distribution.[2]

Several studies have reported important occupational differences in the risk of SARS-CoV-2 infection and death,[3–7] but there have been relatively few systematic comparisons of death rates in different occupations. Infections in healthcare workers have received the most attention,[8 9] with evidence that intensive care unit workers who care for patients with COVID-19 are at elevated risk. However, other occupations may also be at increased risk, particularly those that involve contact with the public.[10] In particular, age-standardised mortality rates (ASMRs) for death involving COVID-19 by occupation are high among taxi drivers and chauffeurs, bus and coach drivers, chefs, sales and retail assistants, and social care workers.[11]

Occupational inequality in COVID-19 mortality is a major public health problem,[10 12] but it is challenging to determine the extent to which working conditions drive these raised risks. Occupational differences in COVID-19 mortality could be caused by non-workplace factors such as living conditions at home or poor underlying (prepandemic) health. Deprivation, poor health and occupation are all linked.[13] For example, people working in low-paid, insecure jobs are also likely to experience poor housing conditions and overcrowding. They may also have low prepandemic health status, which is likely to increase COVID-19 mortality.[14 15] As a result, the elevated COVID-19 mortality rates may not solely be due to workplace exposures to SARS-CoV-2 but could be driven by non-workplace factors. This distinction is crucial for public health policy. If the excess risk is due to workplace exposures, preventive interventions in the workplace may help reduce inequalities in COVID-19 mortality. If the excess risk is also due to non-workplace factors such as living conditions at home (which may be associated with occupation, for instance because of insecure work, low pay, etc), then additional interventions would be required.

In this study, we estimated occupational differences in COVID-19 mortality in England during 2020. We have examined how much these differences changed after adjustment for non-workplace factors, using Cox proportional hazard models.

METHODS

Data

We used individual-level data from the Public Health Data Asset. This dataset is based on the 2011 Census in England, linked via the National Health Service (NHS) number to death records, Hospital Episode Statistics and the General Practice Extraction Service Data for Pandemic Planning and Research (GDPPR). To obtain NHS numbers, the 2011 Census was linked to the 2011–2013 NHS Patient Registers. It was first linked deterministically using 24 different matching keys, based on a combination of forename, surname, date of birth, sex and geography (postcode or unique property reference number). Probabilistic matching was then used to attempt to match records that were not linked deterministically, using 13 different combinations of personal identifiers. Candidate matches were assigned to Census records using the Felligi-Sunter probabilistic matching method. Of the 53 483 502 Census records, 50 019 451 were linked deterministically. A total of 555 291 additional matches were obtained using probabilistic matching (overall linkage rate: 94.6%).

We restricted our sample to individuals who were aged 31–55 years at the time of the 2011 Census and were therefore likely to be in stable employment both in 2011 and 2020 (by which time they were aged 40–64 years). Of the 17 407 025 people enumerated at the 2011 Census in England and Wales, aged 31–55 years in 2011, we excluded 675 447 people (3.9%) who could not be linked deterministically or probabilistically to the NHS Patient register and 318 254 individuals (1.9%) who had died between the Census and 24th January 2020. An additional 2 073 530 people (12.6%) were not linked to the English primary care records, either because they did not live in England in 2019 (the Census included people living in England and Wales) or because they were not registered with the NHS (see sample flow diagram in online supplemental table S1). After excluding 43 894 people who did not live in private households, our sample consisted of 14 295 900 individuals.

We examined the differences between occupation groups in the risk of death involving COVID-19 during the 11 months from 24 January to 28 December 2020.

Outcomes

Individuals in the study population were followed up from 24 January until 28 December 2020 for death involving COVID-19 (either in hospital or out of hospital), defined as confirmed or suspected COVID-19 death as identified by one of two International Classification of Diseases, 10th revision codes (U07.1 or U07.2) mentioned anywhere on the medical certificate of cause of death.

Exposure

The main exposure was the occupation at the time of the 2011 Census. Occupations are coded using a hierarchical classification, under the Standard Occupation Classification (SOC) 2010.[7] The most detailed classification (unit group, with four-digit codes) includes 369 categories, while the most aggregated (major group, with onr-digit codes) has only nine groups.

We derived a hybrid classification based on the submajor groups (two-digit codes), which include 25 categories. We broke down some submajor groups into minor groups (three-digit codes) and unit groups (four-digit codes) to assess selected occupations that have previously been shown to have high COVID-19 mortality, such as taxi drivers, security guards or care home workers.[4] Our final classification contained 41 categories (online supplemental table S2) in appendix). We also derived a classification of essential workers, based on the classification developed for a recent study using data from the UK Biobank.[3]

Because we used the occupation recorded at the 2011 Census, our exposure variable may be misclassified for some participants, since people may have left the labour force or changed occupation since 2011. To estimate the extent of misclassification, we analysed occupational mobility across major (one-digit SOC codes) groups between 2011 and 2019 using data from Understanding Society.[16] However, this analysis has two main limitations. First, because of the sample size of the data we used, we could not measure mobility using the same occupational classification as we used in our main analysis. Second, the pandemic may have caused some changes in occupation that we cannot capture.

Covariates

First, we aimed to estimate the occupational differences in COVID-19 mortality after adjusting for confounding. Sociodemographic factors likely to affect occupational choice, such as age, sex, ethnicity, education and geographical factors are associated with COVID-19 outcomes.[15 17–20] We therefore included these as potential confounding factors in a directed acyclic graph of the relationship between occupation and COVID-19 mortality (figure 1). We then assessed whether the remaining differences were due to work-related exposures.[21] Because we had no direct information on work-related exposures, we assessed these indirectly by adjusting for living conditions that are related to occupation (eg, because of insecure work, low pay, etc—see figure 1). We interpreted the residual effects of occupation as being due to workplace exposure to SARS-CoV-2. Many such indicators of living conditions are associated with the risk of COVID-19 infection or severe outcomes, including socioeconomic status and deprivation[15 22] and household composition.[23] Prepandemic health is also known to affect the risk of severe COVID-19 outcomes[14 15] and is associated with occupation.[24] One could argue that prepandemic health could be seen as a mediator and as a confounder of the relationship between occupation and COVID-19 outcomes, since health can determine occupational choices.[25] However, we only observe health status before the

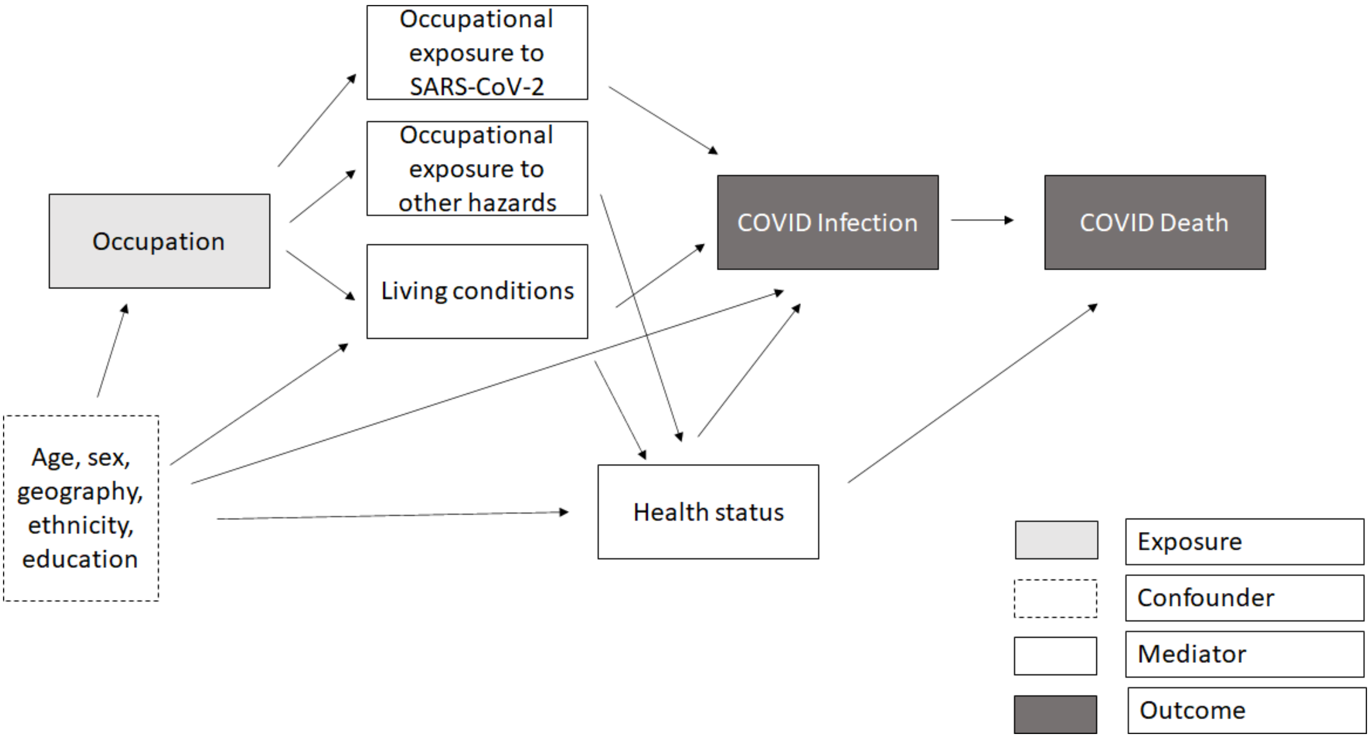

Figure 1 Directed acyclic graph (DAG) of the relationship between occupation and COVID-19 mortality.

pandemic, rather than the health status before individuals took their job. Therefore, in our main model, we consider prepandemic health as being caused by occupation, but we also present results from models adjusted for prepandemic health. In any case, living conditions and prepandemic health are all potential confounders of the association between workplace exposure to SARS-CoV-2 and COVID-19, because they may be associated with the risk of COVID-19 mortality, either through the propensity to become infected or the propensity to die once infected (figure 1). All covariates are summarised in online supplemental table S3 in the appendix. Geographical factors were based on the postcode from the GDPPR; sociodemographic characteristics were based on the 2011 Census; body mass index (BMI) and comorbidities were derived from the primary care and hospitalisation data following the definitions adopted by the QCOVID risk prediction model.[14]

Statistical analyses

For the period from 24 January 2020 to 28 December 2020, we calculated annualised ASMRs for death involving COVID-19 for each occupation using the European Standard Population as a reference population .[25]

First, to estimate the effect of occupation on COVID-19 mortality, we used Cox proportional hazard models to adjust for confounding factors (region, ethnicity, education, in addition to age and sex). Second, we assessed whether the differences were due to work-related exposures. As we do not observe workplace exposure to SARS-CoV-2, we adjusted for non-workplace factors (living conditions and prepandemic health) that are related to occupation and interpreted the residual effects of occupation as being due to workplace exposure. While non-workplace factors are mediating the relationship between occupation and COVID-19 death, they are also likely to confound the relationship between work-related exposures and COVID-19 death, which is what we aim to identify. To that end, we estimated five models, sequentially adjusting for additional covariables to assess how they might confound or mediate differences in the risk of death involving COVID-19 between occupations (see figure 1). Our first model was only adjusted for age. The second model also adjusted for geographical factors (region, population density and rural–urban classification) to account for the differential spread of the virus in different areas. The third model further adjusted for other confounding factors, ethnicity and education, which are related both to occupation and to COVID-19 risk. The fourth model also controlled for non-workplace factors (living conditions), including socioeconomic factors (Index of Multiple Deprivation, household deprivation, household tenancy and house type) and household composition (household size, children in the household and overcrowding). Finally, the last model adjusted for prepandemic health (BMI, chronic kidney disease, learning disability, cancer or immunosuppression, and other conditions; see online supplemental table S3 for details on all the covariates). We used corporate managers and directors as the reference category, because it is a large group with a low absolute risk.[11]

As supplementary analyses, we also used alternative occupational classifications (major groups, essential workers classification). We also estimated HRs using all other occupations (rather than corporate managers and directors) as the reference group.

RESULTS

Characteristics of the study population

Our analytical sample comprised 14 295 900 people aged 40–64 years (mean age 52 years, 51% female) who were alive on 24 January 2020, living in private households in England in 2019, were employed in 2011 and completed the 2011 Census. Between 24 January and 28 December 2020, 4552 people (0.003%) died from a cause related to COVID-19; characteristics of these individuals are summarised in table 1 (further details

Table 1 Characteristics of the study population and those who died from a cause related to COVID-19

		Population	COVID-19 deaths
		n (%)	n (%)
Sex	Male	6 964 839 (48.72)	2970 (65.25)
	Female	7 331 061 (51.28)	1582 (34.75)
Age	Mean (SD)	52.19 (6.96)	57.05 (5.84)
Ethnicity	Bangladeshi	76 776 (0.54)	77 (1.69)
	Black African	235 255 (1.65)	191 (4.20)
	Black Caribbean	179 185 (1.25)	132 (2.90)
	Chinese	82 158 (0.57)	14 (0.31)
	Indian	411 615 (2.88)	282 (6.20)
	Mixed	172 402 (1.21)	51 (1.12)
	Other	363 548 (2.54)	214 (4.70)
	Pakistani	210 678 (1.47)	196 (4.31)
	White British	11 762 187 (82.28)	3218 (70.69)
	White other	802 096 (5.61)	177 (3.89)
Occupation	Administrative occupations	1 252 049 (8.76)	302 (6.63)
	Bus and coach drivers	71 019 (0.50)	67 (1.47)
	Business and public service associate professionals	904 207 (6.32)	159 (3.49)
	Business, media and public service professionals	717 153 (5.02)	148 (3.25)
	Care workers and home carers	397 620 (2.78)	188 (4.13)
	Caring personal service occupations	470 906 (3.29)	75 (1.65)
	Caring personal services	218 019 (1.53)	68 (1.49)
	Cleaners and domestics	339 484 (2.37)	170 (3.73)
	Corporate managers and directors	871 555 (6.10)	187 (4.11)
	Culture, media and sports occupations	259 159 (1.81)	39 (0.86)
	Customer service occupations	198 344 (1.39)	67 (1.47)
	Elementary administration and service occupations	46 135 (0.32)	19 (0.42)
	Elementary administration occupations	133 518 (0.93)	57 (1.25)
	Elementary cleaning occupations	80 885 (0.57)	42 (0.92)
	Elementary security occupations	182 686 (1.28)	120 (2.64)
	Elementary storage occupations	212 772 (1.49)	120 (2.64)
	Elementary trades and related occupations	277 084 (1.94)	141 (3.10)
	Food preparation and hospitality trades	246 335 (1.72)	105 (2.31)
	Health and social care associate professionals	178 277 (1.25)	51 (1.12)
	Health professionals	579 622 (4.05)	145 (3.19)
	Large goods vehicle drivers	139 955 (0.98)	69 (1.52)
	Leisure, travel and related personal service occupations	299 868 (2.10)	89 (1.96)
	Managers and directors in retail and wholesale	254 033 (1.78)	82 (1.80)
	Other elementary services occupations	279 992 (1.96)	145 (3.19)
	Other managers and proprietors	549 931 (3.85)	171 (3.76)
	Plant and machine operatives	144 901 (1.01)	78 (1.71)
	Process operatives	167 785 (1.17)	101 (2.22)
	Process, plant and machine operatives	225 487 (1.58)	105 (2.31)
	Protective service occupations	187 510 (1.31)	27 (0.59)
	Sales occupations	774 754 (5.42)	284 (6.24)
	Science, engineering and technology associate professionals	214 286 (1.50)	72 (1.58)
	Science, research, engineering and technology professionals	627 758 (4.39)	103 (2.26)
	Secretarial and related occupations	457 792 (3.20)	86 (1.89)
	Skilled agricultural and related trades	136 740 (0.96)	32 (0.70)
	Skilled construction and building trades	542 774 (3.80)	171 (3.76)
	Skilled metal, electrical and electronic trades	522 167 (3.65)	214 (4.70)
	Taxi and cab drivers and chauffeurs	103 956 (0.73)	124 (2.72)
	Teaching and educational professionals	626 682 (4.38)	93 (2.04)
	Textiles, printing and other skilled trades	125 133 (0.88)	52 (1.14)
	Transport and mobile machine drivers and operatives	132 163 (0.92)	78 (1.71)

in online supplemental table S4). Using data from Understanding Society, we estimated the proportions of people who remained in the same major occupation group between 2011 and 2019 was 72.4% for men and 68.1% for women (see online supplemental table S5). For men, it ranged from 42.1% for those working in sales and customer service occupations to 78.9% for those working in professional occupations. For women, it ranged from 57.9% for process, plant and machine operatives to 73.9% for those working in professional occupations.

Age-standardised mortality rates

Table 2 shows the annualised ASMRs for death involving COVID-19 for men aged 40–64 years. The ASMRs were highest among those working as taxi and cab drivers or chauffeurs at 119.7 deaths per 100 000 men (95% CI 98.0 to 141.4) over

Table 2 Annualised age-standardised mortality rates for death involving COVID-19 per 100 000 adults aged 40 to 64 years, by sex and occupation

Occupation	Men	Women
Taxi and cab drivers and chauffeurs	119.7 (98.0 to 141.4)	–
Other elementary occupations	106.5 (84.5 to 132.4)	36.5 (29.1 to 45.3)
Care workers and home carers	99.2 (74.5 to 129.4)	39.4 (32.7 to 46.1)
Secretarial and related occupations	96.2 (57.0 to 152.0)	15.6 (12.0 to 19.9)
Elementary security occupations	93.8 (74.7 to 116.2)	40.5 (28.3 to 56.2)
Bus and coach drivers	90.8 (69.0 to 117.2)	–
Cleaners and domestics	84.2 (62.2 to 111.4)	40.0 (32.7 to 47.3)
Customer service occupations	75.2 (52.2 to 104.8)	25.3 (17.3 to 35.7)
Van drivers	71.6 (57.5 to 85.8)	–
Caring personal services	69.1 (45.0 to 101.3)	23.0 (16.5 to 31.2)
Process operatives	65.2 (51.2 to 81.8)	46.9 (30.3 to 69.2)
Food preparation and hospitality trades	64.3 (50.4 to 81.0)	28.2 (19.0 to 40.1)
Sales occupations	63.8 (51.7 to 75.9)	30.4 (25.9 to 34.9)
Elementary storage occupations	63.5 (51.4 to 75.6)	37.8 (20.6 to 63.4)
Mobile machine and other drivers	59.3 (46.4 to 74.5)	–
Health and social care associate professionals	59.2 (38.2 to 87.5)	20.1 (13.1 to 29.4)
Health professionals	54.2 (41.7 to 69.2)	18.7 (14.8 to 23.2)
Elementary trades and related occupations	53.3 (43.2 to 64.9)	45.2 (32.4 to 61.3)
Leisure, travel and related personal service occupations	53.2 (38.5 to 71.7)	22.8 (16.7 to 30.5)
Elementary administration occupations	52.0 (38.2 to 69.3)	–
Administrative occupations	51.2 (42.6 to 59.8)	17.0 (14.4 to 19.6)
Elementary cleaning occupations excl. cleaners and domestics	48.9 (32.7 to 70.1)	51.7 (27.5 to 88.4)
Textiles, printing and other (excl. food prep and hospitality) skilled trades	48.8 (35.0 to 66.2)	23.0 (11.4 to 41.2)
Assemblers and construction operatives	48.2 (37.9 to 60.4)	36.7 (24.0 to 53.2)
Large goods vehicle drivers	47.3 (36.3 to 60.5)	–
Plant and machine operatives	45.3 (34.6 to 58.2)	77.2 (43.1 to 126.8)
Science, engineering and technology associate professionals	44.3 (33.8 to 57.1)	20.1 (10.4 to 35.2)
Managers and directors in retail and wholesale	43.9 (33.4 to 56.7)	23.6 (15.0 to 35.5)
Skilled metal, electrical and electronic trades	40.4 (34.9 to 45.9)	–
Other managers and proprietors	37.9 (31.0 to 44.8)	20.5 (15.3 to 26.9)
Caring personal service occupations excluding care workers and home carers	–	15.4 (11.9 to 19.7)
Skilled construction and building trades	31.5 (26.7 to 36.2)	–
Business, media and public service professionals	27.3 (22.0 to 32.6)	17.2 (12.6 to 23.1)
Teaching and educational professionals	26.0 (18.9 to 34.8)	11.2 (8.3 to 14.9)
Culture, media and sports occupations	25.8 (17.3 to 36.8)	–
Corporate managers and directors excl. those in retail and wholesale	25.6 (21.5 to 29.7)	12.5 (8.5 to 17.7)
Business and public service associate professionals	25.6 (20.8 to 30.4)	14.0 (10.4 to 18.6)
Skilled agricultural and related trades	23.4 (15.3 to 34.0)	–
Science, research, engineering and technology professionals	20.9 (16.8 to 25.6)	10.3 (5.0 to 18.7)
Protective service occupations	19.5 (12.5 to 28.8)	–

Deaths involving COVID-19 occurring between 24 January 2020 and 28 December 2020. Ninety-five per cent CIs are reported. Mortality rates are standardised to the 2013 European Standard Population and annualised. Occupation classification and associated SOC 2010 codes are given in online supplemental table S2).
Data are not reported if there were 10 or fewer deaths.
SOC, Standard Occupation Classification.

the period, followed by other elementary occupations at 106.5 (95% CI 84.5 to 132.4) and care workers and home carers at 99.2 (95% CI 74.5 to 129.4) (table 2). The ASMRs were lowest among those working in the protective service occupations at 19.5 (95% CI 12.5 to 28.8), followed by science, research, engineering and technology professionals at 20.9% (95% CI 16.8 to 25.6) and skilled agricultural and related trades occupations at 23.4 deaths per 100 000 men (95% CI 15.3 to 34.0).

For women aged 40–64 years, the ASMRs for death involving COVID-19 were greatest among those working as plant and machine operatives at 77.2 deaths per 100 000 women (95% CI 43.1 to 126.8), followed by elementary cleaning occupations excl. cleaners and domestics (51.7 (95% CI 27.5 to 88.4)), process operatives at (46.9 (95% CI 30.3 to 69.2)), elementary trades and related occupations (45.2 (95% CI 32.4 to 61.3)) and elementary security operations at (40.5 (95% CI 28.3 to 56.2)) (table 2). The ASMRs were lowest among those working in the teaching and educational professions at 11.2 (95% CI 8.3 to 14.9), followed by corporate managers and directors, excluding those in retail and wholesale at 12.5 (95% CI 8.5 to 17.7) and business and public service associate professionals at 17.2 (95% CI 12.6 to 23.1).

Adjusted HRs

The age-adjusted HRs, relative to corporate managers, indicated large differences in the risk of death involving COVID-19 between occupations for both men and women (figure 2, light grey bar). For men, adjustment for confounding factors (grey bar) strongly attenuated the HRs for many occupations, but many remained at elevated risk. For instance, after adjusting for confounding, men working in secretarial and related occupations remained at elevated risk of death involving COVID-19 with an adjusted HR of 2.67 (95% CI 1.64 to 4.36), compared with an age-adjusted HR of 3.89 (95% CI 2.39 to 6.34). Care workers and home carers were also at higher risk of death involving COVID-19, with an adjusted HR of 2.36 (95% CI 1.72 to 3.22), compared with an age-adjusted HR of 3.75 (95% CI 2.75 to 5.12). Men working in caring personal services, as social care associate professionals or as health professionals were also at elevated risk, with adjusted HR of 1.74 (95% CI 1.35 to 2.25), 1.78 (95% CI 1.16 to 2.71) and 1.68 (95% CI 1.25 to 2.27), respectively. Working as a taxi, bus or van driver was also associated with a higher risk of COVID-19 mortality, even after adjusting for confounding (adjusted HR of 1.82 (95% CI 1.41

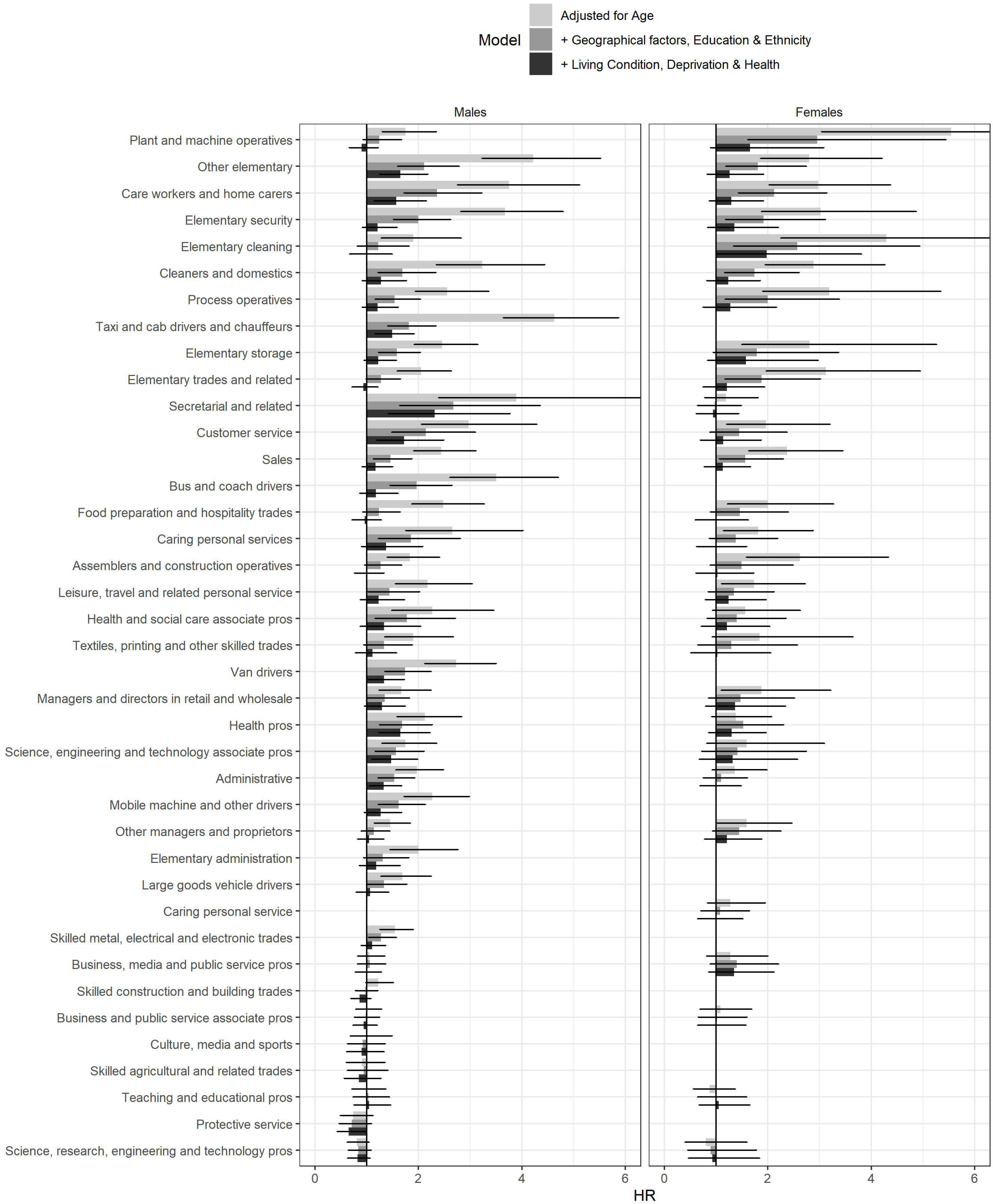

Figure 2 HRs for death involving COVID-19 for adults aged 40–64 years, compared with corporate managers and directors, by sex. Note: geographical factors include region, population density and urban/rural classification. Living conditions include Index of Multiple Deprivation decile group, household deprivation, social grade, household tenancy, type of accommodation, household size, multigenerational household, household with children, health factors include body mass index, chronic kidney disease (CKD), learning disability, cancer and immunosuppression, and other conditions. See online supplemental table S3 for more details. Occupation classification and associated SOC 2010 codes are given in online supplemental table S2). Numerical results can be found in online supplemental table S6 and S7.

to 2.34), 1.96 (95% CI 1.45 to 2.65) and 1.85 (95% CI 1.22 to 2.81), respectively).

To assess the role of workplace exposure to SARS-CoV-2, we further adjusted for living conditions and prepandemic health status, which can mediate the effect of occupation on COVID-19 outcomes but are not due to workplace exposure to SARS-CoV-2. After adjusting for living conditions and prepandemic health, many occupations were no longer at excess risk. For instance, the fully adjusted HR for bus and coach drivers fell to 1.18 (95% CI 0.86 to 1.61). A notable exception is health professionals, for whom adjustment for socioeconomic factors did not affect the HRs.

For women, adjusting for confounding also greatly attenuated the estimated difference in risk between occupations. The highest age-adjusted HRs were observed for plant and machine operatives (5.54 (95% CI 3.04 to 10.10))) and those working in elementary cleaning occupations (4.29 (95% CI 2.25 to 8.18). Adjusting for confounders reduced the magnitude of the HRs for most occupations, with the HR falling to 2.96 (95% CI 1.61 to 5.45) for plant and machine operatives and 2.57 (95% CI 1.34 to 4.94) for those working in elementary cleaning occupations. Adjusting for living conditions reduced further the HRs towards one. For many occupations, the HRs are of similar magnitude to those for men, but less precise because of smaller numbers.

For most occupations, confounding factors and mediating factors other than workplace exposures explained 70%–80% of the excess age-adjusted HRs. Adjusting for socio-economic status had the largest impact on the confounding-adjusted HRs (online supplemental table S6 and S7) for men and women, respectively). A notable exception is health professionals, for whom adjustment for socioeconomic factors did not affect the HRs. Adjusting for health conditions known to be related to COVID-19 mortality[14 15] had little impact on the HRs.

HRs obtained when using all other occupations (rather than corporate managers and directors) as a reference group were similar. The unadjusted HRs were slightly lower, but the adjusted estimates were similar to those in our main analyses (online supplemental table S8).

Table 3 shows the HRs for essential workers compared with non-essential workers as the reference category. Overall, essential workers were at higher risk of death involving COVID-19 than non-essential workers, and most categories of essential workers also had higher mortality. Once again, the differences were generally much attenuated after adjusting for potential confounding and mediating factors other than workplace exposure; a notable exception is healthcare professionals. We also report HRs for major groups, compared with directors and managers, in online supplemental table S9.

DISCUSSION

By combining data from the 2011 Census with electronic health records, the Public Health Data Asset enabled us to analyse occupational differences in COVID-19 mortality. Information on occupation is not available in traditional electronic health records in the UK, and the Census is the only source of population-wide occupation data. With data for over 14 million people aged 40–64 years who were living in England at the beginning of the pandemic, we were able to estimate COVID-19 mortality for detailed occupational groups and to estimate whether the differences in mortality are driven by workplace-related factors, or by other confounding and mediating factors. We found large age-adjusted differences in COVID-19 mortality across occupations. For most occupations, confounding factors and mediating factors other than workplace exposure to SARS-CoV-2 explained about 70%–80% of the excess age-adjusted HRs.

The main limitation of our study is that the information on occupation is 9 years out of date. Therefore, our exposure is likely to be misclassified for a proportion of people, because they have left the labour force or changed occupation since 2011, especially during the pandemic. To mitigate measurement error, we restricted our analysis to people aged 40–64 years, who had a relatively high occupational stability, as shown in our analysis of a large longitudinal household survey. Exposure misclassification is nonetheless likely to bias the estimated HRs. However, the HRs are high for many occupations found to be at elevated risk

Table 3 HRs for death involving COVID-19 for adults aged 40–64 years, compared with non-essential workers, by sex

	Men			Women		
Occupation	Age adjusted	Adjusted for confounders	Fully adjusted	Age adjusted	Adjusted for confounders	Fully adjusted
All essential workers*	1.45 (1.34–1.56)	1.31 (1.21–1.41)	1.22 (1.13–1.32)	1.16 (1.05–1.28)	1.10 (0.99–1.21)	1.06 (0.96–1.17)
Taxi and cab drivers and chauffeurs	3.08 (2.56–3.70)	2.01 (1.67–2.43)	1.39 (1.14–1.70)	3.94 (1.634–9.48)	2.59 (1.075–6.26)	2.45 (1.014–5.92)
Support staff	2.39 (1.68–3.41)	1.98 (1.39–2.83)	1.74 (1.22–2.49)	0.95 (0.673–1.34)	2.07 (0.860–5.00)	0.78 (0.550–1.10)
Bus and coach drivers	2.33 (1.81–3.00)	1.53 (1.19–1.98)	1.11 (0.85–1.45)	2.95 (1.226–7.12)	1.05 (0.860–1.28)	1.73 (0.716–4.18)
Sanitary workers	1.84 (1.46–2.32)	1.12 (0.89–1.42)	1.18 (0.93–1.50)	1.78 (1.473–2.16)	1.36 (1.164–1.59)	1.09 (0.892–1.33)
Social care	1.83 (1.51–2.20)	1.54 (1.27–1.85)	1.27 (1.04–1.53)	1.62 (1.390–1.89)	1.16 (0.480–2.79)	1.18 (1.010–1.39)
Van drivers	1.81 (1.48–2.22)	1.28 (1.05–1.57)	1.26 (1.03–1.55)	1.59 (0.661–3.84)	1.09 (0.709–1.66)	1.27 (0.526–3.06)
Health associate professionals	1.65 (1.26–2.16)	2.11 (1.60–2.78)	1.86 (1.41–2.46)	0.92 (0.746–1.15)	1.07 (0.914–1.26)	1.22 (0.969–1.54)
Food retail and distribution	1.41 (1.22–1.63)	1.26 (1.09–1.46)	1.14 (0.98–1.32)	1.39 (1.187–1.63)	0.89 (0.633–1.26)	1.02 (0.867–1.20)
Other transport workers	1.21 (1.02–1.43)	1.07 (0.91–1.27)	1.10 (0.93–1.30)	0.36 (0.115–1.11)	1.33 (1.052–1.67)	0.31 (0.098–0.95)
Health professionals	1.21 (0.82–1.78)	1.88 (1.27–2.79)	1.45 (0.97–2.15)	0.25 (0.079–0.76)	0.79 (0.623–1.00)	0.45 (0.145–1.42)
Food production	1.12 (0.86–1.45)	1.17 (0.90–1.52)	1.15 (0.89–1.50)	1.48 (0.968–2.26)	0.50 (0.161–1.56)	1.15 (0.750–1.77)
Education	0.63 (0.47–0.84)	0.86 (0.64–1.16)	0.91 (0.68–1.23)	0.56 (0.446–0.70)	0.31 (0.099–0.96)	0.83 (0.653–1.05)
Police and protective services	0.45 (0.31–0.67)	0.59 (0.40–0.88)	0.60 (0.40–0.88)	0.38 (0.123–1.19)	0.44 (0.141–1.37)	0.50 (0.160–1.54)

Note: fully adjusted Cox regression models include geographical factors (region, population density and urban/rural classification), ethnicity, socioeconomic characteristics (Index of Multiple Deprivation decile group, household deprivation, educational attainment, social grade, household tenancy, type of accommodation, household size, multigenerational household and household with children), health (body mass index, chronic kidney disease, learning disability, cancer or immunosuppression and other conditions). See online supplemental table S3 for more details.

*Essential workers include all types of essential workers listed in the table. The HR for essential workers is obtained using a separate model, where the exposure is a binary variable.

in official estimates based on occupational data from the death certificates.[11] Misclassification of occupation would be constant across our various analyses and could not explain the substantial decrease in most HRs after adjustment for confounders. However, the confounders may also likely be misclassified. As adjustment for confounders strongly reduced the estimated occupational associations, if we had more accurate or detailed confounder data, the adjustment may have driven the HR estimates even closer towards unity.

Another limitation is that the outcome variable, death involving COVID-19, may also be misclassified, as not all deaths involving COVID-19 may have been captured on death certificates. Conversely, not all deaths for which COVID-19 was mentioned on the death certificate may have involved the disease. We cannot exclude some non-differential misclassification among some occupations, such as healthcare workers. If deaths for healthcare workers are more likely to have been wrongly attributed to COVID-19 than for people from other occupations, then we would expect the HRs to be upward biased. In addition, some deaths may not have been registered by the end of the study period if they had been sent to a coroner. If this disporportionally affects some occupational groups such as healthcare workers, then we would expect the HRs for that group to be biased towards one.

Finally, the dataset only included people who were enumerated at the 2011 Census and were registered with the NHS in 2019. Therefore, it excluded recent migrants and people from marginalised groups who are not registered with the NHS. We were also unable to identify all emigrants who had left the country and were no longer at risk.

Our age-adjusted results are consistent with official estimates of COVID-19 mortality by occupation group.[11] However, these elevated risks were greatly attenuated after adjustment for non-workplace factors, such as geographic factors, sociodemographic factors and prepandemic health. A recent study based on the UK Biobank found that compared with non-essential workers, medical support staff and healthcare professionals had the highest risk of severe COVID-19.[3] We also found that, among men, healthcare professionals were at increased risk of death involving COVID-19, but the fully adjusted HRs for healthcare professionals in our study were similar to those working as care workers, taxi drivers or in secretarial occupations. Our results are also consistent with US studies documenting higher mortality rates in essential workers, such as transportation/logistics workers, healthcare workers[5,26] and retail workers.[27]

Our findings are also generally consistent with a recent analysis of data from the UK Coronavirus Infection Survey (CIS), which found increased risks for a similar list of occupations.[28] This analysis found that the occupational differences largely disappeared after adjustment for other factors, but the adjustment included factors that are likely to be inherent to working conditions (inability to work at home and inability to socially distance at work) and are therefore on the causal pathway linking occupational exposure and infection. Thus, our adjusted findings are not directly comparable with those obtained from CIS.

Our age-standardised mortality rates and age-adjusted HRs confirm that there is a wide variation in the risk of death involving COVID-19 between occupations. Adjusting for confounding factors (geography, ethnicity and education) reduced HRs for most occupations, suggesting that part of the age-adjusted differences is due to confounding factors. Differences in COVID-19 mortality by geographical factors are well documented,[19] and disparities in mortality between ethnic groups have been reported and attributed to a wide range of factors.[17,18,23] Workplace exposure to SARS-CoV-2 is only one of several factors driving the differences in the risk of death involving COVID-19 between occupations: living conditions or prepandemic health may also contribute to the observed differences. Adjusting for these factors reduced the HRs for most occupations, and for many, the adjusted HRs were no linger different from one. However, people who worked in occupations that involved contacts with patients (eg, health and social care workers) or the public (eg, bus and taxi drivers, and retail workers) remained at elevated risk of death involving COVID-19. For health professionals, adjusting for socioeconomic factors did not affect the HRs, because they tend to be less disadvantaged.

Other occupations that do not involve contact with patients or the public may also have increased risks due to specific working conditions (eg, overcrowding in the workplace, lack of ventilation, lack of personal protective equipment, etc),[29] but our analyses indicate that these relative risks are generally small after adjustment for confounding. This could be due to furlough, remote working or other interventions implemented by employers to limit infections, which could have limited workplace exposure in occupations such as teaching and educational professionals and administrative occupations. Our results do not mean that infection is not occurring in specific workplaces. While workplace outbreaks have been reported in various industries,[30,31] these are not sufficient to produce sector-wide increased risks after adjustment for non-workplace factors.

Our analyses have confirmed that many occupations have elevated risks of COVID-19 mortality. However, these associations were greatly attenuated, for many occupations, after adjustment for measures of deprivation and geographical factors, suggesting that differences in risk between occupations are a result of a complex mix of different factors. Several occupations showed increased risks, even after comprehensive adjustment, and it is likely that working conditions played a role. However, our findings also indicate that non-workplace factors also play a major role. Preventive measures are needed to reduce workplace exposures, but also to reduce exposures outside the workplace, including overcrowding, inadequate housing and deprivation.

Author affiliations
[1]Health Analysis Division, Office for National Statistics, Newport, UK
[2]Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, UK
[3]Methodology Division, Office for National Statistics, Newport, UK
[4]Centre for Biostatistics, School of Health Sciences, Faculty of Biology, Medicine and Health, School of Health Sciences, The University of Manchester, Manchester, UK
[5]Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK
[6]Cancer Survival Group, Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
[7]Centre for Occupational and Environmental Health, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
[8]Epidemiology and Population Health, London School of Hygiene, London, UK

Twitter Martie van Tongeren @martievt

Acknowledgements We would like to thank Raymond Agius, Sheila Bird, David Coggon, Steve Robertson and Lesley Rushton for their comments on the draft manuscript.

Contributors Study conceptualisation was led by VN and NP. All authors contributed to the development of the research question, study design, with development of statistical aspects led by VN, NP and PP. VN and PP were involved in data specification, curation and collection. VN and PP conducted and checked the statistical analyses. All authors contributed to the interpretation of the results. VN and NP wrote the first draft of the paper. All authors contributed to the critical revision of the manuscript for important intellectual content and approved the final version of the manuscript. VN had full access to all data in the study and takes responsibility of the integrity of the data and the accuracy of the data analysis.

The lead author (VN) is the garantor and affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Funding This work was supported by funding through the National Core Study 'PROTECT' programme, managed by the Health and Safety Executive on behalf of HM Government and by a grant from the Colt Foundation (CF/05/20).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was obtained from the National Statistician's Data Ethics Advisory Committee (NSDEC(20)12).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. The data used in this study are available on the Office for National Statistics (ONS) Secure Research Service for Accredited researchers as the Public Health Research Database. Researchers can apply for accreditation through the Research Accreditation Service.

ORCID iDs
Vahe Nafilyan http://orcid.org/0000-0003-0160-217X
Daniel Ayoubkhani http://orcid.org/0000-0001-6352-0394
Sarah Rhodes http://orcid.org/0000-0002-5837-801X
Martie van Tongeren http://orcid.org/0000-0002-1205-1898
Neil Pearce http://orcid.org/0000-0002-9938-7852

REFERENCES

1 ONS. Deaths registered Weekly in England and Wales, provisional: week ending 5 February 2021, 2021. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregisteredweeklyinenglandandwalesprovisional/weekending5february2021
2 ONS. Coronavirus (COVID-19) related deaths by occupation, England and Wales coronavirus (COVID-19) related deaths by occupation, England and Wales: deaths registered between 9 March and 28 December 2020, 2020. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/bulletins/coronaviruscovid19relateddeathsbyoccupationenglandandwales/deathsregisteredbetween9marchand28december2020
3 Mutambudzi M, Niedwiedz C, Macdonald EB, *et al*. Occupation and risk of severe COVID-19: prospective cohort study of 120 075 UK Biobank participants. *Occup Environ Med* 2020. doi:10.1136/oemed-2020-106731. [Epub ahead of print: 09 Dec 2020].
4 ONS. Coronavirus (COVID-19) related deaths by occupation, England and Wales: deaths registered between 9 March and 28 December 2020, 2021. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/bulletins/coronaviruscovid19relateddeathsbyoccupationenglandandwales/deathsregisteredbetween9marchand28december2020
5 Hawkins D, Davis L, Kriebel D. COVID-19 deaths by occupation, Massachusetts, March 1–July 31, 2020. *Am J Ind Med* 2021;64:238–44.
6 Magnusson K, Nygård K, Methi F. Occupational risk of COVID-19 in the 1st vs 2nd wave of infection. *merRxiv* 2020.
7 Chen Y-H, Glymour M, Riley A. Excess mortality associated with the COVID-19 pandemic among Californians 18–65 years of age, by occupational sector and occupation: March through November 2020. *PLOS One* 2021;16.
8 Sim MR. The COVID-19 pandemic: major risks to healthcare and other workers on the front line. *Occup Environ Med* 2020;77:281–2.
9 Chadeau-Hyam M, Bodinier B, Elliott J, *et al*. Risk factors for positive and negative COVID-19 tests: a cautious and in-depth analysis of UK Biobank data. *Int J Epidemiol* 2020;49:1454–67.
10 Burdorf A, Porru F, Rugulies R. The COVID-19 (coronavirus) pandemic: consequences for occupational health. *Scand J Work Environ Health* 2020;46:229–30.
11 ONS, "Coronavirus (COVID-19) related deaths by occupation, England and Wales: deaths registered between 9 March and 28 December 2020, 2021[Online]. Available. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/bulletins/coronaviruscovid19relateddeathsbyoccupationenglandandwales/deathsregisteredbetween9marchand28december2020
12 Agius R. Covid-19 and health at work. *Occup Med* 2020;70:349–51.
13 Michael M. *Fair society, healthy lives : the Marmot Review : strategic review of health inequalities in England post-2010*. London: UK Government, 2010.
14 Clift AK, Coupland CAC, Keogh RH, *et al*. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *BMJ* 2020;371:m3731.
15 Williamson EJ, Walker AJ, Bhaskaran K. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 2020;584.
16 University of Essex, Institute for social and economic research, UK Data Service. Understanding Society: waves 1-10, 2009-2019 and Harmonised BHPS: waves 1-18, 1991-2009., 2020. Available: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6669#!#10
17 Ayoubkhani D, Nafilyan V, White C. Ethnic minority groups in England and Wales - factors affecting the size and timing of elevated COVID-19 mortality : a retrospective cohort study linking Census and death records. *Medrxiv* 2020.
18 Mathur R, Rentsch CT, Morton CE. Ethnic differences in COVID-19 infection, hospitalisation, and mortality: an OpenSAFELY analysis of 17 million adults in England. *medRxiv* 2020:p. 2020.09.22.20198754.
19 ONS. Analysis of geographic concentrations of COVID-19 mortality over time, England and Wales: deaths occurring between 22 February and 28 August, 2020. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/analysisofgeographicconcentrationsofcovid19mortalityovertimeenglandandwales/deathsoccurringbetween22februaryand28august2020
20 Public Health England. *Disparities in the risk and outcomes from Covid-19*. PHE publications, 2020.
21 Lan F-Y, Wei C-F, Hsu Y-T, *et al*. Work-Related COVID-19 transmission in six Asian countries/areas: a follow-up study. *PLoS One* 2020;15:e0233588.
22 Lan F-Y, Filler R, Mathew S, *et al*. Sociodemographic risk factors for coronavirus disease 2019 (COVID-19) infection among Massachusetts healthcare workers: a retrospective cohort study. *Infect Control Hosp Epidemiol* 2021:1–6.
23 Nafilyan V, Islam N, Ayoubkhani D, *et al*. Ethnicity, household composition and COVID-19 mortality: a national linked data study. *J R Soc Med* 2021;114:182–211.
24 Katikireddi SV, Leyland AH, McKee M, *et al*. Patterns of mortality by occupation in the UK, 1991–2011: a comparative analysis of linked census and mortality records. *The Lancet Public Health* 2017;2:e501–12.
25 Eurostat. *Revision of the European Standard Population — Report of Eurostat's task force*. Luxembourg: Publications Office of the European Union, 2013.
26 Yea-Hung C, Glymour M, Riley A. "Excess mortality associated with the COVID-19 pandemic among Californians 18–65 years of age, by occupational sector and occupation: March through October 2020,". *Medrxiv,* 2021.
27 Lan F-Y, Suharlim C, Kales SN, *et al*. Association between SARS-CoV-2 infection, exposure risk and mental health among a cohort of essential retail workers in the USA. *Occup Environ Med* 2021;78:237–43.
28 ONS. Coronavirus (COVID-19) infection survey: characteristics of people testing positive for COVID-19 in England, 22 February 2021, 2021. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsinthecommunityinengland/characteristicsofpeopletestingpositiveforcovid19inengland22february2021
29 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.
30 Herstein JJ, Degarege A, Stover D, *et al*. Characteristics of SARS-CoV-2 transmission among meat processing workers in Nebraska, USA, and effectiveness of risk mitigation measures. *Emerg Infect Dis* 2021;27:1032–8.
31 Waltenburg MA, Rose CE, Victoroff T, *et al*. Coronavirus disease among workers in food processing, food manufacturing, and agriculture workplaces. *Emerg Infect Dis* 2021;27:243–9.