medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Relative contribution of leaving home for work or education, transport, shopping and other activities on risk of acquiring COVID-19 infection outside the household in the second wave of the pandemic in England and Wales

Hoskins S, Beale S, Aldridge RW, Navaratnam AMD, Smith C, French CE, Yavlinsky A, Nguyen V, Byrne T, Kovar J, Fragaszy E, Fong WLA, Geismar C, Patel P, Johnson AM, Hayward AC - on behalf of the Virus Watch Collaborative

Affiliations

Institute of Health Informatics, University College London, UK (AMDN, MS, VN, SB, IB, TB, EF, WF, CG, PP, AA, RWA); Institute of Epidemiology and Health Care, University College London, London, UK (JK, SB, ACH); Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK (EF).

Abstract

Background: With the potential for and emergence of new COVID-19 variants, such as the reportedly more infectious Omicron, and their potential to escape the existing vaccines, understanding the relative importance of which non-household activities increase risk of acquisition of COVID-19 infection is vital to inform mitigation strategies.

Methods: Within an adult subset of the Virus Watch community cohort study, we sought to identify which non-household activities increased risk of acquisition of COVID-19 infection and which accounted for the greatest proportion of non-household acquired COVID-19 infections during the second wave of the pandemic. Among participants who were undertaking antibody tests and self-reporting PCR and lateral flow tests taken through the national testing programme, we identified those who were thought to be infected outside the household during the second wave of the pandemic. We used exposure data on attending work, using public or shared transport, using shops and other non-household activities taken from monthly surveys during the second wave of the pandemic. We used multivariable logistic regression models to assess the relative independent contribution of these exposures on risk of acquiring infection outside the household. We calculated Adjusted Population Attributable Fractions (APAF - the proportion of non-household transmission in the cohort thought to be attributable to each exposure) based on odds ratios and frequency of exposure in cases.

Results: Based on analysis of 10475 adult participants including 874 infections acquired outside the household, infection was independently associated with: leaving home for work (AOR 1.20 (1.02 – 1.42) p=0.0307, APAF 6.9%); public transport use (AOR for use more than once per week 1.82 (1.49 – 2.23) p<0.0001, APAF for public transport 12.42%); and shopping (AOR for shopping more than once per week 1.69 (1.29 – 2.21) P=0.0003, APAF for shopping 34.56%). Other non-household activities such as use of hospitality and leisure venues were rare due to restrictions and there were no significant associations with infection risk.

Conclusions: A high proportion of the second wave of the pandemic was spent under conditions where people were being advised to work from home where possible, and to minimize exposure to shops, and a wide range of other businesses were subject to severe restrictions. Vaccines were being rolled out to high-risk groups. During this time, going to work was an important risk factor for infection but public transport use likely accounted for a lot of this risk. Only a minority of the cohort left home for work or used public or shared transport. By contrast, the majority of participants visited shops and this activity accounted for about one-third of non-household transmission.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

The COVID-19 pandemic has led to levels of hospitalisation and mortality that are unprecedented in recent history. Governments around the world have imposed strict impositions on social mixing to control the virus. The relative importance of these restrictions is challenging to estimate as empirical evidence to date relies on ecological studies where it can be difficult to disentangle the effects of multiple interventions introduced at different stages of the pandemic in different countries and regions. A review of such empirical studies from the first wave of the pandemic suggests that school and workplace closures, closure of businesses and banning of public events had the greatest impact on transmission, especially when introduced early. (1)

The effectiveness of restrictions depends on the amount of transmission that occurs in different settings and whilst transmission can take place in any setting the relative importance of different non-household settings and activities to spread of infection is poorly understood. (2)

Occupation is known to be an important predictor of COVID-19 infection and mortality risk, with a high proportion of the differential risk between occupations likely to be related to differential ability to work from home during periods of intense COVID-19 transmission. (3)

Transmission on public transport has also been shown to occur but the relative importance of this compared to other exposures is unknown and since transport is used as a means of travelling to undertake other activities, it has proved challenging to untangle the risk from that of other out of home exposures. (4)

Although it is clear that the importance of different venues such as hospitality, retail and leisure on population infection rates depends on both the likelihood of transmission occurring within a particular environment and the frequency with which people visit that setting, the relative importance of these settings has proved difficult to assess.(5) During periods of intense control essential shops have remained open but the amount of transmission occurring in such settings is unknown.

We aimed to understand the relative importance of different activities and settings in the transmission of COVID-19 in England and Wales during the second wave of the pandemic, a period with a relatively intensive range of non-pharmaceutical interventions (NPI's) including advice to work from home where possible, closure of a range of non-essential businesses such as hospitality and leisure venues and restrictions on social gatherings.

Methods

The analyses are based on the Virus Watch Cohort, the detailed methodology of this cohort is described elsewhere (6). Briefly, the study recruits whole households with detailed baseline information, weekly surveys of symptoms and self-reporting of positive COVID-19 tests

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

(PCR or lateral flow) conducted through the national tracing programme, linkage to the national testing data set, and monthly questionnaires on contact and activity patterns. A subset of the adult cohort have also undertaken antibody testing through venous blood draws since October 2020 and monthly finger-prick testing since March 2021.

Within an adult subset of the Virus Watch community cohort study, we identified a subcohort of participants who completed two monthly behavioural surveys during the second wave of the pandemic (completed during the periods 1/12/2020 – 10/12/2020 and 17/02/2021 – 28/02/2021). These asked about the frequency of going to work, using public transport, going to retail settings and visiting other non-household settings in the week prior to the survey. These were averaged to give the average weekly frequency of these activities. We also asked about the number of known close contacts outside the household during these weeks, where close contact was defined as being within 2m of someone for more than 15 minutes.

We further restricted the cohort to those who were taking part in the antibody test study. Anyone who tested positive for nucleocapsid antibody due to likely infection versus vaccination between 01/10/2020 and 01/05/2021 was considered to have been infected in the second wave of the pandemic unless they had previously reported a positive PCR or lateral flow test prior to the second wave. Participants who tested positive for antibodies on venous blood samples taken before the second wave were excluded. Anyone testing PCR or lateral flow positive during the second wave (01/10/2020 - 01/05/2021) was also considered to have been infected during the second wave.

In order to focus analyses on risk factors for non-household transmission we further restricted cases to those which were the only case in the household (no other antibody or PCR or lateral flow cases in the household) and, in households with more than one PCR/lateral flow confirmed case, to i) the first case in the household based on the earliest date of symptom onset in the house, or ii) where there were multiple positive people in the household reporting the same symptom onset date, all cases were included as co-primaries, or iii) if no symptom data were available, based on the earliest date of PCR or lateral flow test in the house using dates from the National Testing data, or if National testing date data were not available, using the Virus Watch given date or imputed middle of the week date when the test was self-reported. Where we could not identify who was the likely first case in the household these infections were excluded from the analyses.

We undertook univariate analyses comparing the proportion with evidence of infection acquired outside the household, according to different weekly frequency of going to work, using public or shared transport, using retail venues and other non-household contacts. We conducted multivariate logistic regression models including these variables as well as age, gender, vaccination status, social deprivation, region and ethnicity. For activities associated with risk of non-household transmission we calculated stratum specific univariate and multivariate population attributable fractions during the second wave of the pandemic. We used the formula PAF = p * (1-1/Relative Risk) where p=proportion of those with COVID-19

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

acquired outside the household who had the exposure of interest and where adjusted and unadjusted odds ratios were taken as proxies of relative risk (based on the rare outcome assumption).

Results

Of the 10,858 Virus Watch participants who completed the two monthly surveys and were members of the antibody cohort, 1,257 (12%) were identified as having had Covid-19 (343 were identified by either PCR or lateral flow test and 914 by likely infection-induced antibody results alone). Of all Covid-19 cases identified via PCR or lateral flow test, we estimate up to 18% of these may be due to household transmission based on the timing of illnesses and swabs in other household members (Table 1).

The non-household transmission study therefore included 10475 participants; 874 COVID-19 cases identified through PCR, lateral flow and/or positive antibody tests who were thought to have been infected outside the household during the second wave of the pandemic and 9601 uninfected participants (Table 1). The breakdown of the cohort by age, gender, index of multiple deprivation, region, ethnicity and vaccination status are shown in table 2 with univariate analyses of the risk of infection acquired outside the household in each group. The highest levels of infection were seen in working-aged adults, women, poorer areas, non-white ethnic groups, London and the unvaccinated.

Supplementary table S1 shows the detailed breakdown of public or shared transport exposures in relation to risk of infection. On univariate analysis, all forms of public or shared transport, with the exception of using an airplane which was rare, were associated with an increased risk of COVID-19 infection. Table S2 shows the detailed breakdown of non-work, non-public or shared transport use and out of household activities and associations with non-household acquired infections. The only significant association during this period was with essential retail, although exposure to other settings such as leisure and hospitality was rare during this period.

Table 3 shows the relative impact of leaving home to go to work or education, using public or shared transport, retail and other non-household activities on the risk of infection acquired outside the household. Models are adjusted for all variables in the table in addition to age, gender, ethnicity, region, vaccination and index of multiple deprivation quintile. Table 3 also shows the unadjusted and adjusted population attributable fractions for leaving home for work, public or shared transport and retail.

The odds ratio for leaving home for work or education at least once per week was 1.72 but this association was considerably weakened after controlling for public transport use and other risk factors - adjusted odds ratio 1.20 (1.02 – 1.42) p=0.0307. Across the cohort leaving home for work accounted for 18% of non-household acquired infections (PAF) but this was reduced to 7% after controlling for transport use and other variables.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The odds ratio for using public or shared transport more than once per week compared to no usage was 2.35 but this was significantly reduced after controlling for going to work or education and other risk factors (adjusted OR 1.82 (1.49 – 2.23) p<0.0001). Across the cohort public or shared transport use accounted for 16% of infections acquired outside the household, though this was reduced to 12% after controlling for going to work and other variables.

The odds ratio for using shops more than once per week compared to no usage was 2.01 but reduced after controlling for other variables (adjusted OR 1.69 (1.29 – 2.21) P=0.0003). Across the cohort, shopping accounted for 42% of infections acquired outside the household and 35% after controlling for other variables.

Notably, despite significant univariate associations between ethnicity and social deprivation and COVID risk, the full final multivariate model (table S5) including work, public or shared transport use, shopping and other variables showed no significant difference in risk of acquiring COVID-19 outside the household by ethnicity (p= 0.12) or social deprivation quintile (p=0.71).

Supplementary tables S3 and S4 show the unadjusted and adjusted odds ratios and PAFs for infection acquired outside the household in the working-age population (18-64) and in those aged > 64 respectively. Of note, among those of working age, leaving home for work or education and using public or shared transport each accounted for around 10% of non-household acquired infections (aPAFs 8.99% and 10.94%, respectively) while using shops accounted for 32% (aPAF 31.87%). For those aged 65 years and above, using public or shared transport (aPAF 14.06%) and shopping (aPAF 38.41%) accounted for the greatest proportion of non-household acquired infections. Leaving home for work was rare in this age group and not significantly associated with risk of infection.

Supplementary table S5 shows the effect of leaving home for work, public or shared transport use and shopping after additionally controlling for the number of close contacts outside the home and the previously described adjustment variables. The remaining risk from attending work is further ameliorated (aOR 1.10 95% CI 0.92 – 1.32 p= 0.30). Adjusted odds ratios for public transport and shopping were minimally affected by adding the number of non-household close contacts as a covariate.

Discussion

The study demonstrates that leaving home for work or education, public transport and shopping were important independent risk factors for acquiring COVID-19 outside the household during the second wave of the pandemic in England and Wales. Although those going to work had a substantially higher risk of infection much of this was explained by public or shared transport use and other variables. This suggests that a high proportion of the risk associated with going to work or education was due to exposure on public or shared transport. Public or shared transport use remained a strong independent risk factor after

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

controlling for other variables. Shopping was also an important risk factor for acquiring COVID-19 outside the household and, because it was a very common exposure, accounted for a high proportion of infections acquired outside the household in adults. Other non-household activities such as visiting hospitality or leisure venues were not significantly associated with acquiring COVID-19 outside the household but were rare during this period of intense restrictions. The risk of infection acquired outside the household was strongly associated with the number of close contacts outside the household. Controlling for this further ameliorated the effect of attending work suggesting this is mediated by close contact at work. Controlling for close contact made little difference to associations with public or shared transport and shopping suggesting these exposures are not mediated by recognized close contact and may represent more distant aerosol-based transmission (7). It was interesting that the effect of ethnicity and social deprivation was not seen after accounting for work, transport use, shopping and other variables – suggesting different patterns of exposure to work, associated public or shared transport use and use of shops may account for differential infection rates.

By restricting our analysis to those in the cohort with antibody test results we could ensure that infection could be ascertained in all participants (although antibody waning may lead to loss of detectable nucleocapsid antibody). Due to limited testing capacity nationwide during the first wave of the pandemic, it is possible that we have included some individuals with positive antibodies as a result of infection prior to October 2020. By excluding those with positive PCR or lateral flow or antibody tests prior to October 2020, we sought to minimise this potential over-ascertainment bias. While we sought to identify cases who were the first or only case of COVID-19 in the household it is possible that there was some misclassification error because of uncertainties in the timing of infection and failure to identify all infections in a household. Activities and behaviours are self-reported and therefore subject to recall bias, we tried to minimise this by asking about activities in the previous seven days. These activities were sampled at two points during the second wave and may not be representative of the activities throughout the second wave.

Both self-reported and linked data on test results from the national testing system also allowed ascertainment of infections. Maximising ascertainment of COVID-19 infections supports accurate assessment of the relative importance of risk factors. A further strength was the household structure of the cohort allowing us to focus analyses onto risk factors for non-household transmission and largely eliminate confounding effects that act on household transmission, although misclassification will have occurred. Outcomes and exposures were both measured during the same wave of the pandemic, although for those with antibody results only it was not possible to ascertain whether they were infected during the first or the second wave of the pandemic. Population attributable fractions are influenced by the frequency of exposures within the cohort which may not be representative of the entire adult population. For example, if our cohort includes fewer people going to work than on average, then this will lead to underestimation of the PAF related to going to work.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The research suggests that working from home and consequently avoiding the need to use public or shared transport has a significant impact on risk at individual and population level. We could not ascertain the risk associated with transmission in hospitality and leisure venues as such exposures were rare, suggesting that regulations restricting their use was effective in reducing transmission. Shopping which remained a common exposure was an important contributor to risk at individual and population levels.

During periods of intense COVID-19 transmission, increasing the proportion of people who work from home, facilitating active transport such as cycling or walking in those who need to go to work, and enabling people to shop for essential goods online would be expected to make a highly significant impact on transmission and risk of severe disease. Although high vaccination rates reduce the need for intense non pharmaceutical interventions these measures remain important in poorly vaccinated countries and may become important in the event of COVID-19 resurgence due to waning immunity, increased population mixing or emergence of new variants such as the Omicron variant.

Ethics

The Virus Watch study has been approved by the Hampstead NHS Health Research Authority Ethics Committee. Ethics approval number - 20/HRA/2320.

Funding

The research costs for the study have been 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 supported by the Wellcome Trust through a Wellcome Clinical Research Career Development Fellowship to RA [206602].

Conflicts of interest

ACH serves on the UK New and Emerging Respiratory Virus Threats Advisory Group. AMJ and ACH are members of the COVID-19 transmission sub-group of the Scientific Advisory Group for Emergencies (SAGE). AMJ 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 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.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

**Virus Watch Collaborative:

Susan Michie, Pia Hardelid, Linda Wijlaars, Eleni Nastouli, Moira Spyer, Ben Killingley, Ingemar Cox, Vasileios Lampos, Rachel A McKendry, Tao Cheng, Yunzhe Liu, Jo Gibbs, Richard Gilson, Anne M Johnson, Alison Rodger

Centre for Behaviour Change, University College London, London, UK (SM); Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, UK (PH, LW, EN, MS, BK); Francis Crick Institute, London, UK (EN, MS); Health Protection and Influenza Research Group, Division of Epidemiology and Public Health, University of Nottingham School of Medicine, Nottingham, UK (BK); University College London Hospital, London, UK (BK); Department of Computer Science, University College London, London, UK (IC, VL); London Centre for Nanotechnology and Division of Medicine, London, University College London (RM); SpaceTimeLab, Department of Civil, Environmental and Geomatic Engineering, University College London, London, UK (TC, YL); Institute for Global Health, University College London, London, UK (JG, RG). Institute for Global Health, University College London, London, UK (AMJ, AR).

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

1)Mendez-Brito A, El Bcheraoui C, Pozo-Martina Systematic review of empirical studies comparing the effectiveness of non-pharmaceutical interventions against COVID-19. J Infect. 2021 Sep; 83(3): 281–293

2)PHE Transmission group. Factors contributing to risk of SARS-CoV2 transmission associated with various settings https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/945978/S0921_Factors_contributing_to_risk_of_SARS_18122020.pdf, Last accessed 25/10/2021

3)SAGE Environmental Modelling Group: COVID-19 risk by occupation and workplace, 11 February 2021 https://www.gov.uk/government/publications/emg-covid-19-risk-by-occupation-and-workplace-11-february-2021

4)Gartland N, Fishwick, Coleman A, Davies K, Hartwig A, Johnson S, Van tongeren M. Transmission and control of SARS-CoV-2 on ground public transport: A rapid review of the literature to date. (https://figshare.manchester.ac.uk/articles/preprint/Transmission_and_control_of_SARS-CoV-2_on_ground_public_transport_A_rapid_review_of_the_literature_to_date_/16622974

5)Sage Environmental Modelling group: Insights on transmission of COVID-19 with a focus on the hospitality, retail and leisure sector. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/982865/S1194_Transmission_in_hospitality_retail_leisure.pdf

6) Hayward A, Fragaszy E, Kovar J, Nguyen V, Beale S, Byrne T, Aryee A, Hardelid P, Wijlaars L, Fong WLE, Geismar C, Patel P, Shrotri M, Navaratnam AMD, Nastouli E, Spyer M, Killingley B, Cox I, Lampos V, McKendry RA, Liu Y, Cheng T, Johnson AM, Michie S, Gibbs J, Gilson R, Rodger A, Aldridge RW. 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 Jun 23;11(6):e048042. doi: 10.1136/bmjopen-2020-048042. PMID: 34162651; PMCID: PMC8230990.

7)Leung NHL. Transmissibility and transmission of respiratory viruses. Nature Reviews Microbiology.2021; 19: 528–545

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1: Proportion of overall infection attributable to household and non-household transmission

Identification of primary, co-primary and secondary cases	**Full monthly survey AB cohort (n=10,858)**
Number of positive PCR/lat flow	343
b) Secondary cases (no symptom data but someone in the house tests (+) first, by National Testing. If national testing data date was not available then the date of onset was imputed as the middle of the week when the participant reported they tested positive).	16(5%)
a)Secondary case (someone else in home who is PCR/Lateral Flow positive has symptoms prior to case)	47 (14%)
Upper Limit Household transmission (proportion of PCR confirmed cases that are thought to be secondary cases) (a+b)	63/343=18%
c)Primary case: Only case in a home	189 (55%)
d) Co-primary case: Symptom date the same as earliest symptom in the home	85 (25%)
e) Primary case: no symptom data but tests first in the home by National Testing data. If National testing data not available, by VW imputed mid-week date.	6 (2%)
Non-household transmission - proportion of PCR/lateral flow confirmed cases that are thought to be due to non-household transmission (c+d+e/a+b+c+d+e)	280/343=82%
AB (+) PCR (-) cases	914
e) Primary AB case: only AB (+) PCR (-) case in the home (i.e. no other AB (+) or PCR (+) case in the home)	594 (65%)
f) Potential secondary case: AB (+) PCR (-) case with someone else infected in the home either as - another AB (+) PCR (-) case in the home or - a (+) PCR case in the home	 279 41 Total 320 (35%)
Non-household transmission cohort = 10858-(a+b+f)	10,475

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2 Participant characteristics and risk of infection acquired outside the household during the pandemic second wave

Characteristic	**Category**	**N=10,475 (% in category)**	**Number of infections acquired outside the household n=874 (% within category)**	**Unadjusted OR**	**95% CI**	**P**
Vaccine status	Yes No	9,502 (91%) 973 (9%)	736 (8%) 138 (14%)	1.00 1.97	 1.62 – 2.39	<0.0001
Age	Under 18 Working Age 65 and above Missing	383 (4%) 5,412 (52%) 4,669 (45%) 11	23 (6%) 634 (12%) 217 (5%)	1.00 2.08 0.76	 1.35 – 3.19 0.49 – 1.19	<0.0001
Sex	Male Female Missing	4,538 (43%) 5,913 (57%) 24	308 (7%) 563 (10%)	1.00 1.44	 1.25 – 1.67	<0.0001
Ethnic group	White White Other Asian Black Mixed Other Missing	9,475 (91%) 589 (6%) 240 (2%) 55 (1%) 59 (1%) 49 (<1%) 8	774 (8%) 54 (9%) 31 (13%) 5 (9%) 6 (10%) 4 (8%)	1.00 1.13 1.67 1.12 1.27 0.99	 0.85 – 1.52 1.14 – 2.45 0.45 – 2.83 0.55 – 2.97 0.36 – 2.79	0.2302
Deprivation score (IMD quintile) 1= most deprived	1 2 3 4 5 Missing	778 (7%) 1,538 (15%) 2,099 (20%) 2,815 (27%) 3,230 (31%) 15	71 (9%) 165 (11%) 177 (8%) 225 (8%) 236 (7%)	1.27 1.52 1.17 1.10 1.00	0.97 – 1.68 1.24 – 1.88 0.95 – 1.43 0.91 – 1.33	0.0026

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Region	East Midlands	1,067 (10%)	83 (8%)	1.00		<0.0001
	E.of England	2,236 (21%)	160 (7%)	0.91	0.69 – 1.20	
	London	1,125 (11%)	159 (14%)	1.95	1.48 – 2.58	
	North East	554 (5%)	51 (9%)	1.20	0.83 – 1.73	
	North West	1,233 (12%)	96 (8%)	1.00	0.74 – 1.36	
	South East	1,840 (18%)	146 (8%)	1.02	0.77 – 1.35	
	South West	948 (9%)	54 (6%)	0.72	0.50 – 1.02	
	Wales	255 (2%)	24 (9%)	1.23	0.76 – 1.98	
	W. Midlands	588 (6%)	49 (8%)	1.08	0.75 – 1.56	
	Yorkshire & The Humber	614 (6%)	52 (8%)	1.09	0.76 – 1.58	
	Missing	15				

Table 3 Unadjusted and adjusted odds ratios and population attributable fractions for non-household COVID-19 acquisition

Activity and frequency of occurrence			**Positive PCR lateral flow or AB**	**Univariate and multivariate analyses**		**Population attributable fraction**
Activity	**Weekly frequency**	**Total in AB cohort (n=10,475)**	**Number of COVID cases (n=874)**	**OR (95% CI) ,p**	**Adjusted OR (95% CI), p**	**(%) Raw (adjusted)**
Leaving home for work or education	No Yes	7,266 (69%) 3,209 (31%)	507 (7%) 367 (11%)	1 1.72 (1.49 – 1.98) p<0.001	1.00 1.20 (1.02 – 1.42) P=0.0307	17.58 (Adj-6.99)
Weekly frequency of using public transport	0 >0 -1 >1	7,621 (73%) 1,733 (17%) 1,121 (11%)	539 (7%) 165 (10%) 170 (15%)	1.00 1.38 (1.15 – 1.66) 2.35 (1.95 – 2.83) P<0.001	1.00 1.24 (1.03 – 1.49) 1.82 (1.49 – 2.23) p<0.0001	5.19 (Adj-3.65) 11.17(Adj-8.76) Total 16.37 (Adj 12.42)
Weekly frequency of any retail	0 >0 -1 >1	1,560 (15%) 3,030 (29%) 5,885 (56%)	78 (5%) 234 (8%) 562 (10%)	1.00 1.59 (1.22 – 2.07) 2.01 (1.57 – 2.56) P<0.001	1.00 1.45 (1.09 – 1.92) 1.69 (1.29 – 2.21) P=0.0003	9.93 (Adj-8.31) 32.31 (Adj-26.25) Total 42.25 (Adj 34.56)

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Weekly frequency of other non household activities	0 >0 – <2 2 – 3 >3 – 5 > 5	823 (8%) 2,454 (23%) 2,396 (23%) 2,620 (25%) 2,182 (21%)	51 (6%) 184 (8%) 191 (8%) 248 (9%) 200 (9%)	1.00 1.23 (0.89 – 1.69) 1.31 (0.95 – 1.81) 1.58 (1.16 – 2.16) 1.53 (1.11 – 2.10) P<0.0066	1.00 0.99 (0.71 – 1.39) 0.88 (0.63 – 1.23) 0.97 (0.69 – 1.36) 0.87 (0.61 – 1.23) P=0.6188	 3.94 5.17 10.41 7.93 Total 27.45 (adjusted not estimated)

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary materials

Table S1 Risk of infection according to type and frequency of public or shared transport use

			Positive PCR/Lat Flow or AB	**Univariate analysis**		
Activity undertaken	**Weekly frequency**	**All participants (N=10,475)**	**Covid infection (N=874)**	**Odds Ratio**	**95% CI**	**p**
Used a taxi or car shared with someone outside the household	No Yes	8,395 (80%) 2,080 (20%)	638 (8%) 236 (11%)	1.00 1.56	 1.33 – 1.82	<0.0001
Used a bus	0 >0 – 2 >2	9,708 (93%) 604 (6%) 163 (2%)	766 (8%) 86 (14%) 22 (14%)	1.00 1.94 1.82	 1.52 – 2.46 1.16 – 2.87	0.0001
Used an over-ground train or tram	0 >0 – 2 >2	10,099 (96%) 307 (3%) 69 (1%)	814 (8%) 48 (16%) 12 (17%)	1.00 2.11 2.40	 1.54 – 2.90 1.28 – 4.49	<0.0001
Used an underground train	0 >0 – 2 >2	10,172 (97%) 251 (2%) 52 (1%)	812 (8%) 54 (22%) 8 (15%)	1.00 3.16 2.09	 2.32 – 4.31 0.98 – 4.47	<0.0001
Used an airplane	No Yes	10,401 (99%) 74 (1%)	864 (8%) 10 (14%)	1.00 1.72	 0.88 – 3.37	0.1353

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table S2 Risk of infection according to frequency of non-work or education and non-public or shared transport activities outside the household

			Positive PCR/Lat Flow or AB	**Univariate analysis**		
Activities	**Weekly frequency**	**All participants N=10,475**	**Covid infection N=874**	**Odds Ratio**	**95% CI**	**p**
Played a team sport outdoors	No Yes	10,319 (99%) 156 (1%)	862 (8%) 12 (8%)	1.00 0.91	 0.51– 1.65	0.7642
Went to a theatre, cinema, concert or sports event	No Yes	10,430 (99%) 45 (<1%)	868 (8%) 6 (13%)	1.00 1.69	 0.72 – 4.01	0.2602
Went to a shop for essential items	0 >0 – 2 >2 – 4 >4 – 5 >5	1,748 (17%) 6,273 (60%) 2,063 (20%) 219 (2%) 172 (2%)	90 (5%) 541 (9%) 202 (10%) 24 (11%) 17 (10%)	1.00 1.74 1.99 2.27 2.02	 1.38 – 2.19 1.55 – 2.59 1.41 – 3.64 1.17 – 3.48	<0.0001
Went to a shop for non-essential items	No Yes	8,187 (78%) 2,288 (22%)	677 (8%) 197 (9%)	1.00 1.05	 0.89 – 1.23	0.6034
Went to a bar, pub, club, disco	No Yes	10,361 (99%) 114 (1%)	864 (8%) 10 (9%)	1.00 1.06	 0.55 – 2.03	0.8689
Ate at a restaurant, café or canteen	No Yes	10,100 (96%) 375 (4%)	834 (8%) 40 (11%)	1.00 1.33	 0.95 – 1.86	0.1104
Went to a party	No Yes	10,446 (99%) 29 (<1%)	872 (8%) 2 (7%)	1.00 0.81	 0.19 – 3.43	0.7717
Went to a place of worship	No Yes	10,124 (97%) 351 (3%)	849 (8%) 25 (7%)	1.00 0.84	 0.55 – 1.27	0.3894
Went to an outdoor market	No Yes	9,806 (94%) 669 (6%)	816 (8%) 58 (9%)	1.00 1.04	 0.79 – 1.38	0.7540
Went to a gym/indoor sport	No Yes	10,284 (98%) 191 (2%)	858 (8%) 16 (8%)	1.00 1.00	 0.59 – 1.68	0.9866

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Went to a hairdresser, barber, nail salon, beauty parlour	No Yes	10,212 (97%) 263 (3%)	853 (8%) 21 (8%)	1.00 0.95	 0.61 – 1.49	0.8301

Table S3 Unadjusted and adjusted odds ratios and population attributable fractions for non-household COVID-19 acquisition among those of working age (18 - 64 years)

Activity and frequency of occurrence			**Positive PCR lateral flow or AB**	**Univariate and multivariate analyses**		**Population Attributable Fraction (%)**
Activity	Weekly frequency	Total in AB cohort (n=5,412)	Number of COVID cases (n=634)	OR (95% CI), p	Adjusted OR (95% CI), p	Raw (adjusted)
Leaving home for work or education	No Yes	3,008(56%) 2,404 (44%)	318 (11%) 316 (13%)	1 1.28 (1.08 – 1.51) P=0.0035	1.00 1.22 (1.02 – 1.47) P=0.0329	10.90 (Adj 8.99)
Weekly frequency of using public transport	0 >0 -1 >1	3,850 (71%) 900 (17%) 662 (12%)	396 (10%) 117 (13%) 121 (18%)	1.00 1.30 (1.04 – 1.62) 1.95 (1.56 – 2.44) P<0.001	1.00 1.19 (0.96 – 1.50) 1.72 (1.35 – 2.19) P=0.0001	4.26 (Adj 2.95) 9.29 (Adj 7.99) Total 13.56 (Adj 10.94)

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Weekly frequency of any retail	0 >0 - 1 >1	636 (12%) 1,540 (28%) 3,236 (60%)	50 (8%) 169 (11%) 415 (13%)	1.00 1.44 (1.04 – 2.01) 1.72 (1.27 – 2.34) P=0.0006	1.00 1.38 (0.99– 1.94) 1.60 (1.16 – 2.22) P=0.0097	8.14 (Adj 7.34) 27.4 (Adj 24.55) Total 35.55 (Adj 31.87)
Weekly frequency of other non household activities	0 >0 – <2 2 – 3 >3 – 5 > 5	349 (6%) 1,073 (20%) 1,210 (22%) 1,457 (27%) 1,323 (24%)	37 (11%) 125 (12%) 136 (11%) 171 (12%) 165 (12%)	1.00 1.11 (0.75 – 1.64) 1.07 (0.73 – 1.57) 1.12 (0.77 – 1.63) 1.20 (0.82 – 1.75) P=0.8404	1.00 0.91 (0.61 – 1.35) 0.78 (0.52 – 1.16) 0.77 (0.51 – 1.15) 0.79 (0.53 – 1.21) P=0.5716	1.95 1.40 2.89 4.34 Total 10.58 Adjusted not estimated

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table S4 Unadjusted and adjusted odds ratios and population attributable fractions for non-household COVID-19 acquisition among those of retired age (65 years and above)

Activity and frequency of occurrence			**Positive PCR lateral flow or AB**	**Univariate and multivariate analyses**		**Population Attributable Fraction (%)**
Activity	Weekly frequency	Total in AB cohort (n=4,669)	Number of COVID cases (n=217)	OR (95% CI) ,p	Adjusted OR (95% CI), p	Raw (adjusted)
Leaving home for work or education	No Yes	4,220(90%) 449 (10%)	189 (4%) 28 (6%)	1 1.42 (0.94 – 2.14) P=0.1073	1.00 1.28 (0.82 – 1.99) P=0.2806	3.82 (Adj 2.82)
Weekly frequency of using public transport	0 >0 -1 >1	3,515(75%) 795(17%) 359 (8%)	132 (4%) 46 (6%) 39 (11%)	1.00 1.57 (1.11 – 2.22) 3.12 (2.15 – 4.54) P<0.001	1.00 1.29 (0.89 – 1.87) 2.07 (1.36 – 3.14) P=0.0042	7.69 (Adj 4.77) 12.21 (Adj 9.29) Total 19.91 (Adj 14.06)
Weekly frequency of any retail	0 >0 -1 >1	669 (14%) 1,400 (30%) 2,600 (56%)	17 (3%) 57 (4%) 143 (6%)	1.00 1.62 (0.94 – 2.82) 2.23 (1.34 – 3.72) P=0.0014	1.00 1.59 (0.89 – 2.83) 1.77 (1.01 – 3.09) P=0.1083	10.05 (Adj 9.75) 36.35 (Adj 28.67) Total 46.40 (Adj 38.41)

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Weekly frequency of other non household activities	0 >0 – <2 2 – 3 >3 – 5 > 5	434 (9%) 1,340 (29%) 1,050 (22%) 1,048 (22%) 797 (17%)	14 (3%) 57 (4%) 49 (5%) 66 (6%) 31 (4%)	1.00 1.33 (0.74 – 2.42) 1.47 (0.80 – 2.69) 2.02 (1.12 – 3.63) 1.21 (0.64 – 2.31) P=0.0477	1.00 1.12 (0.59 – 2.12) 1.07 (0.55 – 2.07) 1.47 (0.77 – 2.80) 0.88 (0.43 – 1.78) P=0.2195	6.52 (Adj 2.81) 7.22 (Adj 1.48) 15.36 (Adj 9.72) 2.48 (adj not estimated) Total 31.57 (Adj not estimated)

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table S5 – Multivariate analysis of number of contacts outside household, non household activities and risk of infection (all variables are mutually adjusted)

		PCR or lateral flow or AB positive		
Activities	**Weekly frequency**	**Adjusted OR**	**(CI)**	**p**
Number of contacts outside the home	0 >0 - 2 >2 - 5 < 5 - 10 >10	1.00 0.93 1.15 1.34 1.35	 0.78 – 1.12 0.91 – 1.45 0.96 – 1.87 0.99 – 1.83	0.0431
Leaving home for work or education	No Yes	1.00 1.10	 0.92 – 1.32	0.2967
Weekly frequency of using public transport	0 >0 -1 >1	1.00 1.23 1.76	 1.01 – 1.49 1.43 – 2.16	<0.0001
Weekly frequency of any retail	0 >0 -1 >1	1.00 1.46 1.69	 1.10 – 1.93 1.29 – 2.22	0.0003
Weekly frequency of other activities	0 >0 - >1 2 – 3 >3 – 5 > 5	1.00 1.00 0.87 0.95 0.83	 0.72 – 1.39 0.62 – 1.22 0.68 – 1.34 0.59 – 1.19	0.4742
Vaccine status	Yes No	1.00 1.99	 1.59 – 2.49	<0.0001
Age	Under 18 Working Age 65 and above	1.00 3.60 1.61	 2.17 – 5.97 0.94 – 2.76	 <0.001 0.082
Ethnic group	White White Other Asian Black Mixed Other	1.00 0.67 0.97 0.66 0.69 0.55	 0.49 – 0.92 0.64 – 1.46 0.26 – 1.69 0.29 – 1.66 0.19 – 1.59	0.1196

medRxiv preprint doi: https://doi.org/10.1101/2021.12.08.21267458; this version posted December 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Deprivation score (IMD quintile) 1= most deprived	1 2 3 4 5	0.92 1.13 1.02 1.02 1.00	0.69 – 1.23 0.91 – 1.42 0.83 – 1.26 0.84 – 1.24	0.7067
Region	East Midlands East of England London North East North West South East South West Wales West Midlands Yorkshire & The Humber	1.00 0.96 1.62 1.17 0.99 1.04 0.77 1.27 1.12 1.19	 0.73 – 1.28 1.19 – 2.19 0.80 – 1.70 0.72 – 1.35 0.78 – 1.39 0.54 – 1.10 0.78 – 2.07 0.76 – 1.62 0.82 – 1.72	0.0027
Sex	Male Female	1.00 1.36	 1.17 – 1.58	<0.001