SARS-CoV-2 infections in migrants and the role of household overcrowding: a causal mediation analysis of Virus Watch data

Yamina Boukari ⓘ ,[1] Sarah Beale ⓘ ,[1,2] Vincent Nguyen,[1,2,3] Wing Lam Erica Fong ⓘ ,[1] Rachel Burns,[1] Alexei Yavlinsky,[1] Susan Hoskins,[2] Kate Lewis ⓘ ,[4] Cyril Geismar,[2] Annalan MD Navaratnam,[1,2] Isobel Braithwaite,[1] Thomas E Byrne,[1] Youssof Oskrochi ⓘ ,[1] Sam Tweed,[1] Jana Kovar,[1] Parth Patel ⓘ ,[1] Andrew Hayward,[2] Robert Aldridge ⓘ [1]

ABSTRACT

Background Migrants are over-represented in SARS-CoV-2 infections globally; however, evidence is limited for migrants in England and Wales. Household overcrowding is a risk factor for SARS-CoV-2 infection, with migrants more likely to live in overcrowded households than UK-born individuals. We aimed to estimate the total effect of migration status on SARS-CoV-2 infection and to what extent household overcrowding mediated this effect.

Methods We included a subcohort of individuals from the Virus Watch prospective cohort study during the second SARS-CoV-2 wave (1 September 2020–30 April 2021) who were aged ≥18 years, self-reported the number of rooms in their household and had no evidence of SARS-CoV-2 infection pre-September 2020. We estimated total, indirect and direct effects using Buis' logistic decomposition regression controlling for age, sex, ethnicity, clinical vulnerability, occupation, income and whether they lived with children.

Results In total, 23 478 individuals were included. 9.07% (187/2062) of migrants had evidence of infection during the study period vs 6.27% (1342/21 416) of UK-born individuals. Migrants had 22% higher odds of infection during the second wave (total effect; OR 1.22, 95% CI 1.01 to 1.47). Household overcrowding accounted for approximately 36% (95% CI −4% to 77%) of these increased odds (indirect effect, OR 1.07, 95% CI 1.03 to 1.12; proportion accounted for: indirect effect on log odds scale/total effect on log odds scale=0.36).

Conclusion Migrants had higher odds of SARS-CoV-2 infection during the second wave compared with UK-born individuals and household overcrowding explained 36% of these increased odds. Policy interventions to reduce household overcrowding for migrants are needed as part of efforts to tackle health inequalities during the pandemic and beyond.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Migrants in England and Wales may be at greater risk of exposure to SARS-CoV-2 due to unique risk factors, including over-representation in front-line jobs, an increased likelihood of living in multigenerational households and difficulties in accessing primary care. Research shows that migrants in high-income countries have been disproportionally infected with SARS-CoV-2. It is likely that, due to their pre-existing vulnerabilities, this is similarly the case for migrants in England and Wales; however, quantitative evidence addressing this is lacking.

WHAT THIS STUDY ADDS

⇒ We investigated the effect of being a migrant on SARS-CoV-2 infection during the second wave of the pandemic in a cohort in England and Wales. We also determined the proportion of the effect mediated by household overcrowding after controlling for age, sex, ethnicity, clinical vulnerability, occupation, income and the presence of children in the household. Migrants had 22% higher odds of being infected with SARS-CoV-2 than their UK-born counterparts, and household overcrowding accounted for approximately 36% of these increased odds.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our findings highlight the role of household overcrowding in the disproportionate impact of SARS-CoV-2 infections on migrants. They also demonstrate the urgent need for policy interventions that improve housing conditions as part of efforts to reduce health inequalities. Further research investigating other causes of migrants' over-representation in infection cases is also needed to inform further targeted policy interventions.

BACKGROUND

Globally, the UK has the fifth largest migrant (non UK-born) population comprising approximately 9.57 million people in 2020.[1] [2] Migrants in the UK may be at greater risk of SARS-CoV-2 infection due to pre-existing vulnerabilities such as their over-representation in front-line jobs (eg, in healthcare, hospitality, retail and delivery sectors), increased use of public transport and increased likelihood of living in multigenerational households.[3] Barriers to

accessing primary care are well documented for migrants[4–6] and may negatively impact vaccine uptake, thus potentially putting migrants at greater risk of infection and severe outcomes from the combination of greater exposure and undervaccination.[7] Migrants in high-income countries have been over-represented in SARS-CoV-2 infections, hospitalisations and deaths.[8–10] UK-focused quantitative evidence is limited but does suggest inequalities. In England, a study showed a greater increase in all-cause mortality for migrants versus non-migrants from 21 March 2020 to 8 May 2020 when compared with previous years' deaths.[11]

The built environment is a wider determinant of health.[12] Household overcrowding is a potential marker of social deprivation and is associated with an increased risk of infectious diseases, including tuberculosis, influenza-related illnesses, pneumonia and acute respiratory illness, and mental health problems.[13] Growing UK-focused evidence links household overcrowding to SARS-CoV-2 infection and other COVID-19-related outcomes.[14 15] In England and Wales, individuals who participated in the Virus Watch study and lived in overcrowded households, as defined using the persons-per-room (PPR) methodology, had higher odds of testing positive for SARS-CoV-2 infection via PCR and antibody tests than individuals living in underoccupied households.[14] Similar findings were reported from the COVIDENCE UK study[15] and studies that used related measures such as household size when controlling for various demographic, social, behavioural and comorbidity characteristics[16–19] or area-level housing indicators.[20] Household size also played a role in differences in COVID-19 outcomes for South Asian groups after adjusting for sociodemographic and clinical factors.[21 22]

Household overcrowding is particularly relevant to migrants. In London, 13%–16% of migrant households were overcrowded compared with just 4% of UK-born households between 2016 and 2018.[23] Outside of London, the overcrowding rates were lower with 2% of UK-born households being overcrowded compared with 5%–8% of migrant households. Despite the lack of UK-focused studies, in Europe and the USA, household overcrowding is a reported risk factor for SARS-CoV-2 exposure in migrants,[24–26] thus highlighting the need for investigation in a UK-based sample. We aimed to examine the odds of SARS-CoV-2 infection for migrants versus UK-born individuals during the second COVID-19 wave, and whether household overcrowding, as determined using the PPR methodology, mediated the effect of migration status on SARS-CoV-2 infection.

METHOD
Study setting
We used data from Virus Watch, a prospective community cohort study of COVID-19 in England and Wales from 1 June 2020.[27] Virus Watch included 58 628 individuals as of 28 July 2022. Our analysis was restricted to the second wave, from 1 September 2020 to 30 April 2021, as migrants were likely to have high exposure risk early in the pandemic and because testing was not widespread during the first wave. Households were recruited from 24 June 2020 to March 2022 and asked to complete a postenrolment baseline survey containing demographic, medical history, financial and occupation questions. Individuals received a weekly illness survey via email to collect information on self-reported acute symptoms, vaccination status and PCR or lateral flow test results. Households also received a monthly survey of demographic, health-related, environmental and behavioural/psychosocial questions. Within the larger study, a subcohort of adults (the laboratory cohort) received monthly antibody testing.

Virus Watch cohort data were linked to the second-generation surveillance system (SGSS) containing laboratory SARS-CoV-2 test results from swabs taken during hospitalisation (pillar 1) and community testing (pillar 2).[28] The linkage period was March 2020–August 2021 for pillar 1, and June 2020–November 2021 for pillar 2. Linkage was conducted by NHS Digital using the name, date of birth, address and NHS number variables, which were sent in March 2021.

Participants
Participants were aged ≥18 years and reported the number of rooms in their household in the February 2021 survey. Participants with evidence of SARS-CoV-2 infection before the start of the second wave (September 2020) were excluded as first infection, rather than reinfection, was our focus.

Exposure and outcome
Country of birth was the exposure, defined as migrant (ie, a non-UK country of birth reported at enrolment) or UK-born (a UK country of birth). The outcome was evidence of first SARS-CoV-2 infection during the analysis period defined as either a PCR or lateral flow test self-reported during a weekly survey, a positive test for SARS-CoV-2 anti-nucleocapsid antibodies, a positive test for anti-spike antibodies or a positive PCR test identified via the SGSS.

Potential mediator: household overcrowding status
In the February 2021 survey, participating households were asked how many rooms were available for their exclusive use (excluding bathrooms, toilets, halls, landings and cupboards). PPR was calculated by dividing the total number of people in the household (including children) by the number of available rooms, excluding bathrooms or kitchens. For households reporting ≥2 rooms, one room was subtracted from the total, assuming that one of the rooms was a kitchen. For households reporting one room, it was assumed that this room is likely to be a bedroom/studio. Households with PPRs less than one were defined as 'underoccupied', equal to one as 'balanced' and greater than one as 'overcrowded'.[14] The PPR approach is a validated overcrowding measure that has fair agreement with other measures[29] and has previously been used to determine overcrowding status in Virus Watch.[14]

Confounders
Potential confounders were identified using a directed acyclic graph (DAG; online supplemental figure 1 and online supplemental box 1 for the DAGitty code). To provide minimally adjusted unbiased estimates of the total, indirect and direct effects, we controlled for baseline age, sex at birth, ethnicity (white British, white Irish, white Other, mixed, South Asian, other Asian, Black, other and unknown), clinical vulnerability ('not clinically vulnerable', 'clinically vulnerable', 'clinically extremely vulnerable' and 'missing' based on self-reported conditions or medications indicating vulnerability using government criteria, adapted for the clinical variables collected at baseline[30]), baseline total household income (£0–9999, £10 000–£24 999, £25 000–£49 000, £50 000–£74 999, £75 000–£99 999 and £100 000+), occupation (see online supplemental box 2 for details) and whether the household included children.

Other demographic and clinical characteristics
Households were assigned a geographical region (nine English regions and Wales) and a local area-level Index of Multiple

Deprivation quintile (where 1 represents the most deprived and 5 the least) by linking their postcode to the May 2020 ONS Postcode Lookup.[31]

Statistical analysis

We found no evidence of colinearity between ethnicity and migrant status (online supplemental box 3). We used Buis' logistic decomposition regression with bootstrapped standard errors to estimate the total and direct effects of migration status on infection, and the indirect effect mediated through household overcrowding[32] (see online supplemental box 3 for further methodological information). The percentage of the total effect mediated by household overcrowding was estimated using the indirect effect beta coefficient as the numerator and the total effect beta coefficient (in the form of log odds) as the denominator. The 95% CI was estimated via the delta method using the nlcom command in Stata. As the model derives the total effect coefficient by summing the direct and indirect effects on the log scale, the percentage is an approximation only (ie, indirect and direct effect ORs do not sum to give exactly the total effect).

Sensitivity analyses

As eligibility was not dependent on households responding to all weekly surveys throughout the analysis period, we conducted a sensitivity analysis including only participants who had either linked data or, for those without linked data, self-reported follow-up for every full week of the analysis period (although this may also bias towards households who were healthy enough to respond each week). Positive SARS-CoV-2 anti-nucleocapsid antibody or anti-spike antibody tests during the study period may indicate evidence of older SARS-CoV-2 infection prior to the second wave or postvaccination seroconversions. Therefore, we carried out a sensitivity analysis using only swab-confirmed infections. We also conducted a sensitivity analysis where household overcrowding was represented by the continuous PPR variable.

Individuals with missing country-of-birth responses were classified as UK-born; however, in a sensitivity analysis, we excluded these individuals. There were no missing age data and 363 participants (1.5%) had missing data for sex, 478 (2.0%) for ethnicity, 1405 (6.0%) for clinical vulnerability, 4120 (17.5%) for household income and 3400 (14.5%) for occupation. All missing values were included in the model under a 'missing' category. We conducted sensitivity analyses to assess the effect of missing data using a complete-case analysis and multivariate imputation by chained equations using the mice package with 5 datasets and 50 iterations per dataset (see online supplemental box 4 for the included predictor variables).

Based on a priori assumptions, we modelled ethnicity as a confounder of the effect of migration status on infection. However, ethnicity is complex and can encompass country of birth,[33] which creates overlap with our migrant exposure. To explore this, we conducted a sensitivity analysis without adjusting for ethnicity.

Tools and reporting

R V.4.1.2 was used for data cleaning and multiple imputation. Mediation analysis was carried out using the ldecomp command in Stata V.17.0. This analysis was written up in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology checklist (online supplemental table 1).

RESULTS

Of 58 628 individuals in the Virus Watch cohort on 28 July 2022, 23 478 (40.0%) individuals met our inclusion criteria (figure 1).
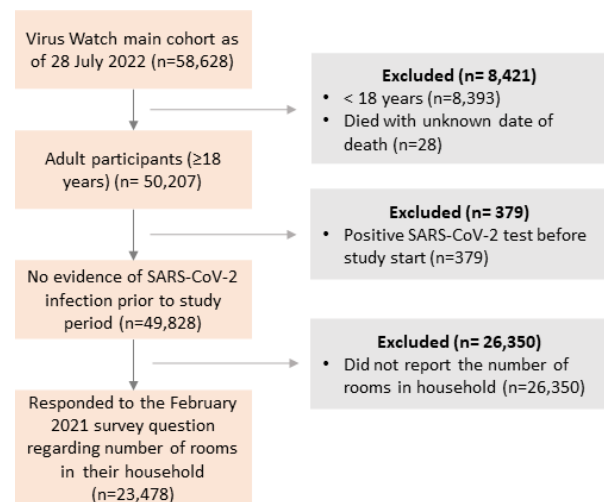

Figure 1 Flow diagram of participant eligibility.

Demographic characteristics

Migrants were generally younger (median age: 53 vs 63 years in migrants and non-migrants, respectively) with a higher percentage identifying as female (59.1% female migrants vs 54.1% female non-migrants; table 1). Migrants identified predominantly with a minority ethnic group (75.6%) vs white British (23.5%) and were less likely to have missing ethnicity data than UK-born individuals (table 1). Over 40% of migrants were situated in London compared with only 9.5% of UK-born individuals, while over 40% of the UK-born individuals were in

Table 1 Cohort demographic and household characteristics

Characteristic	Overall N=23 478	UK-born N=21 416	Migrant N=2062
Age			
Mean (SD)	59 (15)	59 (15)	53 (16)
Median (IQR)	62 (19)	63 (19)	53 (26)
<30	1379 (5.9%)	1241 (5.8%)	138 (6.7%)
30–39	1718 (7.3%)	1341 (6.3%)	377 (18.3%)
40–49	2537 (10.8%)	2163 (10.1%)	374 (18.1%)
50–59	4397 (18.7%)	4022 (18.8%)	375 (18.2%)
60+	13 447 (57.3%)	12 649 (59.1%)	798 (38.7%)
Sex			
Female	12 810 (54.6%)	11 592 (54.1%)	1218 (59.1%)
Male	10 276 (43.8%)	Omitted to avoid disclosure	Omitted to avoid disclosure
Other or unknown	392 (1.7%)	Omitted to avoid disclosure	Omitted to avoid disclosure
Ethnicity			
White British	20 456 (87.1%)	19 972 (93.3%)	484 (23.5%)
Minority ethnic*	2544 (10.8%)	986 (4.6%)	1558 (75.6%)
Missing	478 (2.0%)	458 (2.1%)	20 (1.0%)
Household overcrowding status			
Underoccupied	21 134 (90.0%)	19 662 (91.8%)	1472 (71.4%)
Balanced	1686 (7.2%)	1320 (6.2%)	366 (17.7%)
Overcrowded	658 (2.8%)	434 (2.0%)	224 (10.9%)
Persons per room			
Mean (SD)	0.49 (0.33)	0.47 (0.29)	0.71 (0.53)
Median (IQR)	0.40 (0.31)	0.40 (0.29)	0.50 (0.67)

*All minority ethnic groups were combined to avoid statistical disclosure. For a full description of demographic characteristics, please see online supplemental table 2.

Table 2 Source of positive test results for all participants with first evidence of SARS-CoV-2 infection within the analysis period (1 September 2020–30 April 2021)

Source of positive test result	Overall N=1529	UK-born n=1342 (87.8%)	Migrant n=187 (12.2%)
Antibody tests*	713 (46.6%)	629 (46.9%)	84 (44.9%)
Swab tests (PCR or lateral flow)	816 (53.4%)	713 (53.1%)	103 (55.1%)

*Tests for SARS-CoV-2 anti-nucleocapsid antibodies and anti-spike antibodies.

the East or South East of England (online supplemental table 2). Migrants generally lived in more deprived areas compared with the UK-born group, but more migrants lived in households with higher total incomes (online supplemental table 2). Missing income data were more common for UK-born individuals. The percentages of clinically and extremely clinically vulnerable individuals in each group were similar, with more missing data in the UK-born group. Higher percentages of migrants worked in all occupations versus UK-born individuals, apart from in outdoor trade-related and transport and machines-related occupations. More UK-born individuals were not in employment versus migrants and missing occupation data were more common in the UK-born versus migrant group.

In the migrant group, the median number of rooms per household was 5 compared with 6 in the UK-born group (online supplemental table 2). 10.9% of migrants lived in overcrowded housing compared with 2.0% of the UK-born group. Migrants were less likely to live in underoccupied housing than UK-born individuals (71.4% vs 91.8%, respectively; table 1). Across the full cohort, the majority of individuals lived in households where the PPR was less than 1, with low numbers of participants living

in households where the PPR was 2 or more (online supplemental figure 2).

Evidence of SARS-CoV-2 infection
From 1 September 2020 to 30 April 2021, 1529/23 478 individuals had evidence of SARS-CoV-2 infection (table 2). In both groups, evidence of infection was identified via a swab test in >50% of positive cases, with a slightly higher percentage of swab tests in migrants versus UK-born individuals (55.1% vs 53.1%, respectively).

In the migrant group, 9.07% (187/2062) of individuals had evidence of infection compared with 6.27% (1342/21 416) of UK-born individuals (table 3). Infection rates were highest at over 10% in London for both migrants and non-migrants compared with most other geographical areas outside of London where the infection rates were lower (generally less than 10%). In both groups, the percentage of participants with evidence of infection was highest in individuals living in overcrowded housing (migrant: 15.2% (34/224); UK-born: 9.9% (43/434)) compared with individuals living in under-occupied housing (migrant: 6.9% (102/1472); UK-born: 6.0% (1185/19 662)). Please see online supplemental table 3 for a full breakdown of infections by different demographic characteristics.

Causal mediation analysis
Migrants had 22% higher odds of SARS-CoV-2 infection during the second wave versus UK-born individuals (total effect), determined using logistic decomposition regression adjusted for age, sex, ethnicity, clinical vulnerability, baseline total household income, occupation and the presence of children in the household (OR 1.22, 95% CI 1.01 to 1.47, p=0.041; table 4). An OR of 1.07 (95% CI 1.03 to 1.12, p=0.002) for the indirect

Table 3 Percentage of participants with evidence of SARS-CoV-2 infection in the second wave (1 September 2020 to 30 April 2021)

Characteristic	UK-born			Migrant		
	Overall, N=21 416*	Evidence of SARS-CoV-2 infection		Overall, N=2062	Evidence of SARS-CoV-2 infection	
		No n=20 074 (93.7%)	Yes n=1342 (6.27%)		No n=1875 (90.9%)	Yes n=187 (9.07%)
Age						
<30	1241	1125 (90.7%)	116 (9.3%)	138	123 (89.1%)	15 (10.9%)
30–39	1341	1235 (92.1%)	106 (7.9%)	377	332 (88.1%)	45 (11.9%)
40–49	2163	1969 (91.0%)	194 (9.0%)	374	322 (86.1%)	52 (13.9%)
50–59	4022	3691 (91.8%)	331 (8.2%)	375	343 (91.5%)	32 (8.5%)
60+	12 649	12 054 (95.3%)	595 (4.7%)	798	755 (94.6%)	43 (5.4%)
Sex						
Male	9435	8877 (94.1%)	558 (5.9%)	841	770 (91.6%)	71 (8.4%)
Female	11 592	10 823 (93.4%)	769 (6.6%)	1218	1102 (90.5%)	116 (9.5%)
Ethnicity						
White British	19 972	18 723 (93.7%)	1249 (6.3%)	484	451 (93.2%)	33 (6.8%)
Minority ethnic	986	910 (92.3%)	76 (7.7%)	1558	1404 (90.1%)	154 (9.9%)
Missing	458	441 (96.3%)	17 (3.7%)	20	20 (100.0%)	0 (0.0%)
Geographical region*						
East of England	4889	4611 (94.3%)	278 (5.7%)	319	297 (93.1%)	22 (6.9%)
London	2044	1830 (89.5%)	214 (10.5%)	889	769 (86.5%)	120 (13.5%)
South East	4147	3927 (94.7%)	220 (5.3%)	377	358 (95.0%)	19 (5.0%)
Household overcrowding status						
Underoccupied	19 662	18 477 (94.0%)	1185 (6.0%)	1472	1370 (93.1%)	102 (6.9%)
Balanced	1320	1206 (91.4%)	114 (8.6%)	366	315 (86.1%)	51 (13.9%)
Overcrowded	434	391 (90.1%)	43 (9.9%)	224	190 (84.8%)	34 (15.2%)

*Only the geographical areas with the highest proportion of individuals are included to avoid statistical disclosure. For a full description of infections by each demographic characteristic, please see online supplemental table 3.

Table 4 ORs for total, indirect and direct effects of migration status on SARS-CoV-2 infection

Effect	OR	95% CI	P value
Total	1.22	1.01 to 1.47	0.041
Indirect	1.07	1.03 to 1.12	0.002
Direct	1.13	0.94 to 1.37	0.198

effect indicates that household overcrowding partially mediated the relationship between migration status and SARS-CoV-2 infection, accounting for approximately 36% (95% CI −4% to 77%) of the total effect. A positive, but not statistically significant direct effect of migration status on SARS-CoV-2 infection remained after accounting for the indirect effect of household overcrowding status (OR 1.13, 95% CI 0.94 to 1.37, p=0.198).

Sensitivity analyses

Consistent indirect effect sizes were observed in all the sensitivity analyses, with varying statistical significance (table 5). Total effects and direct effects were also generally consistent across sensitivity analyses, with the exception of the complete case analysis and after excluding individuals with missing country of birth, likely due to power issues with the reduced sample size (table 5).

DISCUSSION

We present findings indicating that migrants had 22% higher odds of being infected with SARS-CoV-2 during the second wave of the pandemic compared with UK-born individuals after controlling for baseline demographic, socioeconomic and clinical confounders. This increased odds of infection aligns with findings from other high-income countries showing migrants' over-representation in SARS-CoV-2 infections[8] and is likely due to amplified prepandemic inequalities. These findings, alongside evidence showing migrants' low COVID-19 vaccine uptake,[7 34] highlight the need to carefully consider delivery and prioritisation of booster COVID-19 vaccines.

We found a significant positive indirect effect, with household overcrowding explaining approximately 36% of the increased odds of SARS-CoV-2 infection in migrants compared with UK-born individuals, which is consistent with Norwegian and American studies linking household overcrowding and SARS-CoV-2 infections.[24 25] A direct effect was found after accounting for mediation by household overcrowding, consistent with complementary mediation whereby the investigated mediator has a significant causal role alongside other unmeasured variables.[35]

A strength of this analysis is the focus on a causal mechanism underlying migrants' increased SARS-CoV-2 infection odds in a substantial group of over 2000 migrants, allowing for specific policy recommendations to help reduce health inequalities. The use of a DAG-informed model facilitated comprehensive adjustment for confounders and interpretation of the mediated effects. Other strengths are the inclusion of self-reported and linked data on SARS-CoV-2 infection, which reduced reliance on participant recall, and the use of multiple sensitivity analyses.

A limitation is that migrants in Virus Watch are not representative of England's migrant population. Lead householders who spoke English and had internet access were eligible, whereas evidence suggests that vulnerable, marginalised migrants have limited access to technology and experience English difficulties.[5 7] Additionally, only households of ≤6 people were eligible, which may induce selection bias given that migrants are more

Table 5 Total, indirect and direct effects from sensitivity analyses

Sensitivity analysis	N	Effect	OR	95% CI	P value
Outcome related					
Restricted denominator	22 496	Total	1.17	0.97 to 1.41	0.098
		Indirect	1.07	1.02 to 1.12	0.003
		Direct	1.09	0.90 to 1.32	0.368
Positive swab tests only	23 479	Total	1.20	0.92 to 1.56	0.174
		Indirect	1.12	1.07 to 1.18	<0.001
		Direct	1.07	0.83 to 1.37	0.609
Exposure related					
Excluding individuals with unknown country of birth	20 748	Total	1.04	0.85 to 1.28	0.687
		Indirect	1.06	1.02 to 1.10	0.002
		Direct	0.98	0.81 to 1.20	0.872
Mediator related					
Use of PPR as mediator	23 478	Total	1.23	1.03 to 1.46	0.023
		Indirect	1.07	1.03 to 1.11	<0.001
		Direct	1.14	0.96 to 1.36	0.132
Confounder related					
Multiple imputation of missing confounder data	23 478	Total	1.21	0.99 to 1.48	0.067
		Indirect	1.07	1.03 to 1.11	0.001
		Direct	1.13	0.92 to 1.38	0.238
Complete-case analysis	18 466	Total	1.07	0.86 to 1.33	0.550
		Indirect	1.07	1.01 to 1.12	0.011
		Direct	1.00	0.80 to 1.26	0.982
Without adjusting for ethnicity	23 478	Total	1.24	1.05 to 1.47	0.011
		Indirect	1.08	1.03 to 1.12	<0.001
		Direct	1.16	0.97 to 1.37	0.095

PPR, persons per room.

likely to live in larger, multigenerational households.[3] We also focused on infections that occurred during the second wave and excluded individuals who were infected in the first wave, which could potentially induce bias given the risk factors for higher exposure faced by migrants versus UK-born individuals early in the pandemic.[9] Consequently, it is likely that with this in mind, we underestimated the effect of migration status on infection and the household overcrowding-mediated indirect effect. Despite PPR being an accepted measure of overcrowding, it is also important to acknowledge that counting rooms, as opposed to bedrooms, can potentially underestimate the extent of overcrowding[29]; however, this underestimation is likely to be similar for both the migrant and UK-born group.

Individuals whose country of birth was unknown (n=2730) were classified as UK-born. While this could introduce misclassification bias whereby true migrants are classified as being UK-born, the impact is likely small and would underestimate the effect of migration status on infection. Additionally, results from the sensitivity analysis excluding individuals with a missing country of birth were generally consistent with the main analysis.

We did not adjust for vaccination status at baseline as no individuals had been vaccinated at the start of the analysis period (roll-out began in England on 8 December 2021). However, by 25 April 2021, 91.5% (22 644 679) of individuals aged ≥45 years had received at least one dose.[36] Evidence suggests that vaccination status is a separate mediator of SARS-CoV-2 infection in migrants, with differential uptake of COVID-19 vaccines across migrant and UK-born groups and underimmunisation in migrants in Europe for both COVID-19 and routine vaccinations.[34] Consequently, adjustment for vaccination status should not influence the indirect effect through household overcrowding.

Other limitations are the exclusion of children (<18 years), a group which requires further research. Additionally, Virus Watch enrolled more individuals aged over 60 years or of white British ethnicity versus England's and Wales's general population, and included more higher income households. These biases likely contribute to an underestimation of the effect of migration status on infection. Finally, a limitation of the methodology that we used to estimate the indirect effect through household overcrowding is that we have potentially adjusted for intermediate confounders (ie, confounders of the mediator-and-outcome relationship that are also caused by the exposure). This has the potential to remove some of the true causal effect of the exposure (migration status) on the outcome (SARS-CoV-2 infection).[37] To assess this, we also estimated a separate total causal effect (without using the Buis methodology) and obtained a similar estimate, which increases confidence in the estimates obtained from the Buis methodology. Nevertheless, we acknowledge that the use of g-methods could be explored in future work.

The disproportionate effect of household overcrowding on individuals in the migrant group compared with UK-born individuals builds on previous Virus Watch results showing household overcrowding as a risk factor for SARS-CoV-2 infection.[14] Household overcrowding has become more common over recent years, particularly in the private and social rental sectors.[38] Migrants are more likely to privately rent their homes and have lower rates of homeownership than non-migrants,[23] which may explain the differential impact on this group. Findings from the two analyses demonstrate the health implications of existing housing inequalities. They highlight the importance of addressing overcrowding as part of a public health strategy to reduce health inequalities, and to ensuring the UK's preparedness for any subsequent waves or future pandemics. Future research is required to examine other potential mediators of the total effect

of migration status on infection to better inform targeted policy interventions across the wider determinants of health.

Efforts to address overcrowding are complex and require engagement with multiple stakeholders, including both in government and the private sector. Short-term efforts to prevent spread include providing hotel accommodation for infected individuals and ensuring that existing advice on preventing spread within the household (eg, masking, adequate ventilation)[39] is accessible to migrant communities. In the medium term, the statutory overcrowding standard that was introduced in 1935 should be revised as the threshold for breaching it is high, with relatively few households found to be statutorily overcrowded, which limits the ability of local authorities to act.[40] In recognition that local authorities may not have access to affordable, adequately sized houses, the revision of the standard should be combined with policies that improve affordability, such as through household income and employment support.[41] In the long term, the current housing stock should be reformed. A policy focus in recent years has been on boosting the supply of privately owned homes and increasing home ownership, with a substantial drop in the availability of social housing.[38] Increasing the social housing supply could provide more regulated, secure and affordable housing, while enabling swifter improvements in response to policy interventions compared with private housing.

To conclude, we show that migrants were over-represented in SARS-CoV-2 infections early in the pandemic, with household overcrowding playing a significant role in driving over-representation. Our findings highlight the implications of inadequate housing on health and underscore the importance of policy interventions for tackling household overcrowding. As we continue to live with COVID-19, it will be important to address these inequalities in health outcomes and housing to ensure than we build back fairer[42] and that we are better prepared for future waves and pandemics.

Twitter Yamina Boukari @BoukariYamina, Youssof Oskrochi @YOskrochi and Parth Patel @pathwithanr

ORCID iDs
Yamina Boukari http://orcid.org/0000-0002-9171-6922
Sarah Beale http://orcid.org/0000-0002-4038-7460
Wing Lam Erica Fong http://orcid.org/0000-0001-5097-2228
Kate Lewis http://orcid.org/0000-0003-1148-1017
Youssof Oskrochi http://orcid.org/0000-0001-6748-0457
Parth Patel http://orcid.org/0000-0001-5207-2776
Robert Aldridge http://orcid.org/0000-0003-0542-0816

REFERENCES

1 United Nations Department of Economic and Social Affairs. International migration 2020 highlights available from. Available: https://www.un.org/en/desa/international-migration-2020-highlights [Accessed 23 Aug 2022].

2 Office for National Statistics. Population of the UK by country of birth and nationality: individual country data. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/datasets/populationoftheunitedkingdombycountryofbirthandnationalityunderlyingdatasheets [Accessed 27 Dec 2021].

3 Public Health England. Beyond the data: understanding the impact of COVID-19 on BAME groups. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/892376/COVID_stakeholder_engagement_synthesis_beyond_the_data.pdf [Accessed 1 Aug 2022].

4 Weller SJ, Crosby LJ, Turnbull ER, *et al*. The negative health effects of hostile environment policies on migrants: A cross-sectional service evaluation of humanitarian Healthcare provision in the UK. *Wellcome Open Res* 2019;4:109.

5 Kang C, Tomkow L, Farrington R. Access to primary health care for asylum seekers and refugees: a qualitative study of service user experiences in the UK. *Br J Gen Pract* 2019;69:e537–45.

6 Zhang CX, Boukari Y, Pathak N, *et al*. Migrants' primary care utilisation before and during the COVID-19 pandemic in England: an interrupted time series analysis. *Lancet Reg Health Eur* 2022;20:100455.

7 Deal A, Hayward SE, Huda M, *et al*. Strategies and action points to ensure equitable uptake of COVID-19 Vaccinations: a national qualitative interview study to explore the views of Undocumented migrants, asylum seekers, and refugees. *J Migr Health* 2021;4:100050.

8 Hayward SE, Deal A, Cheng C, *et al*. Clinical outcomes and risk factors for COVID-19 among migrant populations in high-income countries: a systematic review. *J Migr Health* 2021;3:100041.

9 Jaljaa A, Caminada S, Tosti ME, *et al*. Risk of SARS-CoV-2 infection in migrants and ethnic minorities compared with the general population in the European WHO region during the first year of the pandemic: a systematic review. *BMC Public Health* 2022;22:143.

10 Warszawski J, Beaumont A-L, Seng R, *et al*. Prevalence of SARS-Cov-2 antibodies and living conditions: the French national random population-based EPICOV cohort. *BMC Infect Dis* 2022;22:41.

11 Public Health England. Disparities in the risk and outcomes of COVID-19. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/908434/Disparities_in_the_risk_and_outcomes_of_COVID_August_2020_update.pdf [Accessed 28 Jul 2022].

12 Public Health England. Chapter 6: wider determinants of health. Available: https://www.gov.uk/government/publications/health-profile-for-england-2018/chapter-6-wider-determinants-of-health [Accessed 01 Aug 2022].

13 World Health Organization. Report of the systematic review on the effect of household crowding on health; Available: https://apps.who.int/iris/bitstream/handle/10665/275838/WHO-CED-PHE-18.02-eng.pdf?sequence=1&isAllowed=y

14 Aldridge RW, Pineo H, Fragaszy E, *et al*. Household overcrowding and risk of SARS-Cov-2: analysis of the virus watch prospective community cohort study in England and Wales [version 1; peer review: awaiting peer review]. *Wellcome Open Res* 2021;6:347.

15 Holt H, Talaei M, Greenig M, *et al*. Risk factors for developing COVID-19: a population-based longitudinal study (COVIDENCE UK). *Thorax* 2022;77:900–12.

16 Raisi-Estabragh Z, McCracken C, Bethell MS, *et al*. Greater risk of severe COVID-19 in black, Asian and minority ethnic populations is not explained by cardiometabolic, socioeconomic or behavioural factors, or by 25(OH)-vitamin D status: study of 1326 cases from the UK biobank. *J Public Health (Bangkok)* 2020;42:451–60.

17 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.

18 Martin CA, Jenkins DR, Minhas JS, *et al*. Socio-demographic heterogeneity in the prevalence of COVID-19 during Lockdown is associated with Ethnicity and household size: results from an observational cohort study. *EClinicalMedicine* 2020;25:100466.

19 Niedzwiedz CL, O'Donnell CA, Jani BD, *et al*. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK biobank. *BMC Med* 2020;18:160.

20 Soltan MA, Varney J, Sutton B, *et al*. COVID-19 admission risk tools should include multiethnic age structures, multimorbidity and deprivation Metrics for air pollution, household overcrowding, housing quality and adult skills. *BMJ Open Respir Res* 2021;8:e000951.

21 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.

22 Mathur R, Rentsch CT, Morton CE, *et al*. Ethnic differences in SARS-Cov-2 infection and COVID-19-related Hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the Opensafely platform. *Lancet* 2021;397:1711.

23 The migration Observatory. Migrants and housing in the UK: experiences and impacts. Available: https://migrationobservatory.ox.ac.uk/resources/briefings/migrants-and-housing-in-the-uk-experiences-and-impacts/ [Accessed 29 Jul 2022].

24 Kjøllesdal M, Skyrud K, Gele A, *et al*. The correlation between socioeconomic factors and COVID-19 among immigrants in Norway: a register-based study. *Scand J Public Health* 2022;50:52–60.

25 Lin Q, Paykin S, Halpern D, *et al*. Assessment of structural barriers and racial group disparities of COVID-19 mortality with spatial analysis. *JAMA Netw Open* 2022;5:e220984.

26 Gosselin A, Warszawski J, Bajos N, *et al*. Higher risk, higher protection: COVID-19 risk among immigrants in France-results from the population-based epicov survey. *Eur J Public Health* 2022;32:655–63.

27 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:e048042.

28 Public Health England. Sources of COVID-19 surveillance systems. Available: https://www.gov.uk/government/publications/national-covid-19-surveillance-reports/sources-of-covid-19-systems [Accessed 05 Aug 2022].

29 Cable N, Sacker A. Validating overcrowding measures using the UK household longitudinal study. *SSM Popul Health* 2019;8:100439.

30 UK Health Security Agency, Department of Health and Social Care. COVID-19: guidance on protecting people defined on medical grounds as extremely vulnerable. Available: https://www.gov.uk/government/publications/guidance-on-shielding-and-protecting-extremely-vulnerable-persons-from-covid-19 [Accessed 12 Aug 2022].

31 Office for National Statistics. National Statistics Postcode lookup. 2020. Available: https://geoportal.statistics.gov.uk/datasets/ab73ec2e38c04599b64b09b3fa1c3333/about [Accessed 12 Aug 2022].

32 Buis ML. Direct and indirect effects in a Logit model. *Stata J* 2010;10:11–29.

33 Katikireddi SV, Lal S, Carrol ED, *et al*. Unequal impact of the COVID-19 crisis on minority ethnic groups: a framework for understanding and addressing inequalities. *J Epidemiol Community Health* 2021;75:970–4.

34 Crawshaw AF, Farah Y, Deal A, *et al*. Defining the determinants of vaccine uptake and Undervaccination in migrant populations in Europe to improve routine and COVID-19 vaccine uptake: a systematic review. *Lancet Infect Dis* 2022;22:e254–66.

35 Zhao XS, Lynch JG, Chen QM. Reconsidering Baron and Kenny: myths and truths about mediation analysis. *J Consum Res* 2010;37:197–206.

36 NHS England. COVID-19 vaccination Statistics: week ending Sunday 25th April 2021. Available: https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2021/04/COVID-19-weekly-announced-vaccinations-29-April-2021.pdf [Accessed 12 Aug 2022].

37 van Zwieten A, Tennant PWG, Kelly-Irving M, *et al*. Avoiding overadjustment bias in social epidemiology through appropriate covariate selection: a primer. *J Clin Epidemiol* 2022;149:127–36.

38 Tinson A, Clair A. Better housing is crucial for our health and the COVID-19 recovery. Available: https://www.health.org.uk/sites/default/files/2021-01/2020%20-%20Better%20housing%20is%20crucial.pdf [Accessed 04 Aug 2022].

39 UK Health Security Agency. People with symptoms of a respiratory infection including COVID-19. Available: https://www.gov.uk/guidance/people-with-symptoms-of-a-respiratory-infection-including-covid-19 [Accessed 17 Aug 2022].

40 Wilson W, Barton C. Overcrowded housing (England). Available: https://researchbriefings.files.parliament.uk/documents/SN01013/SN01013.pdf [Accessed 04 Aug 2022].

41 World Health Organization. WHO housing and health guidelines. Available: https://www.who.int/publications/i/item/9789241550376 [Accessed 01 Aug 2022].

42 Marmot M, Allen J, Goldblatt P, *et al*. The pandemic, socioeconomic and health inequalities in England. In: *Build Back Fairer: The COVID-19 Marmot Review*. London: Institute of Health Equity, 2020.