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

Sarah Beale [1,2], Susan Hoskins[2], Thomas Byrne [1], Wing Lam Erica Fong [1], Ellen Fragaszy [1,3], Cyril Geismar [1,2], Jana Kovar [2], Annalan M D Navaratnam [1,2], Vincent Nguyen [1,2], Parth Patel [1], Alexei Yavlinsky[1], Anne M Johnson [4], Robert W Aldridge [1] and Andrew Hayward [2] on behalf of the Virus Watch Collaborative

Affiliations

1. Centre for Public Health Data Science, Institute of Health Informatics, University College London, UK, NW1 2DA.

2. Institute of Epidemiology and Health Care, University College London, London, UK, WC1E 7HB.

3. Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK, WC1E 7HT.

4. Institute for Global Health, University College London, London, UK, WC1N 1EH.

Abstract

Background: Workers differ in their risk of SARS-CoV-2 infection according to their occupation, but the direct contribution of occupation to this relationship is unclear. This study aimed to investigate how infection risk differed across occupational groups in England and Wales up to October 2021, after adjustment for potential confounding and stratification by pandemic phase.

Methods: Data from 12,182 employed/self-employed participants in the Virus Watch prospective cohort study were used to generate risk ratios for virologically- or serologically-confirmed SARS-CoV-2 infection using robust Poisson regression, adjusting for socio-demographic and health-related factors and non-work public activities. We calculated attributable fractions (AF) amongst the exposed for each occupational group based on adjusted risk ratios (aRR).

Findings: Increased risk was seen in nurses (aRR=1.90 [1.40-2.40], AF=47%); doctors (1.74 [1.26-2.40], 42%); carers (2.18 [1.63-2.92], 54%); teachers (primary = 1.94 [1.44- 2.61], 48%; secondary =1.64, [1.23-2.17], 39%), and warehouse and process/plant workers (1.58 [1.20-2.09], 37%) compared to both office-based professional occupations (reported above) and all other occupations. Differential risk was apparent in the earlier phases (Feb 2020 - May 2021) and attenuated later (June - October 2021) for most groups, although teachers demonstrated persistently elevated risk.

Interpretation: Occupational differentials in SARS-CoV-2 infection risk are robust to adjustment for socio-demographic, health-related, and activity-related potential confounders. Patterns of differential infection risk varied over time, and ongoing excess risk was observed in education professionals. Direct investigation into workplace factors underlying elevated risk and how these change over time is needed to inform occupational health interventions.

Introduction

Notable occupational inequalities in infection risk have emerged during the Coronavirus Disease 19 (COVID-19) pandemic. Research and surveillance data across various global regions have repeatedly indicated elevated risk of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in workers in various essential and/or public-facing industries, such as health and social care, transportation, education, and cleaning and service occupations [1 2 3 4] compared to other workers or the adult population. Occupational differences in the ability to work from home, the frequency and intensity of workplace exposure to other people, environmental features of the workspace, and the implementation of infection control procedures plausibly contribute to differential risk of infection and transmission at work [5 6 7]. However, occupation is intimately linked with other socio-demographic factors such as deprivation, household size, activities outside the workplace and health status, that can compound to influence infection risk[8 9]. Establishing the contribution of work-related exposure to occupational inequalities in infection risk consequently depends on careful consideration of other non-occupational factors.

Few estimates of the effect of occupation on SARS-CoV-2 infection risk or outcomes have comprehensively accounted for sociodemographic confounding beyond age and sex.
Age, sex, geographic factors, education, living conditions, and pre-pandemic health were estimated to account for 70-80% of the effect of occupation on COVID-19 mortality in the UK in 2020[10]. Healthcare, care, and some service and transport occupations (among men) and elementary cleaning and plant workers (among women) demonstrated elevated mortality compared to all other occupations, but the strength of these estimates was greatly attenuated by adjustment. While these findings indicate the importance of comprehensive adjustment, mortality data are strongly affected by clinical risk factors and the impact of work-related factors on differential infection risk cannot therefore be inferred from these findings.
Probability of antigen test positivity differed little across occupations after adjustment for age, sex, region, ethnicity, household composition, deprivation, ability to work from home, use of face coverings at work, and ability to socially distance at work, based on the UK Office for National Statistics (ONS) Coronavirus Infection Survey[11]between early September- early January 2021. However, the inclusion of work-related potential mediators in this analysis precludes disaggregating the impact of occupational and non-occupational factors.

Differential risk across occupations is plausibly influenced by time, due to changes in public health interventions and restrictions - including sectoral closures, social distancing, and infection control in the workplace - as well as fluctuating levels of community transmission across the pandemic and changes in immunity due to infection or vaccination. Preliminary evidence from the UK and Norway suggests that occupational differences in infection risk vary across time, with health [3,12,13,14] and social care workers [12,13] and transport workers [3] demonstrated elevated infection risk during the first pandemic wave and other public-facing occupations including education [12,13], manufacturing [12,13] and

food service as well as transport workers [3] demonstrated elevated risk in the second wave. More recent data including the period of relaxation of pandemic restrictions in the UK are lacking, as are estimates over time comprehensively adjusted for non-occupational factors.

Using data from a prospective community cohort study in England and Wales (Virus Watch) [15], this study aimed to extend current understanding of the direct effect of occupation on SARS-CoV-2 infection risk over time. Specific objectives were: (1) to estimate the relative risk of SARS-CoV-2 infection by occupation across the pandemic, adjusting for socio-demographic and health-related factors and non-work public activities; (2) to investigate whether occupational infection risk differed across pandemic waves; and (3) to estimate the attributable fraction amongst the exposed for different occupations overall and by pandemic wave.

Methods

Ethics Approval

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

Participants

Participants in the current study (*n*=12,182) were an adult sub-cohort of the Virus Watch longitudinal cohort study as of 8/11/2021 (*n*=50,759). Participants were included in the present study if they were (1) ≥16 years, (2) in employment or self-employment and reported their occupation upon study registration, and (3) completed at least one monthly survey between November 2020 and September 2021 concerning their activities across a recent week. Further detail of the full Virus Watch cohort study, including inclusion criteria for the full cohort, can be obtained from the study protocol[15].

Exposure

Occupation was derived based on free-text responses to the Virus Watch baseline survey. Following the protocol recommended by the UK Office for National Statistics (ONS)[16], we performed semi-automatic coding using Cascot Version 5.6.3[17] to assign participants UK Standard Occupational Classification (SOC) 2020 codes[16]. Occupations were then classified into the following groups, which aimed to broadly reflect workplace environments while retaining, as far as possible, ONS-defined occupational skill groupings: administrative and secretarial occupations; healthcare occupations; indoor trade, process & plant occupations; leisure and personal service occupations; managers, directors, and senior officials; outdoor trade occupations; sales and customer service occupations; social care and community protective services; teaching education and childcare occupations; transport and mobile machine operatives; and other professional and associate occupations (broadly office-based professional and associate professional occupations).

Where possible, we also extracted more specific occupational groupings based on three-digit SOC groups for occupations within the essential worker classification[18] and classified by the investigators as public facing/frontline roles. These more detailed occupational groups were included where group sizes exceeded *n*=100 and some SOC groupings were split or combined together to reflect working environment/role, to yield the following included groupings: nurses, doctors, warehouse and process/plant occupations, food preparation and hospitality occupations, teachers (primary), teachers (secondary), teachers (higher), teaching assistants and support occupations, carers, social work and welfare occupations, cleaners, and salespeople/cashiers/shopkeepers.

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

Outcomes

The outcome of interest was binary SARS-CoV-2 infection status (yes/no ever infected) based on any clinical evidence of infection (positive lateral flow (LFT), polymerase chain reaction (PCR), anti-nucleocapsid antibody serological test, or anti-spike antibody serological test in absence of vaccination). Participants were censored after first infection, as susceptibility to reinfections was not the focus of this paper. Please see 'Clinical Outcomes' in the Supplementary Material for further information about clinical data in Virus Watch and how infection status was derived.

Where possible, we attributed results to either the earlier phase of the pandemic (comprising Wave 1 and 2 between February 2020 to May 2021) or the latest phase (comprising Wave 3 from June 2021 to November 2021) based on test date. Waves 1 and 2 were amalgamated into a single phase as it was not possible to attribute specific waves to serology tests conducted during Wave 2, and as mass population testing was largely introduced after the first pandemic wave in England and Wales. Both Waves 1 and 2 included periods of stringent public health restrictions, whilst Wave 3 occurred during the relaxation of public health measures in included regions. Some infections could not be attributed to a particular date as they were based on seropositivity without a prior seronegative result.

Covariates

Where appropriate (see Statistical Analyses), models were adjusted for the following socio-demographic and health-related covariates: age (<30, 30-39, 40-49, 50-59, 60+ years), sex at birth, binary vulnerability status (defined as any condition on the UK NHS/government list of clinically vulnerable conditions [19], obesity, and/or having received an NHS shielding letter), minority ethnicity (White British vs other), geographic region (ONS national region), deprivation based on English or Welsh Indices of Multiple Deprivation Quintile derived from postcode, annual household income (£0-24,900, £25,000-£49,999, £50,000-£75,000, and £75,000+) and household size (excluding participant).

Models were adjusted for non-work public activities based on monthly surveys where participants reported the median number of days that they undertook the following activities across a survey weeks: using transport (using a bus, underground or overground train/tram, taxi, or sharing a car with a non-household member), visiting essential shops, and leisure and social activities (attending the theatre, cinema, concert or sports event; eating in a restaurant, cafe or canteen; going to a bar, pub or club; going to a party; or non-essential shops or personal care services). Responses from November 2020 and February - April 2020 were allocated to Waves 1 and 2, with the second wave used to extrapolate to both early phases of the pandemic. The remaining responses were allocated to Wave 3.

Statistical Analyses

To assess the influence of occupation on SARS-CoV-2 infection risk, we performed Poisson regression with robust standard errors, an established method to estimate risk ratios for binary outcomes [20]. Separate models were conducted for the full pandemic and by wave, with the reference category set as (1) 'Other Professional and Associate Occupations', the largest occupational group in Virus Watch broadly comprising office-based professional occupations (see Supplementary Table S1) with a low absolute infection risk (see Supplementary Table S2), and (2) the full working population of Virus Watch excluding the occupational group under consideration. We identified potential confounders based on a purpose-developed directed acyclic graph (see 'Directed Acyclic Graphs' in Supplementary Materials), with models presented unadjusted and fully adjusted for the following confounders according to our DAG: age, sex, ethnicity, region, deprivation and household size, vulnerability status, and non-work public activities. Vaccination status was not directly included in models due to the inclusion of determining variables (i.e., age, health status, and occupation in the case of vaccination due to UK protocols. No evidence of multicollinearity emerged based on variance inflation factors for any model. We also performed a sensitivity analysis limited to participants who had undergone serological testing (n=7372) to address potential differential access and testing behaviour for virological/antigen testing across occupations; it was only possible to perform this analysis on broad occupational groups across the full study period, and not for specific occupations or by wave due to limited statistical power.

Based on the fully-adjusted models, we calculated attributable fractions for the exposed subpopulations (AFs) using the *punaf* programme in Stata Version 16 [21]. Attributable fractions range from -∞ to 1, with negative values indicating a protective effect and positive values indicating a harmful effect [21]; while negative values are often transformed to express cases prevented in the unexposed group, we did not transform estimates in order to facilitate comparison by leaving all estimates with the same denominator.

Missing data were limited for all included sociodemographic variables (0-6%) and complete cases were included in the final analyses. We conducted a missing data sensitivity analysis by applying multivariate imputation by chained equations (*mice* package in R Version 4.0.3[22]) with 5 datasets with 50 iterations per dataset to socio-demographic variables and re-testing models.

Results

Selection of participants into the current study based on inclusion criteria presented in Figure 1, with demographic features of included participants (*n*=12,182) reported in Table 1.

Occupational Group and Infection Risk

Absolute risk of infection by occupational risk is reported in Supplementary Table S2, and ranged from 11% in transport workers to 21% in healthcare workers across the full pandemic period covered by the study.

Across the full pandemic period covered, healthcare (adjusted risk ratio (aRR)=1.71, 1.46-2.01; adjusted Attributable Fraction (AF) = 42%, 32-50%), indoor trade/process/plant (aRR=1.33, 1.11-1.60; AF=25% etc, 10-37), leisure and personal service (aRR=1.26, 1.01-1.56; AF=20, 1-36%), social care and community protective service (aRR=1.42, 1.16-1.73; AF=29%, 14-42%), and teaching/education/childcare occupations (aRR=1.42, 1.22-1.64; AF=29%, 18-39%) demonstrated elevated infection risk compared to Other Professional and Associate Occupations (Figure 2; see Supplementary Table S3 for adjusted AFs). When limited to participants who underwent serological testing (Supplementary Figure 2), similar groups demonstrated elevated infection risk with the addition of administrative and secretarial occupations (aRR=1.43, 1.09-1.86).

In Waves 1 and 2, administrative and secretarial (aRR=1.24, 1.02-1.51; AF=19%, 2-34%), healthcare (aRR=2.28, 1.90-2.74; AF=56%, 47-63%), indoors trades/process/plant (aRR=1.56, 1.25-1.96; AF=36%, 20-49%), sales and customer service (aRR=1.10-1.85; AF=30%, 9-46%), social care and community protective services (aRR=1.65, 1.30-2.10; AF=39%, 23-52%), and teaching/education/childcare occupations (aRR=1.45, 1.20-1.75; AF=31%, 16-43%) demonstrated elevated risk, while only teaching/education/childcare occupations remained at elevated risk in Wave 3 (aRR=1.41, 1.07-1.87; AF=29%, 7-46%). Across all models, limited effects of adjustment for sociodemographic, health-related and non-workplace activities were observed (Figure 2). Similar results were obtained in sensitivity analyses including imputed sociodemographic data (Supplementary Figure 3a).

Across the full pandemic period, healthcare workers (aRR=1.48, 1.29-1.70; AF=32%, 22-41%) and teaching/education/childcare workers (aRR=1.21, 1.06-1.38; AF=17%, 6-27%) also demonstrated elevated risk of SARS-CoV-2 compared to the working population outside of these occupations (Figure 3; see Supplementary Table S2 for adjusted AFs). Workers in Other Professional and Associate occupations conversely demonstrated lower risk (aRR = 0.78, 0.71-0.87). Similar results were obtained in sensitivity analyses including imputed sociodemographic data (Supplementary Figure 3b) and when limited to participants who underwent serological testing (Supplementary Figure

2), although confidence intervals for teaching/education/childcare workers included one in the later sensitivity analysis.

In Waves 1 and 2, healthcare workers (aRR=1.85, 1.58-2.17; AF=46%, 37-54%), indoor trade/process/plant workers (aRR=1.26,1.03-1.54, AF= 20%, 2-35%) and social care and community protective service workers (aRR=1.26, 1.01-1.57; AF=21%, 1-46%) demonstrated elevated risk of infection and workers in other professional and associate occupations lower risk of infection (aRR=0.69, 0.61-0.79; AF=-45%, -65 - -27%). Conversely, only teaching/education/childcare workers demonstrated excess risk in Wave 3 (aRR=1.50, 1.17-1.91; AF=33%, 15-48%). As above, adjustment for non-work factors had a relatively minimal impact on the strength and direction of effect estimates (Figure 3). Results were similar in corresponding missing data sensitivity analyses, although confidence intervals for the indoor trade/process/plant workers and social care and community protective service workers in Waves 1 and 2 included the value one (Supplementary Figure 3b).

Specific Frontline Occupations

Absolute risk of infection for specific frontline occupations is reported in Supplementary Table S4; carers and primary teachers demonstrated the highest absolute risk (25%) across the full pandemic period.

The following frontline occupational groups demonstrated elevated infection risk compared to 'Other professional and associate' occupations: nurses (RR=1.90, 1.40-2.40; AF=47%, 33-58%); doctors (1.74, 1.26-2.40; AF=42%, 20-58%); carers (2.18, 1.63-2.92; AF=54%, 39-66%); primary teachers (RR=1.94, 1.44- 2.61; AF=48%, 31-62%); secondary teachers (RR=1.64, 1.23-2.17; AF=39%, 19-54%); teaching assistants and support staff (RR=1.40, 1.06-1.85; AF=29%, 6-46%); and warehouse and process/plant workers (RR=1.58, 1.20-2.09; AF=37%, 17-52%) (Figure 4; attributable fractions in Supplementary Table S5). All of these occupational groups demonstrated elevated risk during Waves 1 and 2, along with food preparation and distribution workers (RR=1.73, 1.07-2.81; AF=42%, 6-64%). In Wave 3 only primary teachers demonstrated clear evidence of elevated risk (RR=1.92, 1.09-3.38; AF=48%, 8-70%). Similar results were obtained in sensitivity analysis with imputed sociodemographic data (Supplementary Figure 4a).

Compared to the rest of the working population, the following frontline occupational groups demonstrated elevated infection risk: nurses (RR=1.60, 1.29-1.98; AF=37%, 23-50%); doctors (RR=1.54, 1.12, 2.11; AF=35%, 11-53%); carers (RR=1.81, 1.37-2.38; AF=45%, 27-58%); primary teachers (RR=1.64, 1.23-2.18; AF=39%, 19-54%); secondary teachers (RR=1.37, 1.04-1.80; AF=27%, 4-44%); and warehouse and process/plant workers (RR=1.32, 1.01-1.71; AF=24%, 1-42%) (Figure 5; attributable fractions in Supplementary Table S4). All groups - with the exception of secondary teaching - demonstrated evidence of elevated risk in Waves 1 and 2 beyond variation expected by chance, while primary teachers demonstrated similar evidence of elevated risk in Wave 3

(RR=2.11, 1.22-3.65; AF=53%, 18-73%). Point estimates for secondary teachers indicated elevated risk in Waves 1 and 2 and Wave 3, but confidence intervals included the value one for specific time periods. Similar results were obtained in imputed sensitivity analyses, although confidence intervals for warehouse and process/plant workers included the value one (Supplementary Figure 4b).

Discussion

Key Findings and Interpretation

This study found persistent occupational differences in SARS-CoV-2 infection risk after comprehensive adjustment for non-work-related confounding, including socio-demographic and health-related factors and non-work social activities. Compared to Other Professional and Associate occupations - the largest occupational group in the sample with the lowest infection risk - workers in healthcare, teaching/education/childcare, indoor trade and process/plant, and leisure and personal service occupations demonstrated elevated overall infection risk. In these groups, belonging to their occupation compared to the less risky group accounted for between 20% (for leisure and personal service) to 42% (for healthcare workers) of their infection risk. Healthcare and teaching/education/childcare occupations also demonstrated elevated overall infection risk compared to the remainder of the working population. Most at-risk occupations demonstrated elevated risk in the earlier pandemic phase (Waves 1 and 2) that was later attenuated (Wave 3), with the exception of teaching/education/childcare occupations for whom risk remained elevated.

Where we had sufficient numbers, we also investigated infection risk for specific frontline occupational groups. Nurses, doctors, carers, teachers (primary and secondary) and warehouse and process/plant occupations demonstrated elevated risk compared to Other Professional and Associate occupations and the rest of the working population and across the study period. Belonging to their occupation compared to the less risky group accounted for between 37% (for warehouse/process/plant workers) to 54% (in carers) of their risk of infection. Patterns of risk by pandemic phase were similar to above, with clear evidence of elevated risk in Wave 3 emerging only for primary teachers; however, these findings may have been impacted by lack of power to detect modest effects in some groups.

Elevated infection risk in occupational groups with limited ability to work from home and those involving exposure to patients and/or the public echoes findings from the previous studies with more limited adjustment for potential confounding [1 2 3 4]. Across all analyses in the current study, adjustment for sociodemographic and health-related factors and non-work activities had limited impact on estimates. This result differs markedly to prior analysis of occupational differences in COVID-19 mortality [10], where adjustment for socio-demographic and health-related factors substantially reduced the effect of occupation. Occupation plausibly shapes SARS-CoV-2 exposure - and consequently infection risk - by influencing workers' ability to work from home, practise social distancing at work, work in well-ventilated environments, and access appropriate personal protective equipment. The

specific mechanisms and relative contribution of different mitigating factors are likely to differ considerably by occupation, and are an important area for future research. Conversely, clinical factors that influence risk of severe morbidity and mortality once infected may differ across occupations, however the direct effect of occupation itself on severity of infection is likely to be more limited.

Changing patterns of differential infection risk by pandemic phase are likely to be multifactorial. Immunity-related factors that reduce the population of susceptible workers within a given occupation are likely to be important, and include prior infection in early phases of the pandemic, prioritization of some occupational groups (i.e. health and care workers [10,23]) for vaccination, and potential differences in the speed and overall uptake of vaccination between occupations [24]. The removal of remaining public health restrictions in Wave 3 may also have reduced differential risk by increasing overall contact rates and networks, and probability of transmission outside of work due to the increasing range of potential venues for exposure at a time of persistently high community infection rates and reduced mitigations. Direct investigation into potential mediators of this phase effect was beyond the scope of this study, and is warranted to better understand the processes shaping occupational infection risk. Relatedly, investigation into effective mitigation for the ongoing elevated infection risk in teachers is recommended both to address occupational inequalities and to reduce disruption in education settings.

Strengths and Limitations

Strengths of this study include the large and diverse cohort that enabled investigation of infection risk from multiple study-derived and linked sources including both symptomatic testing and serology over multiple pandemic phases. Detailed information around participants' demographic characteristics and activities over time allowed adjustment for a comprehensive series of potential confounders, including non-work-related public activities, informed by a directed acyclic graph.

However, the study has several important limitations. The Virus Watch cohort is demographically diverse but not representative of the UK population, with underrepresentation of some occupational categories limiting the ability to investigate differential risk across all occupational categories. Potential confounders, such as deprivation, are challenging to measure and residual confounding cannot be excluded. Non-work public activities were inferred from self-reported activities across a given survey week, and may not have been an accurate reflection of participants' activity patterns across the entire relevant time period. Furthermore, social and leisure activities may have included work for some occupational groups (e.g. leisure and personal service occupations) but could not be disaggregated; however, the limited effect of adjustment in these models indicates that this was unlikely to be a major source of bias. Occupation was measured in broad categories, and only some specific occupations could be investigated due to small subsample sizes. Relatedly, the number of infections within a given pandemic phase was small for some frontline subsamples. Overall estimates of risk by occupational sector may be driven by particularly risky roles with considerable exposure [8], and further investigation

into specific occupations is recommended. Additionally, inclusion of multiple test types to indicate SARS-CoV-2 positivity allowed for potential detection of asymptomatic or previously untested cases through serology, and detection of early cases through linkage. However, issues impacting the uptake and usage of each test type, including differential access to some tests in given phases of the pandemic, self-selection bias, and compliance with testing instructions may have affected estimates and are difficult to delineate. Notably, results may be influenced by differential testing behaviour between occupations. For example, health care workers undertake regular occupational testing which may lead to an overestimation of their relative risk of infection. However, a sensitivity analysis constrained to those participants who underwent serological testing was not subject to such testing behaviour bias and demonstrated results similar to the main analyses.

Conclusions

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

Funding

This work was supported by funding from the PROTECT COVID-19 National Core Study on transmission and environment, managed by the Health and Safety Executive on behalf of HM Government. The Virus Watch study is supported by the MRC Grant Ref: MC_PC 19070 awarded to UCL on 30 March 2020 and MRC Grant Ref: MR/V028375/1 awarded on 17 August 2020. The study also received $15,000 of Facebook advertising credit to support a pilot social media recruitment campaign on 18th August 2020. This study was also supported by the Wellcome Trust through a Wellcome Clinical Research Career Development Fellowship to RA [206602]. SB and TB are supported by an MRC doctoral studentship (MR/N013867/1). The funders had no role in study design, data collection, analysis and interpretation, in the writing of this report, or in the decision to submit the paper for publication.

Conflicts of interest

AH serves on the UK New and Emerging Respiratory Virus Threats Advisory Group. AJ and AH are members of the COVID-19 transmission sub-group of the Scientific Advisory Group for Emergencies

(SAGE). AJ is Chair of the UK Strategic Coordination of Health of the Public Research board and is a member of COVID National Core studies oversight group.

Data availability

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

References

1 D. Phuong Do RF. U.S. frontline workers and COVID-19 inequities. *Prev Med* 2021; **153**: 106833.

2 Fogh K, Strange JE, Scharff BF, *et al.* Testing Denmark: A Danish nationwide surveillance study of COVID-19. *medRxiv* 2021; 2021.08.10.21261777.

3 Magnusson K, Nygård K, Methi F, Vold L, Telle K. Occupational risk of COVID-19 in the first versus second epidemic wave in Norway, 2020. *Eurosurveillance* 2021; **26**: 2001875.

4 de Gier B, de Oliveira Bressane Lima P, van Gaalen RD, *et al.* Occupation- and age-associated risk of SARS-CoV-2 test positivity, the Netherlands, June to October 2020. *Eurosurveillance* 2020; **25**: 2001884.

5 Zhang M. Estimation of differential occupational risk of COVID-19 by comparing risk factors with case data by occupational group. *American Journal of Industrial Medicine.* 2021; **64**: 39–47.

6 Cabinet Office. COVID-19 and occupation: position paper 48. https://www.gov.uk/government/publications/covid-19-and-occupation-iiac-position-paper-48/covid-19-and-occupation-position-paper-48 (accessed Dec 14, 2021).

7 Office for National Statistics. Which occupations have the highest potential exposure to the coronavirus (COVID-19)? 2020; https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/whichoccupationshavethehighestpotentialexposuretothecoronaviruscovid19/2020-05-11 (accessed Dec 14, 2021).

8 Cevik M, Baral SD. Networks of SARS-CoV-2 transmission. *Science* 2021; **373**: 162–3.

9 Pearce N, Rhodes S, Stocking K, *et al.* Occupational differences in COVID-19 incidence, severity, and mortality in the United Kingdom: Available data and framework for analyses. *Wellcome Open Research* 2021; **6**. DOI:10.12688/wellcomeopenres.16729.1.

10 Nafilyan V, Pawelek P, Ayoubkhani D, *et al.* Occupation and COVID-19 mortality in England: a national linked data study of 14.3 million adults. *medRxiv* 2021; 2021.05.12.21257123.

11 Sutherland E, Willis, Z, & Fordham E. Coronavirus (COVID-19) Infection Survey - Office for National Statistics. 2021; published online Feb 22.

https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsinthecommunityinengland/characteristicsofpeopletestingpositiveforcovid19inengland22february2021 (accessed Dec 14, 2021).

12 Health and Safety Executive. Technical summary of data - Coronavirus (COVID-19) disease reports made by employers 10 April 2020. https://www.hse.gov.uk/statistics/coronavirus/april-to-july-2020-technical-summary-of-data.htm (accessed Dec 14, 2021).

13 Agius R. Disease and death from work: RIDDOR and covid-19. *Occup Med* DOI:10.1093/occmed/kqaa155.

14 Riley S, Eales O, Walters CE, *et al.* REACT-1 round 7 interim report: fall in prevalence of swab-positivity in England during national lockdown. *medRxiv* 2020; 2020.11.30.20239806.

15 Hayward A, Fragaszy E, Kovar J, *et al.* Risk factors, symptom reporting, healthcare-seeking behaviour and adherence to public health guidance: protocol for Virus Watch, a prospective community cohort study. *BMJ Open* 2021; **11**. DOI:10.1136/bmjopen-2020-048042.

16 Office for National Statistics. SOC 2020. https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2020 (accessed Dec 14, 2021).

17 Cascot: Computer Assisted Structured Coding Tool. https://warwick.ac.uk/fac/soc/ier/software/cascot/ (accessed Dec 14, 2021).

18 UK Health Security Agency. Essential workers prioritised for COVID-19 testing; 2021; published online Feb 25. https://www.gov.uk/guidance/essential-workers-prioritised-for-covid-19-testing (accessed Dec 14, 2021).

19 UK National Health Service. Who is at high risk from coronavirus (COVID-19). https://www.nhs.uk/conditions/coronavirus-covid-19/people-at-higher-risk/who-is-at-high-risk-from-coronavirus/ (accessed Dec 14, 2021).

20 Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004; **159**. DOI:10.1093/aje/kwh090.

21 Newson R. PUNAF: Stata module to compute population attributable fractions for cohort studies. *Statistical Software Components* 2015; published online Sept 4. https://ideas.repec.org/c/boc/bocode/s457193.html (accessed Dec 14, 2021).

22 Van Buuren, S. Package 'mice': Multivariate Imputation by Chained Equations. https://cran.r-project.org/web/packages/mice/mice.pdf (accessed Dec 14, 2021).

23 Public Health England. COVID-19 vaccination first phase priority groups. 2020; published online Dec 7. https://www.gov.uk/government/publications/covid-19-vaccination-care-home-and-healthcare-settings-posters/covid-19-vaccination-first-phase-priority-groups (accessed Dec 14, 2021).

24 Nafilyan V, Dolby T, Finning K, *et al.* Differences in COVID-19 vaccination coverage by occupation in England: a national linked data study. *medRxiv* 2021; 2021.11.10.21266124.

Tables

Table 1. Characteristics of Study Participants

	N = 12,182[1]
Occupation	
Administrative & Secretarial	1,559 (13%)
Healthcare	923 (7.6%)
Indoor Trades, Process & Plant	820 (6.7%)
Leisure & Personal Service	578 (4.7%)
Managers, Directors & Senior Officials	1,029 (8.4%)
Other Professional & Associate	4,102 (34%)
Outdoor Trades	259 (2.1%)
Sales & Customer Service	632 (5.2%)
Social Care & Community Protective Services	630 (5.2%)
Teaching, Education & Childcare	1,380 (11%)
Transport & Mobile Machine	270 (2.2%)
Age	
<30	948 (7.8%)
30-39	1869 (15.3%)
40-49	2584 (21.2%)
50-59	3606 (29.6%)
60+	3175 (26.1%)
Sex	
Female	6,690 (55%)
Male	5,466 (45%)
Unknown[2]	26 (0.2%)
Ethnicity	
White British	10,014 (82%)
White Other	1,201 (9.9%)
South Asian	439 (3.6%)
Other Asian	118 (1.0%)
Black	113 (0.9%)
Mixed/Multiple Ethnicity	202 (1.7%)
Other Ethnicity	62 (0.5%)
Unknown[2]	33 (0.3%)
Chronic Condition and/or Obesity	6,489 (53%)
Index of Multiple Deprivation Quintile	
1	1,124 (9.3%)
2	2,126 (18%)
3	2,489 (21%)
4	3,058 (25%)
5	3,266 (27%)
Unknown[2]	119 (1.0%)
Household Income	
£0-£24,999	1,810 (16%)
£25,000-£49,999	3,598 (31%)
£50,000-£74,999	2,780 (24%)
£75,000+	3,235 (28%)

Unknown[2]	759 (6.2%)
Household Size	
1 person	2,479 (20%)
2 people	5,528 (45%)
3 people	1,946 (16%)
4 people	1,718 (14%)
5 people	423 (3.5%)
6 people	88 (0.7%)
Region	
East Midlands	1,050 (8.7%)
East of England	2,469 (20.5%)
London	2,166 (18.0%)
North East	554 (4.6%)
North West	1,276 (10.6%)
South East	2,375 (19.7%)
South West	809 (6.7%)
Wales	235 (1.9%)
West Midlands	581 (4.8%)
Yorkshire and The Humber	548 (4.5%)
Unknown[2]	119 (1.0%)

[1]n (%) of available; [2] n (%) of total

Figures

Figure 1. Flow Diagram of Participant Eligibility

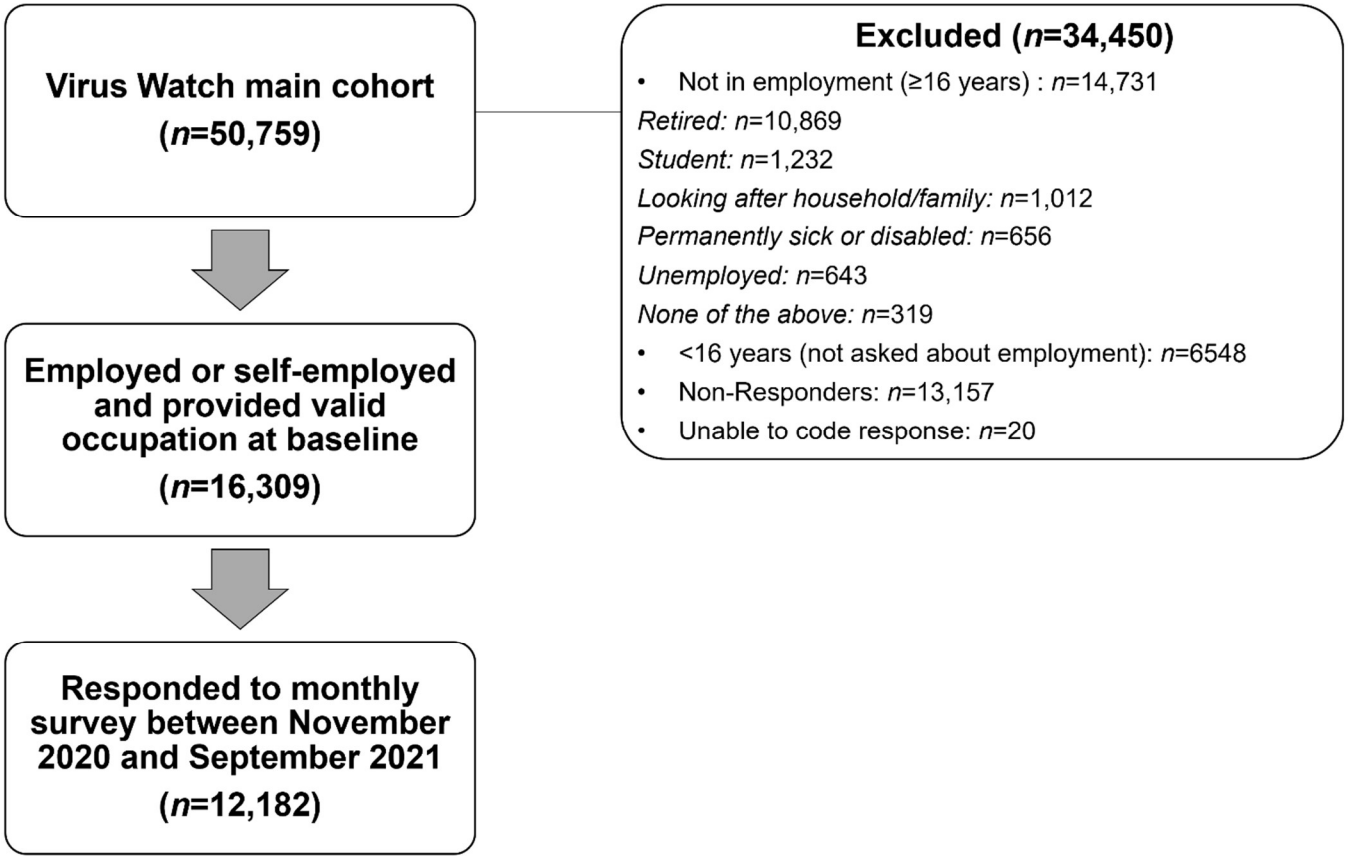

Figure 2. Risk Ratios by Occupational Group (versus Other Professional and Associate)

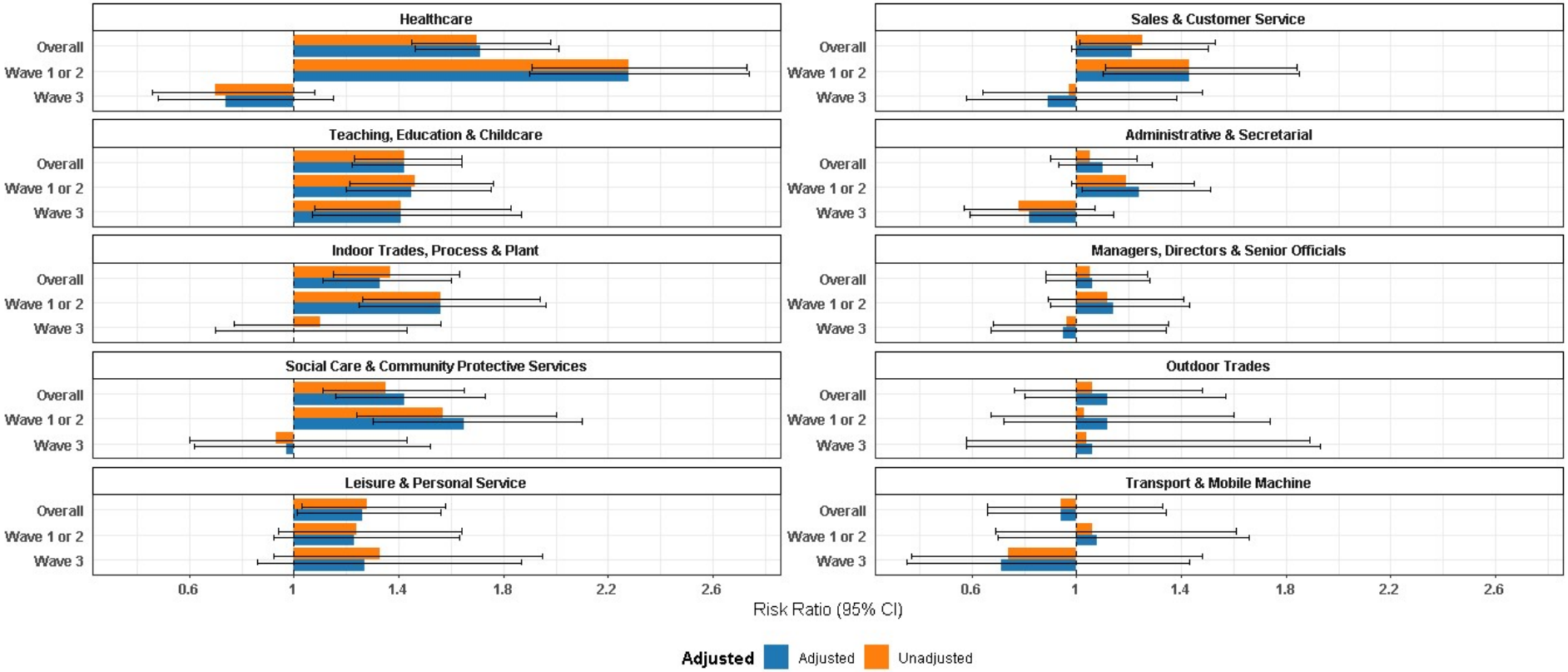

Figure 3. Risk Ratios by Occupational Group (versus Working Population)

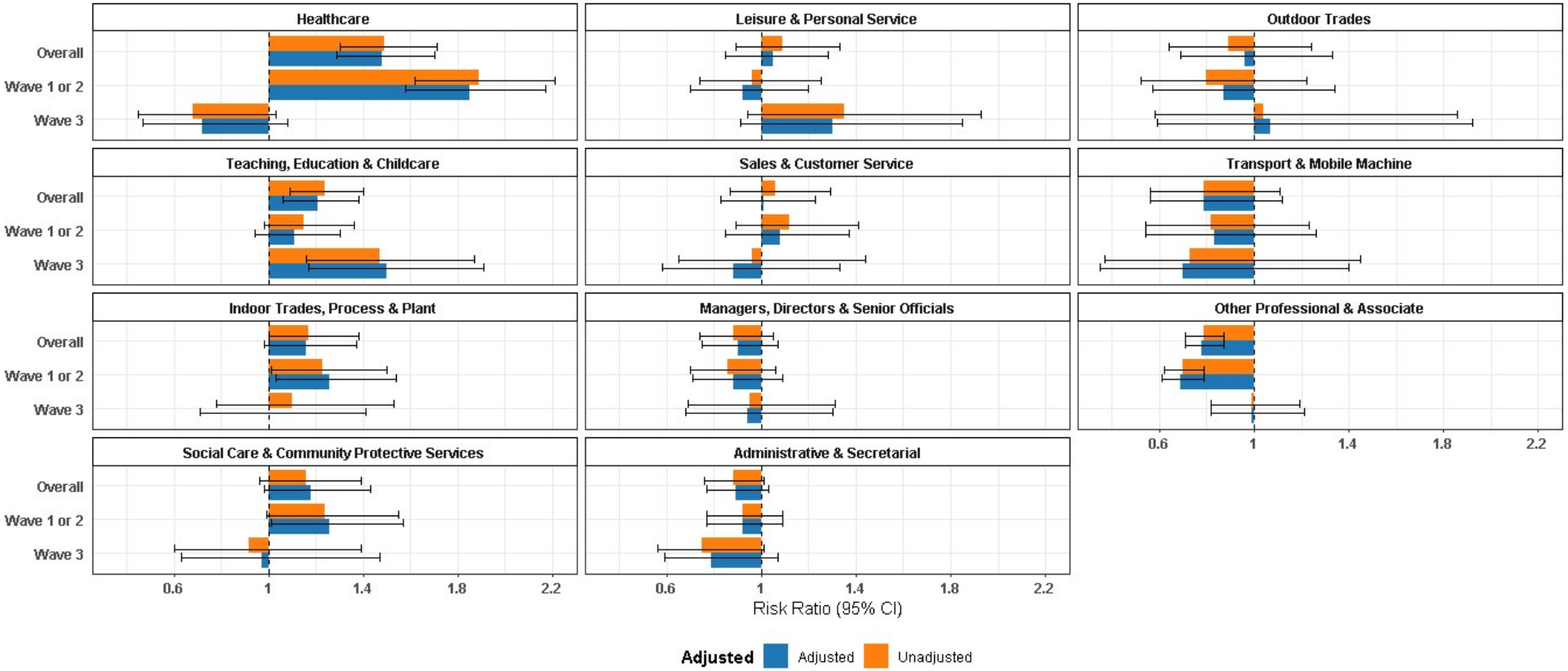

Figure 4. Risk Ratios for Frontline Occupations (versus Other Professional and Associate)

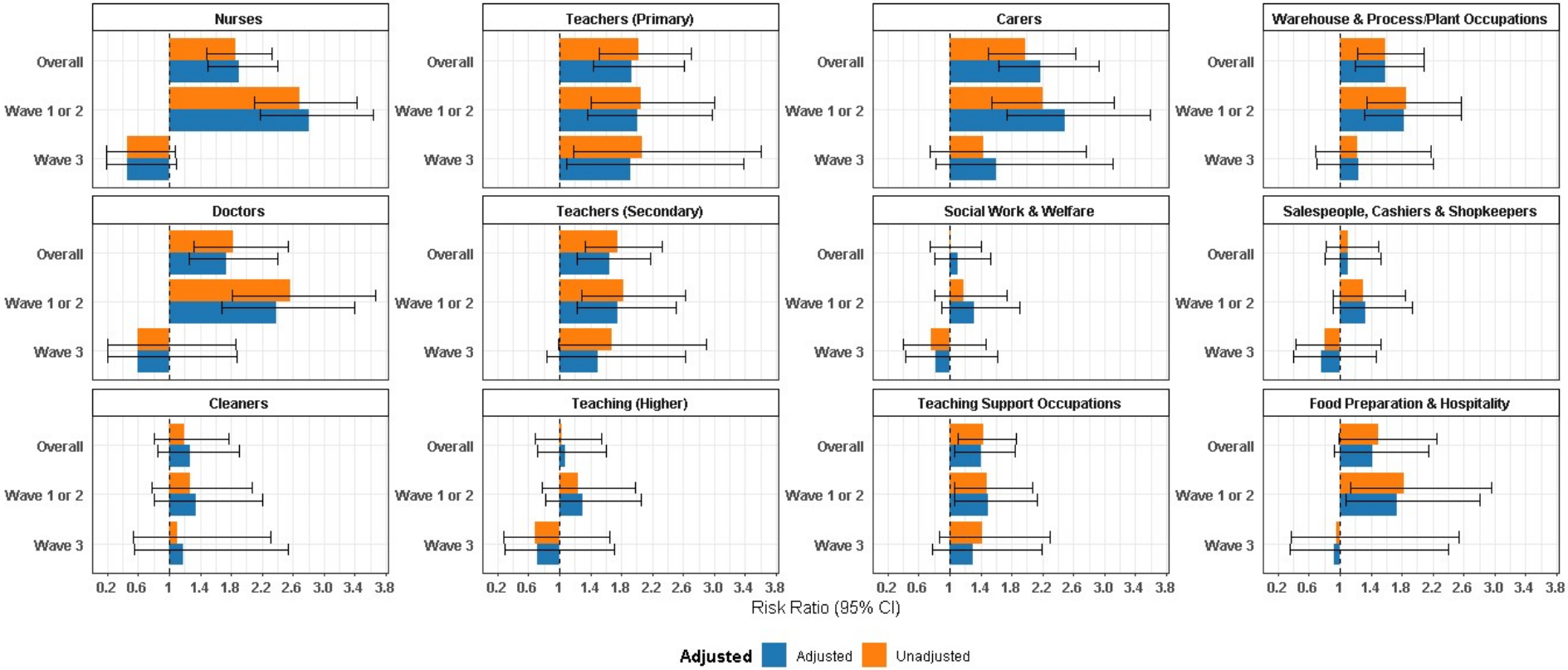
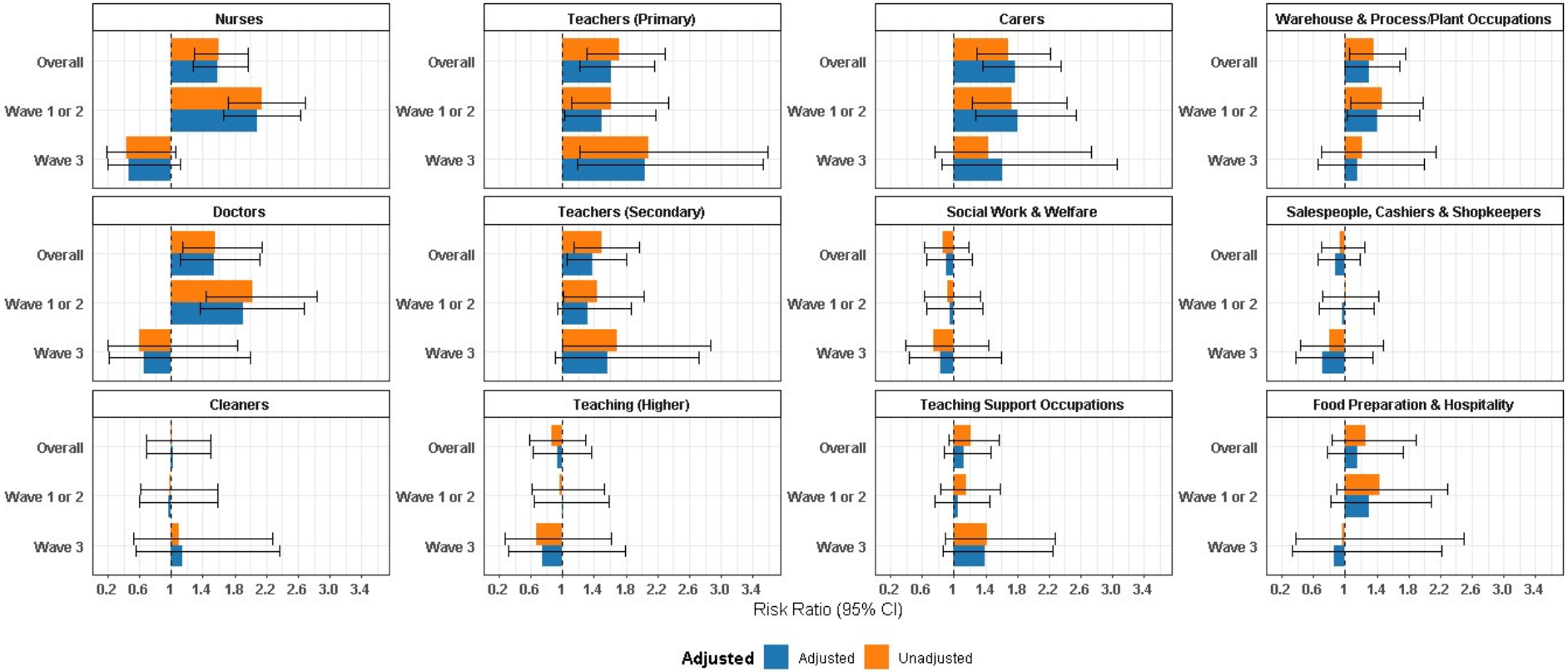
Figure 5. Risk Ratios for Frontline Occupations (versus Working Population)