Risk of Long Covid in people infected with SARS-CoV-2 after two doses of a COVID-19 vaccine: community-based, matched cohort study

Daniel Ayoubkhani,[1,2] Principal Statistician
Matthew L. Bosworth,[1] Senior Statistician
Sasha King,[3] Assistant Methodologist
Koen B. Pouwels,[4,5] Senior Researcher
Myer Glickman,[1] Head of Epidemiology, Climate and Global Health
Vahé Nafilyan,[1,6] Lead Statistician
Francesco Zaccardi,[2] Clinical Epidemiologist
Kamlesh Khunti,[2] Professor of Primary Care Diabetes and Vascular Medicine
Nisreen A. Alwan,[7,8,9] Associate Professor in Public Health
A. Sarah Walker,[4,10] Professor of Medical Statistics and Epidemiology

[1] Health Analysis and Life Events Division, Office for National Statistics, Newport, UK
[2] Leicester Real World Evidence Unit, Diabetes Research Centre, University of Leicester, Leicester, UK
[3] Methodology and Quality Directorate, Office for National Statistics, London, UK
[4] National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, University of Oxford, Oxford, UK
[5] Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK
[6] Faculty of Public Health, Environment and Society, London School of Hygiene & Tropical Medicine, London, UK
[7] School of Primary Care, Population Sciences and Medical Education, Faculty of Medicine, University of Southampton, Southampton, UK
[8] NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK
[9] NIHR Applied Research Collaboration (ARC) Wessex, Southampton, UK
[10] Nuffield Department of Medicine, University of Oxford, Oxford, UK

Corresponding author
Name: Mr Daniel Ayoubkhani
Address: Office for National Statistics, Government Buildings, Cardiff Road, Newport, UK, NP10 8XG
Phone: +44 1633 455825
Email: daniel.ayoubkhani@ons.gov.uk

Alternate corresponding author
Name: Dr Vahé Nafilyan
Address: Office for National Statistics, Government Buildings, Cardiff Road, Newport, UK, NP10 8XG
Phone: +44 1633 455046
Email: vahe.nafilyan@ons.gov.uk

Article type: Brief report
Running title: Long Covid after double-vaccination
Keywords: COVID-19, Coronavirus, Long Covid, Post-COVID Condition, Vaccination

Abstract

We investigated Long Covid incidence by vaccination status in a random sample of UK adults from April 2020 to November 2021. Persistent symptoms were reported by 9.5% of 3,090 breakthrough SARS-CoV-2 infections and 14.6% of unvaccinated controls (adjusted odds ratio 0.59, 95% CI: 0.50-0.69), emphasising the need for public health initiatives to increase population-level vaccine uptake.

Introduction

Long-term symptoms following SARS-CoV-2 infection, often referred to as Long Covid, post-acute COVID-19 syndrome, post-COVID condition, or post-acute sequelae of SARS-CoV-2, affect approximately 2% of the UK population, with two-thirds of these individuals experiencing functional impairment [1]. COVID-19 vaccines reduce rates of SARS-CoV-2 infection [2] and transmission [3] and therefore Long Covid incidence. However, it is unclear to what extent vaccination reduces the risk of developing Long Covid symptoms following breakthrough infection, with mixed evidence to date [4,5].

To 25 January 2022, 16% of the UK population eligible for a second vaccination were yet to receive it [6], while vaccine coverage was lowest in disadvantaged groups, including ethnic minorities and deprived communities, where rates of infection have been highest [7]. Understanding the role of vaccines in Long Covid may therefore aid public health messaging and facilitate informed decision-making regarding vaccine uptake. We investigated whether infection following two doses of a COVID-19 vaccine is associated with a reduction in Long Covid symptoms after 12 weeks, relative to being unvaccinated when infected, using prospective data from a large, random sample of the UK population with routine testing for SARS-CoV-2.

Methods

Study data and design

The main data source was the UK COVID-19 Infection Survey (CIS, ISRCTN21086382, https://www.ndm.ox.ac.uk/covid-19/covid-19-infection-survey/protocol-and-information-sheets), run by the Office for National Statistics (ONS) and comprising a sample of over half a million participants randomly selected from the UK community population (excluding

communal establishments such as hospitals, care homes, halls of residence, and prisons). During the pilot phase of the survey from April to August 2020, households were selected from previous respondents to ONS surveys who had consented to participate in future research, achieving an enrolment rate of 51%. From August 2020, sampling was conducted by random selection from national address lists, with the enrolment rate dropping to 12%. Participants were compensated with a £50 voucher at enrolment and a £25 voucher at each follow-up visit.

Ethical approval was obtained from the South Central Berkshire B Research Ethics Committee (20/SC/0195). At enrolment, adult participants provided written consent, including for optional weekly follow-up visits for one month followed by at least 12 monthly visits in the majority.

We included CIS participants aged 18-69 years who tested positive for SARS-CoV-2, either by polymerase chain reaction test using swabs obtained at study visits (58.7% of infections) or any swab test in national testing programmes (self-reported by study participants), between 26 April 2020 (the start of the CIS) and 30 November 2021 (the latest available data at the time of analysis). We excluded participants who: reported suspected COVID-19 or tested positive for antibodies (in the study or elsewhere) more than two weeks before their first positive swab; reported Long Covid symptoms at any time before their first positive swab; had never responded to the survey question on Long Covid (see 'Outcome' below) following its introduction on 3 February 2021; did not have ≥12 weeks of post-infection follow-up by 30 November 2021; or were single-vaccinated when infected.

Exposure

The exposure of interest was receipt of at least two doses of a COVID-19 vaccine (Oxford/AstraZeneca ChAdOx1 nCoV-19 [AZD1222], Pfizer/BioNTech BNT162b2, or Moderna mRNA-1273) ≥14 days before the first test-confirmed infection. Vaccination status

for participants in England was derived from survey data linked to National Immunisation Management System (NIMS) records, with the latter being prioritised where they conflicted with self-reports. Agreement rates between self-reported CIS data and NIMS records have previously been found to be high for both vaccination type (98%) and date (95% within one week) [8]. Administrative data were not available for participants in Wales, Scotland, and Northern Ireland (13.6%), thus vaccination status was derived solely from self-report. In sensitivity analysis, we restricted the analysis to participants living in England, thereby reducing the risk of exposure misclassification.

Outcome

The primary outcome was Long Covid status according to the survey question: "Would you describe yourself as having 'Long Covid', that is, you are still experiencing symptoms more than 4 weeks after you first had COVID-19, that are not explained by something else?" Participants were also asked whether their symptoms limited their ability to undertake daily activities. The survey questionnaire was administered by trained study workers during face-to-face interviews conducted at participants' homes. We considered participants' first response ≥12 weeks after their first test-confirmed infection. Follow-up time was calculated as the number of days from infection to the first response to the CIS question on Long Covid (either positive or negative) ≥12 weeks later.

Statistical methods

We matched study participants who were double-vaccinated at time of infection to control participants who were unvaccinated when infected and remained so at their first follow-up visit ≥12 weeks later. Double-vaccinated and unvaccinated participants were 1:1 propensity-score matched within calipers of 0.1 points of the propensity score on socio-demographic characteristics: single-year of age, sex, ethnicity (white or non-white), country/region of residence, area deprivation quintile group, and pre-existing health/disability status. To derive the latter, participants were asked: "Do you have any physical or mental health conditions or

illnesses lasting or expected to last 12 months or more (excluding any long-lasting COVID-19 symptoms)?" and "If yes, do any of your conditions or illnesses reduce your ability to carry-out day-to-day activities (a lot, a little, or not at all)?"

Although a 'post-treatment' variable, we also included time from infection to follow-up for Long Covid in the matching set to avoid evaluating Long Covid symptoms in unvaccinated and double-vaccinated participants at different stages of the illness. To assess the robustness of our results to this choice, we performed a sensitivity analysis excluding follow-up time from the matching set.

Continuous variables (age and follow-up time) were modelled as restricted cubic splines, with boundary knots at the 10^{th} and 90^{th} percentiles and an internal knot at the median of the distributions. Large imbalance after matching was identified by absolute standardized differences >10% [9]. We were not able to match on date of infection (a surrogate for SARS-CoV-2 variant); see the Discussion section.

We estimated adjusted odds ratios (aOR) for Long Covid at ≥12 weeks using logistic regression including all covariates from the matching set, comparing participants who were double-vaccinated to those unvaccinated (reference group) when infected, using robust standard errors to account for matching. We interacted the exposure variable (double-vaccinated versus unvaccinated) with time from infection to follow-up for Long Covid (continuous), and with adenovirus vector (Oxford/AstraZeneca) versus messenger ribonucleic acid (mRNA; Pfizer/BioNTech or Moderna) vaccines, to test for effect-modification using a likelihood-ratio test. Statistical analyses were performed using R version 3.6.

Results

Description of the study sample

Of 3,333 eligible participants who were double-vaccinated before their first test-confirmed SARS-CoV-2 infection, 3,090 (92.7%) were 1:1 matched to participants who were unvaccinated when infected (from a pool of 9,854 potential control participants). See **Supplementary Figure 1** for details of the study sample selection. Among double-vaccinated participants, 2,287 (74.0%), 788 (25.5%) and 15 (0.5%) received Oxford/AstraZeneca, Pfizer/BioNTech, and Moderna vaccines, respectively.

Most double-vaccinated participants (3,057, 98.9%) were infected after 17 May 2021, when the Delta variant dominated in the UK, while nearly all unvaccinated participants (3,082, 99.7%) were infected before this date **(Supplementary Figure 2)**. Median follow-up for Long Covid ≥12 weeks after infection among double-vaccinated and unvaccinated participants was 96 (IQR: 90 to 104) and 98 (89 to 109) days, respectively **(Supplementary Figure 3)**. After matching, socio-demographic characteristics were generally well balanced for all variables except age (mean 49 versus 47 years for double-vaccinated versus unvaccinated, absolute standardized difference 19.6%) **(Supplementary Table 1)**.

Long Covid symptoms at follow-up

Long Covid symptoms of any severity were reported by 294 double-vaccinated participants (prevalence 9.5%; 95% CI: 8.5% to 10.6%) versus 452 unvaccinated participants (14.6%; 13.4% to 15.9%), and activity-limiting symptoms by 170 (5.5%; 4.8% to 6.4%) and 268 (8.7%; 7.7% to 9.7%) participants, respectively.

The aOR were 0.59 (0.50 to 0.69) for Long Covid of any severity and 0.59 (0.48 to 0.73) for activity-limiting symptoms in those infected after double vaccination compared with those who were infected when unvaccinated **(Figure 1)**. There was no evidence of heterogeneity by time from infection to follow-up (p=0.65 for symptoms of any severity; p=0.68 for activity-limiting symptoms), or between participants receiving adenovirus vector or mRNA vaccines (p=0.25 for symptoms of any severity; p=0.35 for activity-limiting symptoms).

Sensitivity analysis demonstrated that the aOR increased when removing time from infection to follow-up for Long Covid from the matching set (to 0.68 [0.56 to 0.81] for the primary outcome), and further increased when it was also omitted from the covariate set in adjusted models (0.73 [0.62 to 0.85]) **(Supplementary Table 2)**. However, the aOR remained below 1 in all analyses.

The main analysis results were also insensitive to restricting the study sample to the 2,311 matched pairs (74.8%) for which both the double-vaccinated and unvaccinated participants lived in England (for whom NIMS data were available for linkage), with an aOR of 0.64 (0.53 to 0.78) for the primary outcome **(Supplementary Table 3)**, suggesting that exposure misclassification due to self-reporting of vaccination status is unlikely to have substantially impacted the main results.

Discussion

We found that receiving two COVID-19 vaccinations at least two weeks before SARS-CoV-2 infection was associated with a 41% decrease in the odds of developing Long Covid symptoms at least 12 weeks later, relative to not being vaccinated when infected. Our results extend those already published, whereby the risk of Long Covid was approximately halved in people who were double-vaccinated when infected compared with those who were

unvaccinated, but at four rather than 12 weeks post-infection [4]. Conclusions based on healthcare records rather than self-report (as in our study) are less clear, with vaccination associated with reduced rates of only specific symptoms [5] and diagnoses [10], though under-presentation, under-diagnosis, and under-recording are all possible [11].

The main study strength is that the CIS comprises a large sample of participants randomly selected from the population to minimise selection bias. Participants are routinely tested for SARS-CoV-2 at follow-up visits, therefore our study includes both asymptomatic and symptomatic infections, as well as self-reported tests. We considered participants' first monthly CIS response that was at least 12 weeks after their positive test for SARS-CoV-2, thus time from infection to response could have been any duration from 12 weeks upwards. However, recall bias was not a concern because participants were asked about their current Long Covid status at the time of the follow-up visit (that is, prospective data collection), and we included time from infection to response in the matching set to ensure balanced follow-up time between double-vaccinated and unvaccinated groups.

Although we adjusted for multiple factors related to vaccination uptake [7] and long-term symptoms [12], some unmeasured confounding may remain. In particular, because the question on Long Covid was not introduced until 3 February 2021, shortly after mass COVID-19 vaccination started in the UK on 8 December 2020, one key limitation is that it was not possible to match double-vaccinated and unvaccinated participants on calendar time of infection. Differences in the likelihood of developing Long Covid symptoms between exposure groups may therefore partly reflect changes in the dominant COVID-19 variant or other period effects, such as the introduction of NHS Long Covid assessment and rehabilitation services (though most patients are unlikely to be referred to these inside the first 12 weeks of illness).

Long Covid status was self-reported, so outcome misclassification was possible. Some participants may have been experiencing symptoms because of a health condition unrelated to COVID-19, while others who did have Long Covid may not have described themselves as such (for example, due to the perceived stigma attached to the term [13]). Conversely, self-recognition of Long Covid (participants' perception of the change in their own health compared with pre-infection) may be more reliable than electronic health records in some respects, for example due to differences in healthcare seeking behaviours between socio-demographic groups and Long Covid diagnoses being under-recorded in primary care [11]. Our key exposure was double vaccination, despite third and booster doses now being available, and the study period was before the Omicron variant became widespread. We were not able to investigate participants who were single-vaccinated when infected because nearly all of these received their second dose within the 12-week follow-up period, confounding any relationship between one dose at infection and Long Covid symptoms.

There is potential for survivor bias because our study sample did not include people who were infected but subsequently dropped out of the survey before having had the opportunity to respond to the Long Covid question after it was introduced on 3 February 2021. This loss-to-follow-up may be related to the likelihood of developing or reporting Long Covid symptoms, for example due to ill-health. However, after broadening the study cohort definition by dropping any exclusion criteria dependent on duration of follow-up after infection or response to the Long Covid question, just 3% of the resulting 37,145 participants never responded to the Long Covid question post-infection. Loss to follow-up is therefore unlikely to have materially impacted our findings.

In conclusion, SARS-CoV-2 infection after double vaccination is associated with a reduced risk of developing Long Covid symptoms at least 12 weeks later compared with infection before vaccination, emphasising the need for public health initiatives to increase population-level vaccine uptake. Studies with longer follow-up are needed to assess the impact of

booster doses and the Omicron variant and to evaluate symptom trajectories beyond a single 12-week follow-up visit, particularly given the relapsing nature of Long Covid [14]. Further research into possible biological explanations behind our findings, which may inform therapeutic strategies for Long Covid, is also required.

References

1. Office for National Statistics. Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK: 6 January 2022. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/6january2022. Accessed 10 May 2022.

2. Wei J, Pouwels KB, Stoesser N, et al. SARS-CoV-2 anti-spike IgG antibody responses after second dose of ChAdOx1 or BNT162b2 and correlates of protection in the UK general population. medRxiv [Preprint]. January 14, 2022. Available at: https://doi.org/10.1101/2021.09.13.21263487.

3. Eyre DW, Taylor D, Purver M, et al. The impact of SARS-CoV-2 vaccination on Alpha & Delta variant transmission. medRxiv [Preprint]. October 15, 2021. Available at: https://doi.org/10.1101/2021.09.28.21264260.

4. Antonelli M, Penfold RS, Merino J, et al. Risk factors and disease profile of post-vaccination SARS-CoV-2 infection in UK users of the COVID Symptom Study app: a prospective, community-based, nested, case-control study. Lancet Infect Dis, **2022** ; 22: 43-45.

5. Taquet M, Dercon Q, Harrison PJ. Six-month sequelae of post-vaccination SARS-CoV-2 infection: A retrospective cohort study of 10,024 breakthrough infections. Brain Behav Immun, **2022** ; 103: 154-162.

6. UK Government. Coronavirus (COVID-19) in the UK. Available at: https://coronavirus.data.gov.uk. Accessed 25 January, 2022.

7. Office for National Statistics. Coronavirus and vaccination rates in people aged 18 years and over by socio-demographic characteristic and occupation, England: 8 December 2020 to 31 December 2021. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/bulletins/coronavirusandvaccinationratesinpeopleaged18yearsandoverbysociodemographiccharacteristicandoccupationengland/8december2020to31december2021. Accessed 10 May, 2022.

8. Pritchard E, Matthews PC, Stoesser N, et al. Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom. Nat Med, **2021** ; 27: 1370-1378.

9. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med, **2009** ; 28: 3083-3107.

10. Al-Aly Z, Bowe B, Xie Y. Long Covid after Breakthrough COVID-19: the post-acute sequelae of breakthrough COVID-19. Research Square [Preprint]. 15 November, 2021. Available at: https://doi.org/10.21203/rs.3.rs-1062160/v1.

11. Walker AJ, MacKenna B, Inglesby P, et al. Clinical coding of long COVID in English primary care: a federated analysis of 58 million patient records *in situ* using OpenSAFELY. Br J Gen Pract, **2021** ; 71: e806-e814.

12. Thompson EJ, Williams DM, Walker AJ, et al. Long COVID burden and risk factors in 10 UK longitudinal studies and electronic health records. Nat Commun, **2022** ; 13: 3528.

13. Pantelic M, Ziauddeen N, Boyes M, et al. Long Covid stigma: estimating burden and validating scale in a UK-based sample. medRxiv [Preprint]. 26 May, 2022. Available at: https://doi.org/10.1101/2022.05.26.22275585.

14. Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. EClinicalMedicine, **2021** ; 38: 101019.

Funding

There was no dedicated funding for this study.

Acknowledgements

DA, KK and FZ are supported by the National Institute for Health Research (NIHR) Applied Research Collaboration East Midlands (ARC EM). KK and FZ are also supported by the NIHR Leicester Biomedical Research Centre (BRC). KBP and ASW are supported by the NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance (NIHR200915), a partnership between the UK Health Security Agency (UKHSA) and the University of Oxford. KBP is also supported by the Huo Family Foundation. ASW is also supported by the NIHR Oxford Biomedical Research Centre and is an NIHR Senior Investigator. NAA has lived experience of Long Covid and is a co-investigator on the NIHR-funded STIMULATE-ICP study. The views expressed are those of the authors and are not necessarily those of the National Health Service, the NIHR, the Department of Health and Social Care, or the UK Health Security Agency. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

Footnotes

Data availability: De-identified study data are available to accredited researchers in the ONS Secure Research Service (SRS) under part 5, chapter 5 of the Digital Economy Act 2017. For further information about accreditation, contact research.support@ons.gov.uk or visit:

https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme

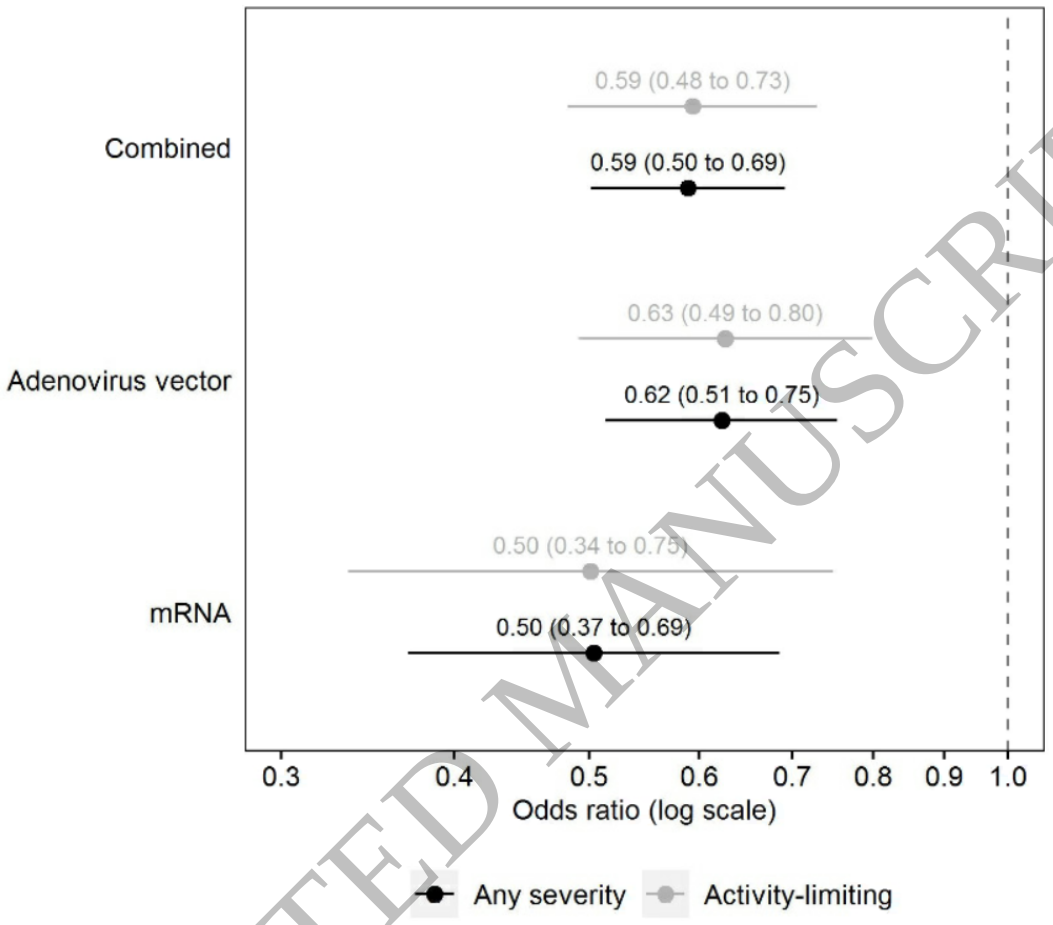
Author contributions: All authors contributed to conceptualising and designing the study. DA, MLB and SK prepared the study data and performed the statistical analysis. All authors contributed to interpretation of the results. DA, MLB and SK were responsible for the first draft of the manuscript. All authors contributed to critical revision of the manuscript. All authors approved the final manuscript.

Potential conflicts of interest: All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/disclosure-of-interest/ and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; KK chairs the Long Covid research-funded group reporting to the Chief Medical Officer, chairs the Ethnicity Subgroup of the UK Scientific Advisory Group for Emergencies (SAGE), and is a Member of SAGE.

Ethical approval: Ethical approval for this study was obtained from the National Statistician's Data Ethics Advisory Committee (NSDEC(20)12). The CIS received ethical approval from the South Central Berkshire B Research Ethics Committee (20/SC/0195).

Patient consent statement: All participants provided written consent at enrolment. Ethical approval for this study was obtained from the National Statistician's Data Ethics Advisory Committee (NSDEC(20)12). The CIS received ethical approval from the South Central Berkshire B Research Ethics Committee (20/SC/0195).

Figure 1: Adjusted odds ratios for Long Covid symptoms ≥12 weeks after first infection, comparing matched study participants who were double-vaccinated or unvaccinated (reference group) before infection



Odds ratios adjusted for socio-demographic characteristics (age, sex, white or non-white ethnicity, country/region of residence, area deprivation quintile group, and self-reported, pre-existing health/disability status) and time from infection to follow-up for Long Covid. Confidence intervals are at the 95% level.