Undervaccination and severe COVID-19 outcomes: meta-analysis of national cohort studies in England, Northern Ireland, Scotland, and Wales

*The HDR UK COALESCE Consortium**

Lancet 2024; 403: 554–66

Published Online January 15, 2024 https://doi.org/10.1016/S0140-6736(23)02467-4

See Comment page 508

*The full author list is provided at the end of the Article

Usher Institute, University of Edinburgh, Edinburgh, UK (S Kerr PhD, W Whiteley PhD, D Weatherill MEng, K Mooney BA, J Davies, Prof C Sudlow DPhil, Prof Sir A Sheikh MD); **Population Data Science Group, Swansea University Medical School, Faculty of Medicine, Health and Life Science, Swansea University, Swansea, UK** (S Bedston PhD, A Akbari MSc, Prof R A Lyons MD); **British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, and Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge, UK** (G Cezard PhD, A Sampri PhD, Prof A Wood PhD); **Centre for Public Health, Queen's University Belfast, Belfast, UK** (S Murphy PhD, D T Bradley PhD, L Patterson PhD, Prof F Kee PhD); **Public Health Agency, Belfast, UK** (D T Bradley, L Patterson, Prof F Kee); **Public Health Scotland, Glasgow, UK** (K Morrison PhD, C Sullivan PhD, J McMenamin MBChB, Prof C Robertson PhD); **Diabetes Research Centre, University of Leicester, UK** (Prof K Khunti PhD); **Institute of Health Informatics, University College London, London, UK** (Prof S Denaxas PhD); **British Heart Foundation Data Science Centre, Health Data Research UK, London, UK** (W Whiteley, Prof A Wood, Prof S Denaxas, T Bolton PhD, S Khan BA, A Keys FBCS, Prof C Sudlow); **Academic Primary Care, School of Medicine, Medical Sciences and Nutrition, University of**

Summary

Background Undervaccination (receiving fewer than the recommended number of SARS-CoV-2 vaccine doses) could be associated with increased risk of severe COVID-19 outcomes—ie, COVID-19 hospitalisation or death—compared with full vaccination (receiving the recommended number of SARS-CoV-2 vaccine doses). We sought to determine the factors associated with undervaccination, and to investigate the risk of severe COVID-19 outcomes in people who were undervaccinated in each UK nation and across the UK.

Methods We used anonymised, harmonised electronic health record data with whole population coverage to carry out cohort studies in England, Northern Ireland, Scotland, and Wales. Participants were required to be at least 5 years of age to be included in the cohorts. We estimated adjusted odds ratios for undervaccination as of June 1, 2022. We also estimated adjusted hazard ratios (aHRs) for severe COVID-19 outcomes during the period June 1 to Sept 30, 2022, with undervaccination as a time-dependent exposure. We combined results from nation-specific analyses in a UK-wide fixed-effect meta-analysis. We estimated the reduction in severe COVID-19 outcomes associated with a counterfactual scenario in which everyone in the UK was fully vaccinated on June 1, 2022.

Findings The numbers of people undervaccinated on June 1, 2022 were 26 985 570 (45·8%) of 58 967 360 in England, 938 420 (49·8%) of 1 885 670 in Northern Ireland, 1 709 786 (34·2%) of 4 992 498 in Scotland, and 773 850 (32·8%) of 2 358 740 in Wales. People who were younger, from more deprived backgrounds, of non-White ethnicity, or had a lower number of comorbidities were less likely to be fully vaccinated. There was a total of 40 393 severe COVID-19 outcomes in the cohorts, with 14 156 of these in undervaccinated participants. We estimated the reduction in severe COVID-19 outcomes in the UK over 4 months of follow-up associated with a counterfactual scenario in which everyone was fully vaccinated on June 1, 2022 as 210 (95% CI 94–326) in the 5–15 years age group, 1544 (1399–1689) in those aged 16–74 years, and 5426 (5340–5512) in those aged 75 years or older. aHRs for severe COVID-19 outcomes in the meta-analysis for the age group of 75 years or older were 2·70 (2·61–2·78) for one dose fewer than recommended, 3·13 (2·93–3·34) for two fewer, 3·61 (3·13–4·17) for three fewer, and 3·08 (2·89–3·29) for four fewer.

Interpretation Rates of undervaccination against COVID-19 ranged from 32·8% to 49·8% across the four UK nations in summer, 2022. Undervaccination was associated with an elevated risk of severe COVID-19 outcomes.

Funding UK Research and Innovation National Core Studies: Data and Connectivity.

Introduction

The UK COVID-19 vaccination programme began on Dec 8, 2020. As of June 1, 2022, the vaccine schedule recommended by the UK's Joint Committee on Vaccination and Immunisation (JCVI) consisted of one dose for those aged 5–11 years, two for 12–15 years, three for 16–74 years, and four for 75 years and older.[1,2] Additional doses were offered to population subgroups in autumn, 2022,[3] and spring, 2023.[4]

In the UK, uptake of the first vaccine dose was high; by January, 2022, more than 90% of the UK population older than 12 years had received a COVID-19 vaccine. However, as more doses have been offered, uptake has declined. In England, uptake of a second dose was lower among those older than 50 years, people of Black ethnicity, men, and those living in high deprivation and urban areas.[5] Vaccine hesitancy has been reported within groups at higher risk of severe COVID-19 outcomes.[5–7]

Licensed COVID-19 vaccines have been shown in clinical trials to be causally effective against infection, hospitalisation, and death, and have been associated with a reduction in adverse COVID-19 outcomes in observational epidemiological studies.[8–15] For example, a study from Israel found vaccine effectiveness of 76% (95% CI 72–79) against COVID-19 hospitalisation and 77% (70–83) against death 14–21 days after receipt of the first dose of the BNT162b2 vaccine, rising to 98% (98–99) 14 days or more after a second dose of BNT162b2.[13] A study in England found vaccine effectiveness of three doses relative to two doses of 85–95%.[14] For long-term

Research in context

Evidence before this study

We searched PubMed, medRxiv, and SSRN on May 22, 2023, using the terms "COVID-19 vaccine uptake" and "COVID-19 under vaccination" with no time restrictions. Several previous papers have studied vaccine uptake in the UK, broadly concluding that younger age, higher socioeconomic deprivation, and non-White ethnicities were associated with lower levels of vaccine uptake. We did not identify any population-level studies in the UK examining the association between undervaccination and severe COVID-19 outcomes.

Added value of this study

We found that undervaccination (receiving fewer than the recommended number of SARS-CoV-2 vaccine doses) was associated with increased risks of severe COVID-19 outcomes across all age groups studied. Our modelling allowed us to estimate that full vaccination (receiving the recommended number of SARS-CoV-2 vaccine doses) at the start of the study period would have been associated with a reduction of 7180 severe COVID-19 outcomes from a total of 40 393 severe events. To our knowledge, this is the first epidemiological analysis using individual-level electronic health records covering the entire population of the UK (aged 5 years or older).

Implications of all the available evidence

Our UK-wide analysis suggests that improved vaccination coverage during the pandemic would have been associated with fewer COVID-19 hospitalisations and deaths in the UK. Achieving target COVID-19 vaccination rates has the potential to reduce the incidence of severe COVID-19 outcomes in the future. The UK's health data environment now has—for the first time—the potential to support the undertaking of analyses on its entire population.

Aberdeen, Aberdeen, UK (Prof Sir L Ritchie MD); **Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK** (Prof C Robertson); **Nuffield Department of Primary Care Health Sciences, University of Oxford** (Prof Sir A Sheikh)

Correspondence to:
Prof Sir Aziz Sheikh, Usher Institute, University of Edinburgh, Edinburgh EH16 4SS, UK
aziz.sheikh@ed.ac.uk

care residents, receipt of a fourth vaccine dose was associated with strong protection of 40% (24–52) against severe COVID-19 outcomes compared with those who received a third dose 84 days or more previously.[15] The Oxford-AstraZeneca COVID-19 vaccine was not routinely used in the UK beyond the first and second doses.

Research has been conducted on the factors influencing COVID-19 vaccine uptake and the effectiveness of recommended doses. However, there is a notable gap in our understanding of the characteristics and COVID-19 outcomes of individuals who are undervaccinated (defined as having fewer than the number of doses recommended by the JCVI), and the association between undervaccination and severe COVID-19 outcomes. This knowledge is crucial to prevent exacerbation of existing inequalities and to provide insights for public health recommendations aimed at promoting vaccine uptake and improving health outcomes. The aim of this study, therefore, was to characterise factors associated with undervaccination and to investigate the association between undervaccination and severe COVID-19 outcomes.

Methods

Study design and population

We followed a prespecified statistical analysis plan (appendix pp 4–17). The datasets consisted of electronic health records with near whole-nation coverage in secure trusted research environments (TREs) in England, Northern Ireland, Scotland, and Wales (appendix p 18). In England, we accessed electronic health records through the National Health Service (NHS) England Secure Data Environment, via the British Heart Foundation (BHF) Data Science Centre's CVD-COVID-UK/COVID-IMPACT consortium. Datasets were linked using the NHS Digital Master Person Service,[16] which aims to match records from different datasets with 99% accuracy for each person to a single unique identifier, the NHS number. This process involves checking NHS numbers present in the records against personal details, such as age, sex, and postcode, which were stored within the Personal Demographics Service. If the NHS number was successfully verified, no further processing was undertaken. In cases where the NHS number could not be verified or was absent, the Master Person Service attempted to match the records to a single NHS number recorded in the Personal Demographics Service using details provided in the submitted data file with information stored in the Personal Demographics Service, yielding a match confidence score.[17] We did not have access to this score. In Northern Ireland and Scotland, the data were deterministically linked using unique patient identifiers—Health and Care Number in Northern Ireland, and Community Health Index in Scotland. In Wales, a combination of deterministic linkage based on NHS number and probabilistic linkage based on personal identifiers was used.[18] In Northern Ireland and Wales, we used the Secure Anonymised Information Linkage Databank.[18,19] In Wales, 80% of the population is linked to the databank. In Scotland, we used the Early Pandemic Evaluation and Enhanced Surveillance of COVID-19 platform.[20] These datasets contained linked primary care, secondary care, mortality, vaccination, and virological testing data from Pillar 1 (laboratory testing within NHS facilities) as well as Pillar 2 (community test facilities set up during the pandemic). A data linkage diagram for each of the four nations of the UK, along with further information on how study variables were obtained from the datasets, is in the appendix (p18).

Cohort follow-up started on June 1, 2022, and ended on Sept 30, 2022. The cohorts for England, Wales, and

See **Online** for appendix

For the **British Heart Foundation Data Science Centre's CVD-COVID-UK/COVID-IMPACT consortium** see https://bhfdatasciencecentre.org/areas/cvd-covid-uk-covid-impact/

Northern Ireland included all individuals registered with a general practitioner or family doctor. The cohort for Scotland included individuals registered with a general practitioner and with recent previous contact with health services via secondary care, lateral flow test or PCR testing, prescribing, vaccinations, or who died during follow-up. Anyone under the age of 5 years at the cohort start date was excluded. We could not carry out a single pooled analysis due to restrictions on individual-level data sharing between national TREs. Therefore, we carried out separate

	England	Northern Ireland	Scotland	Wales
Age group				
5–11	4 593 030 (92·5%)	171 850 (97·4%)	306 239 (79·6%)	168 230 (86·3%)
12–15	2 058 315 (71·1%)	84 660 (82·2%)	123 103 (53·1%)	64 490 (54·6%)
16–17	1 256 235 (91·9%)	44 790 (94·5%)	88 850 (83·0%)	39 200 (70·4%)
18–24	3 277 455 (63·8%)	110 180 (68·3%)	187 288 (49·6%)	82 380 (46·8%)
25–29	2 703 100 (60·8%)	80 570 (66·4%)	183 160 (54·4%)	65 180 (46·4%)
30–34	2 708 990 (56·9%)	84 900 (62·6%)	184 350 (50·3%)	70 390 (44·9%)
35–39	2 348 295 (51·2%)	80 650 (56·5%)	157 143 (43·2%)	61 100 (38·9%)
40–44	1 858 920 (43·4%)	65 790 (47·6%)	122 816 (35·3%)	47 880 (31·6%)
45–49	1 348 475 (35·3%)	50 850 (39·2%)	91 012 (27·3%)	37 080 (25·0%)
50–54	1 090 775 (26·5%)	40 410 (29·4%)	71 811 (18·8%)	32 070 (18·3%)
55–59	855 225 (21·0%)	31 870 (23·4%)	53 338 (13·6%)	25 920 (14·2%)
60–64	610 570 (17·0%)	22 390 (18·6%)	35 290 (10·0%)	17 480 (10·6%)
65–69	395 455 (13·4%)	13 870 (14·2%)	21 782 (7·3%)	10 880 (7·7%)
70–74	258 235 (9·7%)	8980 (10·8%)	14 206 (5·3%)	7 630 (5·5%)
75–79	828 490 (35·2%)	20 850 (29·8%)	29 957 (14·8%)	19 210 (16·7%)
80–84	365 730 (24·7%)	13 000 (29·1%)	19 379 (14·8%)	11 880 (16·2%)
≥85	428 280 (27·7%)	12 820 (31·6%)	20 062 (17·5%)	12 850 (18·9%)
BMI				
<18·5	1 669 840 (72·8%)	..	30 430 (47·8%)	13 750 (47·9%)
18·5–25	4 047 195 (37·5%)	..	145 841 (28·1%)	76 120 (26·1%)
25–30	2 812 845 (28·4%)	..	115 725 (17·9%)	62 220 (17·9%)
30–35	1 415 140 (26·8%)	..	64 384 (15·5%)	37 010 (16·4%)
35–40	575 395 (26·4%)	..	29 732 (15·5%)	17 090 (16·5%)
>40	369 195 (26·4%)	..	19 573 (15·1%)	9210 (16·3%)
Missing	16 095 960 (59·2%)	..	1 304 101 (43·1%)	..
Ethnicity				
White	18 005 235 (39·6%)	..	997 853 (30·3%)	672 460 (30·8%)
Asian	3 352 065 (59·8%)	..	53 737 (43·2%)	27 350 (43·5%)
Black	1 834 835 (74·6%)	..	19 222 (61·0%)	9770 (60·5%)
Mixed	893 990 (67·9%)	..	18 226 (56·3%)	15 630 (59·2%)
Other	1 510 010 (66·9%)	..	16 146 (57·5%)	10 670 (52·4%)
Unknown	1 389 435 (71·9%)	..	604 602 (40·8%)	37 980 (77·9%)
IMD quintile				
1—Most deprived	7 341 955 (60·4%)	224 330 (60·0%)	469 291 (46·1%)	217 590 (43·7%)
2	6 444 300 (52·0%)	207 750 (53·4%)	371 911 (37·6%)	170 490 (35·3%)
3	5 167 530 (43·5%)	193 540 (50·1%)	309 956 (31·9%)	147 230 (32·2%)
4	4 326 475 (37·8%)	178 210 (46%)	277 580 (28·1%)	123 720 (28·0%)
5—Least deprived	3 705 315 (33·3%)	134 610 (38·6%)	260 468 (26·4%)	114 810 (23·9%)
Number of risk groups				
0	23 152 160 (50·4%)	728 740 (63·6%)	1 411 655 (39·6%)	548 450 (38·0%)
1	2 637 180 (32·2%)	94 760 (35·2%)	237 198 (23·5%)	161 240 (27·6%)
2	657 245 (24·6%)	49 870 (27·8%)	40 213 (14·7%)	42 990 (21·2%)
3	266 075 (23·2%)	27 530 (23·2%)	12 857 (13·5%)	12 750 (16·8%)
4	135 145 (24·5%)	17 110 (22·0%)	5088 (14·5%)	5000 (15·6%)
≥5	137 770 (27·5%)	20 390 (21·5%)	2775 (16·5%)	3420 (16·1%)

(Table 1 continues on next page)

	England	Northern Ireland	Scotland	Wales
(Continued from previous page)				
Sex				
Female	12 808 520 (43·4%)	438 460 (46·8%)	780 972 (31·0%)	359 650 (30·3%)
Male	14 177 050 (48·1%)	499 960 (52·7%)	928 814 (37·5%)	414 200 (35·3%)
Urban or rural classification				
Urban	23 794 645 (48·2%)	594 650 (50·4%)	1 292 039 (36·3%)	573 540 (34·0%)
Rural	3 190 925 (32·9%)	343 780 (48·7%)	397 167 (28·6%)	200 310 (29·9%)
Total	26 985 570 (45·8%)	938 420 (49·8%)	1 709 786 (34·2%)	773 850 (32·8%)

Data are n (%). Data were collected on the cohort start date of June 1, 2022. Number of risk groups in Northern Ireland was based on number of different British National Formulary paragraphs prescribed. Counts are rounded to the nearest 5 in England, and to the nearest 10 in Northern Ireland and Wales. Denominators for subgroups are not reported for concision. IMD=index of multiple deprivation.

Table 1: **Numbers and proportions of undervaccinated people by nation**

cohort studies of undervaccination and severe COVID-19 outcomes in each nation, and combined these results in a fixed-effect meta-analysis using inverse variance weighting.

Exposure

Vaccine deficit was defined as the number of doses recommended by the JCVI minus the number of doses received. This included COVID-19 vaccines of any type licensed in the UK, ie, Pfizer-BioNTech (BNT162b2), Oxford-AstraZeneca (ChAdOx1), and Moderna (mRNA-1273). The standard recommended vaccine schedule during the study period was one dose for the age group 5–11 years, two doses for 12–15 years, three doses for 16–74 years, and four doses for 75 years and older (appendix p19). A small proportion of individuals were offered more vaccine doses than normal due to being in higher-risk clinical or demographic groups defined by the JCVI.[1] However, we could not reliably identify these individuals in our datasets. Therefore, we took vaccine deficit to be the standard number of doses recommended by age group minus the number of doses received.

Outcomes

Full vaccination was defined as having received the standard JCVI recommended vaccine schedule and undervaccination was defined as not having received the standard JCVI recommended vaccine schedule.

A severe COVID-19 outcome was defined as COVID-19 hospitalisation or death. COVID-19 hospitalisation was defined as hospitalisation with International Classification of Diseases version 10 codes for COVID-19 (U07.1, U07.2) recorded as the first cause of admission and the admission recorded as an emergency in the secondary care records. COVID-19 death was defined as death with International Classification of Diseases version 10 codes for COVID-19 recorded as the underlying cause of death.

Statistical analysis

The following variables were included as covariates in the common adjustment analysis: age group in years (5–11, 12–15, 16–17, 18–24, and then 5-year bands until 85 or older), sex, ethnicity (White, Black, Asian, Mixed, Other, Unknown), urban or rural classification (defined by different methodologies in each nation), quintiles of index of multiple deprivation (English Index of Multiple Deprivation,[21] Northern Ireland Multiple Deprivation Measure,[22] Scottish Index of Multiple Deprivation,[23] and Welsh Index of Multiple Deprivation[24]), and number of QCovid risk groups (0, 1, and ≥2 in those aged 5–15 years, and 0, 1, 2, 3, 4, and ≥5 in those aged ≥16 years).[25] Brief descriptions of the QCovid risk groups can be found in the appendix (pp 14–17).

Additional covariates included in the extended adjustment analysis differed by nation depending on data availability. These included categorical variables for: health board or region, last positive COVID-19 PCR test (no positive test, 0–13 weeks, 14–26 weeks, or ≥27 weeks), whether the individual had ever been on the shielding list, whether someone in the individual's household had ever been on the shielding list, number of people in the household (1, 2, 3–5, 6–10, or ≥11), number of COVID-19 PCR tests in the last 6 months (0, 1, 2, 3, 4–9, or ≥10), number of positive COVID-19 PCR tests in the last 6 months (0, 1, or ≥2), whether the individual had ever had a COVID-19 hospitalisation, and non-COVID-19 hospitalisation as a time-dependent variable taking a value of 1 if there was an admission, and 0 otherwise.

QCovid risk groups were derived from general practitioner data. Some QCovid risk groups were not available in England, Scotland, and Wales. In Scotland, this included: whether the individual had undergone a bone marrow or stem cell transplant in the last 6 months; whether they had received radiotherapy in the last 6 months; whether they had been prescribed immunosuppressants, oral steroids, or anti-leukotriene or long-acting β-2 agonists four or more times in the last 6 months; whether they had irritable bowel syndrome; and whether they had received a solid organ transplant. In England, the chemotherapy and housing category QCovid risk groups were not available. In Wales, QCovid risk groups for HIV/AIDS were not

available. For QCovid risk groups that were unavailable, everyone in the cohort was assigned to be absent of the risk group. Ethnicity and general practitioner data (appendix p 18) were not available in Northern Ireland. The number of chapters of the British National Formulary (BNF)[26] from which individuals received repeat prescriptions before the vaccination programme was used as a proxy for comorbidity. To be included in the BNF prescription count, a medicine had to be prescribed in each of the two 3-month periods in the 6 months before the study time period. Medications related to contraceptives[26] were removed as these do not indicate an illness. This method was adapted from an approach validated in other multimorbidity studies using administrative data.[27] In Scotland, a small proportion (0·4%) of people did not have data available

	England	Northern Ireland	Scotland	Wales	Meta-analysis
Ethnicity, age 5–11 years					
White	Reference	..	Reference	Reference	Reference
Asian	0·78 (0·77–0·79)	..	0·55 (0·53–0·57)	0·57 (0·54–0·61)	0·75 (0·75–0·76)
Black	2·25 (2·19–2·30)	..	1·11 (1·03–1·20)	0·88 (0·78–1·01)	2·06 (2·02–2·11)
Mixed	1·21 (1·19–1·23)	..	1·17 (1·11–1·24)	0·94 (0·87–1·00)	1·19 (1·17–1·21)
Other	1·10 (1·08–1·12)	..	1·15 (1·05–1·26)	1·18 (1·04–1·34)	1·11 (1·09–1·13)
Number of risk groups, age 5–11 years					
0	Reference	Reference	Reference	Reference	Reference
1	0·68 (0·67–0·68)	0·42 (0·38–0·46)	0·74 (0·72–0·77)	0·64 (0·61–0·66)	0·68 (0·68–0·69)
≥2	0·38 (0·37–0·39)	0·19 (0·17–0·21)	0·48 (0·42–0·54)	0·39 (0·36–0·43)	0·39 (0·38–0·39)
Sex, age 5–11 years					
Female	Reference	Reference	Reference	Reference	Reference
Male	1·02 (1·01–1·02)	0·92 (0·86–0·97)	1·02 (1·00–1·03)	1·02 (1·00–1·05)	1·02 (1·01–1·02)
IMD quintile, age 5–11 years					
1—Most deprived	Reference	Reference	Reference	Reference	Reference
2	0·73 (0·73–0·74)	0·78 (0·70–0·87)	0·79 (0·77–0·81)	0·83 (0·79–0·86)	0·75 (0·74–0·76)
3	0·59 (0·59–0·60)	0·63 (0·56–0·70)	0·63 (0·61–0·64)	0·82 (0·78–0·86)	0·61 (0·60–0·62)
4	0·51 (0·51–0·52)	0·47 (0·43–0·53)	0·54 (0·53–0·56)	0·71 (0·68–0·74)	0·53 (0·52–0·53)
5—Least deprived	0·41 (0·41–0·42)	0·30 (0·27–0·33)	0·41 (0·40–0·42)	0·48 (0·47–0·50)	0·42 (0·41–0·42)
Urban or rural classification, age 5–11 years					
Urban	Reference	Reference	Reference	Reference	Reference
Rural	0·84 (0·83–0·85)	1·14 (1·06–1·21)	1·16 (1·14–1·18)	1·45 (1·41–1·5)	0·92 (0·92–0·93)
Ethnicity, age 12–15 years					
White	Reference	..	Reference	Reference	Reference
Asian	1·19 (1·18–1·20)	..	0·81 (0·77–0·85)	0·78 (0·73–0·83)	1·17 (1·16–1·18)
Black	2·99 (2·95–3·04)	..	1·40 (1·28–1·53)	1·76 (1·55–2·00)	2·89 (2·85–2·94)
Mixed	1·57 (1·55–1·59)	..	1·37 (1·27–1·47)	1·48 (1·38–1·59)	1·56 (1·54–1·58)
Other	1·69 (1·66–1·72)	..	1·78 (1·60–1·99)	1·21 (1·09–1·35)	1·68 (1·66–1·71)
Number of risk groups, age 12–15 years					
0	Reference	Reference	Reference	Reference	Reference
1	0·86 (0·85–0·87)	0·67 (0·63–0·71)	0·87 (0·84–0·89)	0·91 (0·88–0·94)	0·86 (0·86–0·87)
≥2	0·71 (0·69–0·73)	0·49 (0·45–0·53)	0·68 (0·61–0·75)	0·78 (0·72–0·84)	0·71 (0·70–0·73)
Sex, age 12–15 years					
Female	Reference	Reference	Reference	Reference	Reference
Male	1·03 (1·02–1·04)	1·07 (1·04–1·11)	1·06 (1·04–1·08)	1·08 (1·05–1·10)	1·04 (1·03–1·04)
IMD quintile, age 12–15 years					
1—Most deprived	Reference	Reference	Reference	Reference	Reference
2	0·71 (0·70–0·71)	0·73 (0·69–0·77)	0·72 (0·70–0·74)	0·73 (0·71–0·76)	0·71 (0·71–0·72)
3	0·53 (0·53–0·54)	0·62 (0·58–0·65)	0·53 (0·51–0·54)	0·65 (0·63–0·68)	0·54 (0·54–0·54)
4	0·42 (0·42–0·43)	0·45 (0·42–0·47)	0·41 (0·40–0·42)	0·55 (0·53–0·57)	0·43 (0·42–0·43)
5—Least deprived	0·33 (0·32–0·33)	0·29 (0·28–0·31)	0·03 (0·29–0·31)	0·38 (0·37–0·4)	0·32 (0·32–0·33)
Urban or rural classification, age 12–15 years					
Urban	Reference	Reference	Reference	Reference	Reference
Rural	0·84 (0·83–0·84)	0·98 (0·95–1·01)	0·98 (0·96–1·00)	1·03 (1·00–1·06)	0·86 (0·86–0·87)

(Table 2 continued on next page)

	England	Northern Ireland	Scotland	Wales	Meta-analysis
(Continued from previous page)					
Age group, age 16–74 years					
18–24 years	Reference	Reference	Reference	Reference	Reference
16–17 years	7·70 (7·65–7·75)	8·26 (7·93–8·61)	6·01 (5·90–6·12)	2·84 (2·78–2·90)	6·95 (6·91–6·99)
25–29 years	0·84 (0·84–0·84)	0·92 (0·91–0·94)	1·02 (1·01–1·03)	0·97 (0·95–0·98)	0·85 (0·85–0·86)
30–34 years	0·73 (0·73–0·73)	0·79 (0·77–0·80)	0·84 (0·83–0·85)	0·9 (0·89–0·91)	0·74 (0·74–0·74)
35–39 years	0·58 (0·58–0·58)	0·62 (0·61–0·63)	0·63 (0·63–0·64)	0·67 (0·66–0·68)	0·59 (0·58–0·59)
40–44 years	0·43 (0·43–0·43)	0·45 (0·45–0·46)	0·46 (0·46–0·47)	0·48 (0·47–0·48)	0·43 (0·43–0·43)
45–49 years	0·31 (0·31–0·31)	0·34 (0·34–0·35)	0·33 (0·32–0·33)	0·34 (0·33–0·35)	0·31 (0·31–0·31)
50–54 years	0·21 (0·21–0·21)	0·23 (0·23–0·24)	0·22 (0·22–0·22)	0·23 (0·23–0·23)	0·21 (0·21–0·22)
55–59 years	0·16 (0·16–0·16)	0·18 (0·18–0·18)	0·16 (0·16–0·16)	0·17 (0·17–0·17)	0·16 (0·16–0·16)
60–64 years	0·13 (0·13–0·13)	0·14 (0·14–0·15)	0·12 (0·11–0·12)	0·12 (0·12–0·13)	0·13 (0·13–0·13)
65–69 years	0·11 (0·11–0·11)	0·11 (0·11–0·11)	0·08 (0·08–0·09)	0·09 (0·09–0·09)	0·10 (0·10–0·10)
70–74 years	0·08 (0·08–0·08)	0·09 (0·08–0·09)	0·06 (0·06–0·06)	0·06 (0·06–0·07)	0·08 (0·08–0·08)
Ethnicity, age 16–74 years					
White	Reference	..	Reference	Reference	Reference
Asian	1·61 (1·61–1·62)	..	1·3 (1·28–1·32)	1·42 (1·39–1·45)	1·60 (1·60–1·61)
Black	3·74 (3·73–3·76)	..	2·32 (2·26–2·39)	2·85 (2·74–2·96)	3·71 (3·7–3·72)
Mixed	1·96 (1·95–1·97)	..	1·43 (1·39–1·48)	1·82 (1·76–1·88)	1·95 (1·94–1·96)
Other	2·31 (2·30–2·32)	..	2·14 (2·08–2·21)	2·03 (1·96–2·10)	2·31 (2·30–2·31)
Number of risk groups, age 16–74 years					
0	Reference	Reference	Reference	Reference	Reference
1	0·64 (0·64–0·64)	0·54 (0·53–0·55)	0·82 (0·82–0·83)	0·98 (0·97–0·99)	0·67 (0·67–0·67)
2	0·58 (0·58–0·58)	0·49 (0·48–0·50)	0·77 (0·76–0·79)	0·96 (0·95–0·98)	0·61 (0·61–0·61)
3	0·59 (0·58–0·59)	0·47 (0·46–0·48)	0·80 (0·77–0·83)	0·94 (0·91–0·96)	0·61 (0·61–0·62)
4	0·62 (0·62–0·63)	0·44 (0·43–0·45)	0·85 (0·80–0·90)	0·98 (0·94–1·03)	0·65 (0·64–0·65)
≥5	0·67 (0·66–0·68)	0·40 (0·39–0·41)	0·85 (0·76–0·94)	1·00 (0·93–1·07)	0·68 (0·67–0·69)
Sex, age 16–74 years					
Female	Reference	Reference	Reference	Reference	Reference
Male	1·29 (1·29–1·29)	1·32 (1·31–1·33)	1·31 (1·30–1·32)	1·30 (1·29–1·30)	1·29 (1·29–1·30)
IMD quintile, age 16–74 years					
1—Most deprived	Reference	Reference	Reference	Reference	Reference
2	0·71 (0·71–0·71)	0·75 (0·74–0·76)	0·69 (0·68–0·69)	0·74 (0·73–0·75)	0·71 (0·71–0·71)
3	0·54 (0·54–0·55)	0·64 (0·63–0·65)	0·52 (0·51–0·52)	0·65 (0·64–0·65)	0·55 (0·55–0·55)
4	0·43 (0·43–0·43)	0·51 (0·51–0·52)	0·40 (0·39–0·40)	0·53 (0·52–0·54)	0·43 (0·43–0·43)
5—Least deprived	0·33 (0·33–0·33)	0·35 (0·34–0·35)	0·30 (0·30–0·31)	0·41 (0·41–0·42)	0·33 (0·33–0·33)
Urban or rural classification, age 16–74 years					
Urban	Reference	Reference	Reference	Reference	Reference
Rural	0·74 (0·74–0·74)	0·88 (0·87–0·89)	0·93 (0·92–0·93)	1·01 (1·00–1·02)	0·77 (0·77–0·78)
Age group, age ≥75 years					
75–79 years	Reference	Reference	Reference	Reference	Reference
80–84 years	0·56 (0·56–0·57)	0·99 (0·96–1·01)	0·97 (0·95–0·99)	0·94 (0·92–0·97)	0·60 (0·60–0·60)
≥85 years	0·67 (0·67–0·67)	1·12 (1·09–1·15)	1·14 (1·12–1·16)	1·10 (1·07–1·13)	0·71 (0·70–0·71)
Ethnicity, age ≥75 years					
White	Reference	..	Reference	Reference	Reference
Asian	3·64 (3·60–3·67)	..	4·07 (3·73–4·43)	2·06 (1·90–2·24)	3·61 (3·58–3·65)
Black	6·89 (6·78–7·00)	..	4·13 (3·19–5·36)	3·50 (2·87–4·28)	6·85 (6·74–6·96)
Mixed	2·86 (2·79–2·93)	..	3·11 (2·37–4·07)	2·23 (1·79–2·78)	2·85 (2·78–2·92)
Other	3·00 (2·94–3·05)	..	3·49 (2·85–4·28)	2·40 (1·96–2·93)	2·99 (2·94–3·05)

(Table 2 continued on next page)

	England	Northern Ireland	Scotland	Wales	Meta-analysis
(Continued from previous page)					
Number of risk groups, age ≥75 years					
0	Reference	Reference	Reference	Reference	Reference
1	0·87 (0·87–0·87)	0·45 (0·43–0·47)	1·06 (1·04–1·09)	0·92 (0·90–0·95)	0·88 (0·88–0·88)
2	0·91 (0·90–0·91)	0·43 (0·42–0·45)	1·17 (1·14–1·19)	0·97 (0·94–1·00)	0·92 (0·92–0·93)
3	0·97 (0·97–0·98)	0·44 (0·42–0·46)	1·28 (1·24–1·31)	1·04 (1·01–1·08)	0·99 (0·98–1·00)
4	1·06 (1·05–1·07)	0·44 (0·43–0·46)	1·37 (1·32–1·43)	1·05 (1·01–1·10)	1·07 (1·06–1·08)
≥5	1·22 (1·21–1·23)	0·47 (0·45–0·49)	1·56 (1·49–1·64)	1·17 (1·11–1·23)	1·23 (1·22–1·24)
Sex, age ≥75 years					
Female	Reference	Reference	Reference	Reference	Reference
Male	0·87 (0·87–0·87)	0·83 (0·81–0·85)	0·84 (0·82–0·85)	0·79 (0·77–0·81)	0·86 (0·86–0·87)
IMD quintile, age ≥75 years					
1—Most deprived	Reference	Reference	Reference	Reference	Reference
2	0·76 (0·75–0·76)	0·75 (0·72–0·78)	0·74 (0·72–0·76)	0·79 (0·76–0·81)	0·76 (0·75–0·76)
3	0·62 (0·61–0·62)	0·7 (0·68–0·73)	0·65 (0·64–0·67)	0·71 (0·69–0·74)	0·62 (0·62–0·63)
4	0·53 (0·53–0·54)	0·59 (0·57–0·61)	0·55 (0·53–0·56)	0·61 (0·59–0·63)	0·54 (0·53–0·54)
5—Least deprived	0·45 (0·45–0·45)	0·49 (0·48–0·51)	0·44 (0·43–0·45)	0·46 (0·45–0·48)	0·45 (0·45–0·45)
Urban or rural classification, age ≥75 years					
Urban	Reference	Reference	Reference	Reference	Reference
Rural	0·90 (0·89–0·90)	1·20 (1·17–1·22)	1·02 (1·00–1·03)	1·02 (0·01–1·05)	0·92 (0·91–0·92)

Data are OR (95% CI). Data were collected on the cohort start date of June 1, 2022. ORs were adjusted multifactorially, with all variables included in the model simultaneously. Number of risk groups in Northern Ireland was based on number of different British National Formulary paragraphs prescribed. IMD=index of multiple deprivation. OR=odds ratio.

Table 2: **Multifactorial adjusted ORs for undervaccination by age group, individual nation estimates, and meta-analysis**

on their area of residence, and they were excluded from the analysis. There were missing values for ethnicity in England, Scotland, and Wales, and these were placed in the Unknown category. There were no other variables with missing values in our analyses.

In each nation, we created tables of undervaccination for a range of clinical and demographic groups, as well as cumulative plots of number of doses received and number of vaccinations by week and dose number, both stratified by age group in years (5–11, 12–15, 16–74, and ≥75).

In each nation, we separately fitted logistic regressions in the age groups 5–11 years, 12–15 years, 16–74 years, and 75 years or older with undervaccination as the dependent variable. We then fitted Cox models with time to severe COVID-19 outcome as the dependent variable in the age groups 5–15 years, 16–74 years, and 75 years or older. Individuals were censored at non-COVID-19 death, deregistration, or end of the study period. Vaccine deficit was included as a time-dependent exposure, changing levels on the date an individual received a vaccine dose that put them into a different category. We carried out analyses with a common set of adjustments, and an extended analysis that included further adjustments using additional variables that varied by nation depending on availability. In England, due to the significant computational demands of running the analyses in the very large English population, the Cox model included all cases (individuals who had a severe COVID-19 outcome) and 50 controls selected randomly without replacement per case, with weighting to account for this sampling. Cases were assigned a weight of 1, and controls were assigned a weight equal to the inverse of their probability of being sampled. We made plots and carried out tests of the proportional hazards assumption for the Cox models using Schoenfeld residuals. We calculated variance inflation factors to assess collinearity, as well as performance and model selection metrics including concordance and the Akaike and Bayesian information criteria.

We carried out a fixed-effect meta-analysis of the nation-specific results from both the logistic and Cox models with common adjustments using inverse variance weighting, as the same methods and data definitions were used in each nation and similar effects were anticipated. We calculated I^2 heterogeneity statistics. We estimated the reduction in severe COVID-19 outcomes by the end of follow-up that would have been associated with a counterfactual scenario in which everyone was fully vaccinated on June 1, 2022, using adjusted hazard ratios (aHRs) to approximate risk ratios (appendix p 19).

The number of individuals in the cohort for England exceeded the population size estimated by the 2021 UK census, mainly due to migration and some individuals registered at multiple general practices. We performed a sensitivity analysis by applying weights by age, sex, and

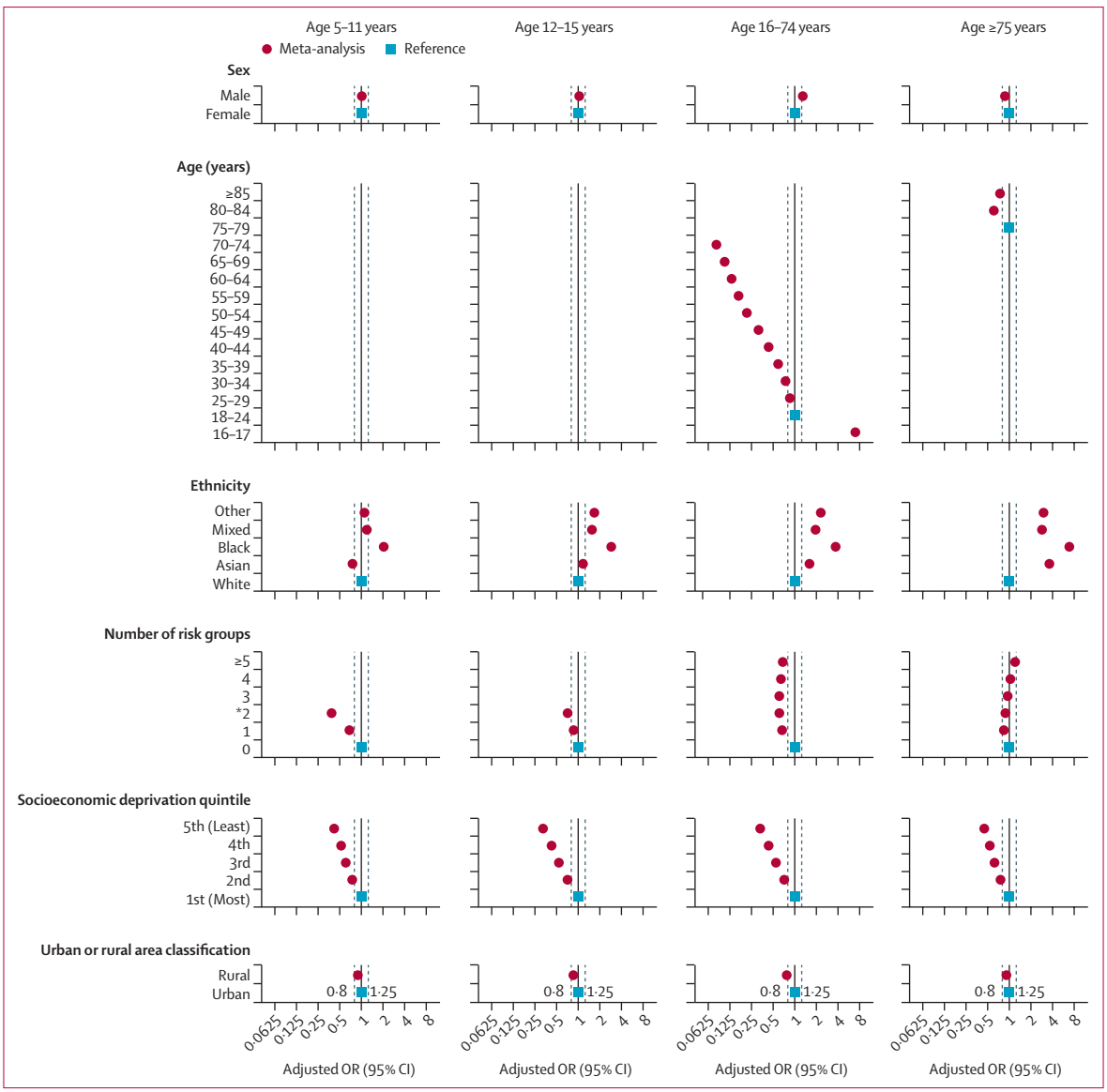

Figure 1: **Multifactorial adjusted ORs for undervaccination on June 1, 2022: meta-analysis of individual nation estimates**
Ethnicity data were not available in Northern Ireland. Number of risk groups in Northern Ireland was based on number of different British National Formulary paragraphs prescribed. OR=odds ratio. *Indicates ≥2 risk groups in the 5–11 and 12–15 age groups, and exactly 2 in the other age groups.

geographical region so that numbers reflected the most recent UK census (appendix pp 4–17).

Ethics and permissions

An ethics and permissions statement can be found in the appendix (p 20).

Reporting

This study is reported in accordance with the Reporting of Studies Conducted using Observational Routinely-Collected Data guidelines (appendix pp 21–25).[28,29] Patient and public contributors were involved in the design, interpretation, and reporting of this study (appendix pp 26–27).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

The cohorts consisted of 58·9 million individuals in England, 1·9 million in Northern Ireland, 5·0 million in Scotland, and 2·4 million in Wales. Table 1 shows undervaccination in population subgroups by country. The number of individuals undervaccinated as of June 1, 2022 was 26 985 570 (45·8%) in England, 938 420 (49·8%) in Northern Ireland, 1 709 786 (34·2%) in

	Number of events	Person-time, 1000 person-years	Event rate, per 1000 person-years	Unadjusted HR (95% CI)	Adjusted HR (95% CI)
5–15 years age group					
0	75	538	0·14	Reference	Reference
1	482	1737	0·28	1·62 (1·30–2·02)	1·28 (0·98–1·67)
2	135	521	0·26	1·59 (1·22–2·06)	2·41 (1·76–3·30)
16–74 years age group					
0	13 796	10 663	1·29	Reference	Reference
1	2163	2755	0·79	0·59 (0·56–0·62)	1·26 (1·19–1·32)
2	554	480	1·15	0·86 (0·79–0·93)	1·88 (1·71–2·06)
3	2404	3523	0·68	0·53 (0·51–0·55)	1·50 (1·42–1·57)
≥75 years age group					
0	12 361	1580	7·83	Reference	Reference
1	5968	326	18·33	2·61 (2·54–2·69)	2·70 (2·61–2·78)
2	1075	44	24·66	3·16 (2·97–3·36)	3·13 (2·93–3·34)
3	196	7	27·92	3·74 (3·25–4·30)	3·61 (3·13–4·17)
4	1184	83	14·24	1·80 (1·69–1·91)	3·08 (2·89–3·29)

Counts are rounded to the nearest 5 in England and to the nearest 10 in Scotland, Northern Ireland, and Wales. Counts of below 10 were suppressed in accordance with statistical disclosure rules implemented by trusted research environments. Counts suppressed in this manner were imputed as 5 when calculating totals. In the 5–15 years age group, event counts exclude Northern Ireland and Wales due to low numbers. Adjustments were included for: age group, sex, ethnicity, urban or rural classification, deprivation, and number of risk groups. Number of risk groups in Northern Ireland was based on number of different British National Formulary paragraphs prescribed. HR=hazard ratio.

Table 3: **Severe COVID-19 (COVID-19 hospitalisation or death) events in each age group by vaccine deficit**

Scotland, 773 850 (32·8%) in Wales, and 30 407 626 (44·4%) in the population aggregated across all four nations. Weekly vaccinations by dose and age group, cumulative vaccine uptake, and undervaccination plots are shown in the appendix (pp 28–39).

We found that higher levels of deprivation, lower number of risk groups in those aged 5–74 years, non-White ethnicity, and being male in the age range 5–74 years were generally associated with higher chances of being undervaccinated in the common adjustment analysis (table 2, figure 1). Effect sizes were similar across countries for level of deprivation, age, and sex. There was some heterogeneity across countries in effect size estimates associated with ethnicity, particularly for the Black and Asian categories in the 5–11 years and 12–15 years age groups.

Adjusted odds ratio gradients across different ethnic groups were more pronounced in older compared with younger age groups. Younger age was strongly associated with undervaccination. The results in each country were broadly similar for the common and extended adjustment analyses (appendix pp 40–45).

There were 33 885 severe COVID-19 events in England, 1220 in Northern Ireland, 3718 in Scotland, and 1570 in Wales. Event counts in undervaccinated and fully vaccinated people by nation and age group can be found in the appendix (p 46). A cumulative risk plot by week aggregated across the four nations can be seen in the appendix (p 49). aHRs for a composite outcome of COVID-19 hospitalisation or death in each age group are shown for the common adjustment analysis in table 3 and figure 2 (full results in appendix pp 46–83). Vaccine deficit was associated with an elevated risk of severe COVID-19 outcomes in all age groups and in all countries as well as the meta-analysis, particularly in those aged 75 years and older. aHRs for severe COVID-19 outcomes in the meta-analysis for the 16–74 years age group were 1·26 (95% CI 1·19–1·32) for a vaccine deficit of one dose, 1·88 (1·71–2·06) for two doses, and 1·50 (1·42–1·57) for three doses. In the 75 years and older age group, these were 2·70 (2·61–2·78) for a vaccine deficit of one dose, 3·13 (2·93–3·34) for two doses, 3·61 (3·13–4·17) for three doses, and 3·08 (2·89–3·29) for four doses. In Northern Ireland and Wales, there was an insufficient number of events in the 5–15 years age group to estimate the aHRs. In some analyses in Wales, there was also an insufficient number of events to estimate aHRs for some or all ethnic groups. The sensitivity analysis using population weights in England made little difference to the results.

A higher number of risk groups was strongly associated with severe COVID-19 outcomes across all age groups in individual country analyses and the meta-analysis. Results in each nation were broadly similar in the common and extended adjustment analyses; in particular, aHRs associated with vaccine deficit were not affected by the additional adjustments in each country (appendix pp 51–63 for common adjustment analysis, pp 64–80 for extended adjustment analyses). Meta-analysis estimates and I^2 heterogeneity statistics can be seen in the appendix (pp 81–83). Many of the I^2 statistics were high. Schoenfeld residuals, variance inflation factors, and model performance metrics from fitted models are available upon request from the corresponding author.

The expected reduction in number of severe COVID-19 outcomes over 4 months of follow-up, associated with a counterfactual scenario in which everyone was fully vaccinated on June 1, 2022, by age group and country are in the appendix (p 50). There was little difference in this calculation between the common and extended adjustment analyses. Based on the extended adjustment analysis, we estimated that if everyone had been fully vaccinated on June 1, 2022, and with all else equal, this would have been associated with 210 (95% CI 94–326) fewer events in the 5–15 years age group in England and Scotland, 1544 (1399–1689) fewer events in the 16–74 years age group in total, and 5426 (5340–5512) fewer events in the 75 years and older age group in total.

Discussion

In this study of COVID-19 vaccination across the whole UK population (aged 5 years and older), we found that undervaccination was associated with higher risk of severe COVID-19 outcomes relative to full vaccination. We estimate that in a counterfactual scenario in which everyone in the UK was fully vaccinated on June 1, 2022, there would have been an associated reduction in severe

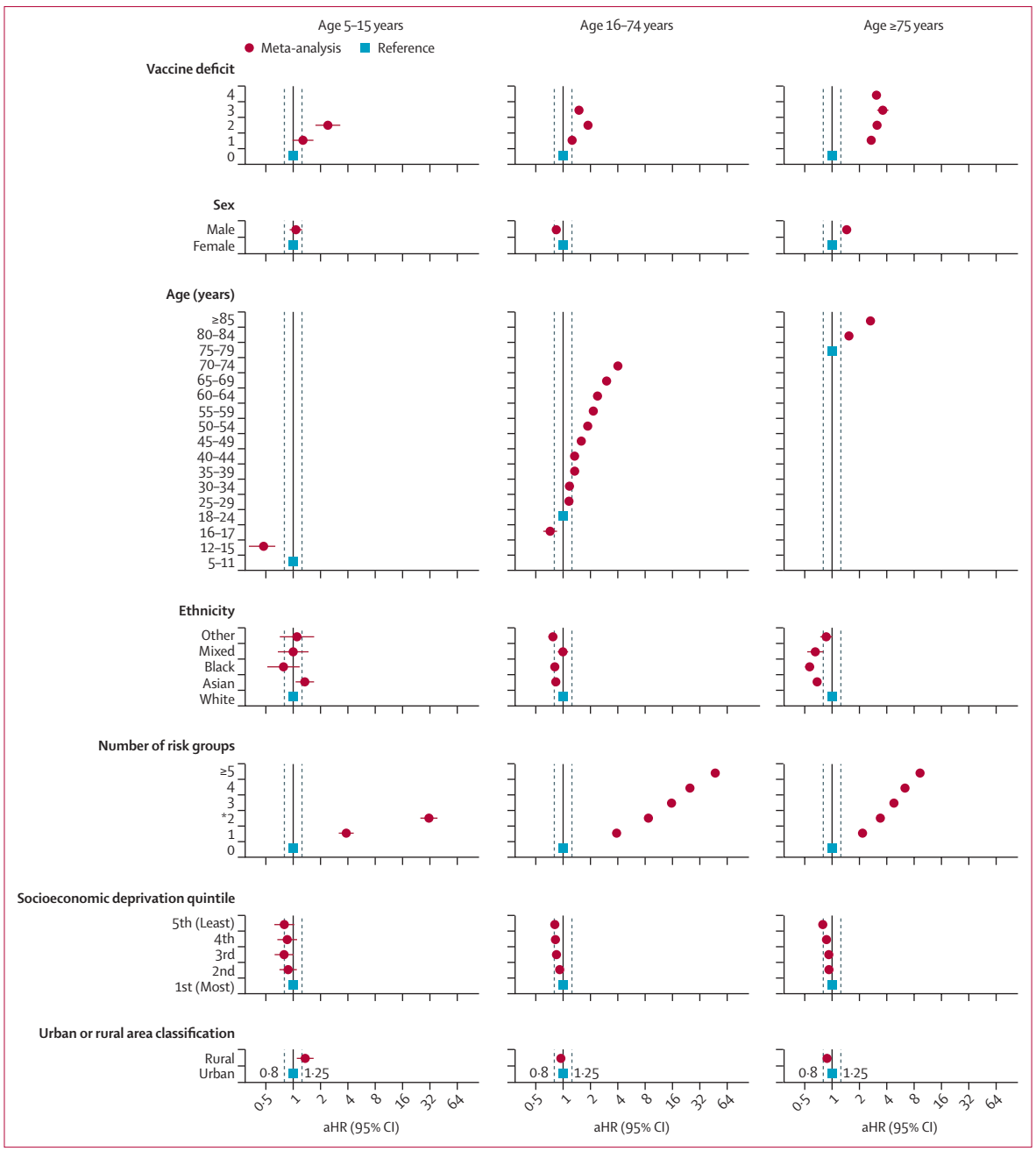

Figure 2: **Multifactorial aHRs for COVID-19 hospitalisation or death: meta-analysis of individual nation estimates**
Ethnicity data were not available in Northern Ireland. Number of risk groups in Northern Ireland was based on number of different BNF paragraphs prescribed. aHR=adjusted hazard ratio.

COVID-19 outcomes by the end of follow-up of 210 (95% CI 94–326) in the 5–15 years age group, 1544 (1399–1689) in the 16–74 years age group, and 5426 (5340–5512) in the 75 years and older age group, from a total of 40 393 severe COVID-19 outcomes (14 156 in the undervaccinated).

Although COVID-19 vaccine uptake and effectiveness has been studied extensively in the UK and elsewhere,[5–16] fewer studies have looked at the particular association between vaccine deficit and severe COVID-19 outcomes. This study provides precise estimates of the association between full vaccination for the entire population of the UK and incidence of severe COVID-19 outcomes. It also provides information on which population subgroups had the highest levels of undervaccination, which can be used to inform policy, public health, and research agendas.

In the analysis with undervaccination as the outcome, there was some heterogeneity across countries in effect

size estimates for number of risk groups. This could be due to different data availability for QCovid risk groups in each nation, and the fact that BNF chapters were used as a proxy for comorbidities in Northern Ireland. In Scotland, adjusted odds ratios for undervaccination were higher in the 80–84 years and 85 years and older age groups compared with other nations. This could have been due to residual confounding from number of risk groups—Scotland had a higher proportion of individuals in these age groups with no risk groups compared with the other nations.

In the analysis of severe COVID-19 outcomes, being male in the 75 years and older age group was associated with an increase in aHR of severe COVID-19 outcomes relative to being female. However, in the 16–74 years age group this effect was reversed in England, Scotland, and Wales (appendix p 47). This could have been due to residual confounding from age; younger males were more likely to be undervaccinated, and younger age groups were less likely to have an event. Our estimates for the 16–74 years and 75 years and older age groups show that being unvaccinated (strictly maximum dose deficit) was associated with similar or lower hazard ratio for severe COVID-19 outcomes compared with being vaccinated but having a vaccine deficit of at least one dose. This association could be due to vaccine waning and the fact that the most recent dose for those with a vaccine deficit frequently occurred many months before the study start date. The association could also be due to an uncontrolled selection effect for healthier individuals being more likely to be unvaccinated.

Unadjusted severe COVID-19 outcome event rates were not always highest in countries with the highest levels of undervaccination at the start of the study period. This could be because the differences in overall rates of undervaccination were largely driven by younger people who were less susceptible to severe COVID-19 outcomes, and levels of undervaccination across different countries might have converged quickly during the study period. The effect of being undervaccinated on severe COVID-19 outcomes was notably larger than the effect of ethnicity or socioeconomic status. However, although age was most strongly associated with undervaccination, ethnicity and socioeconomic status were also associated with undervaccination. We estimated that the full vaccination counterfactual scenario was associated with an approximately 50% reduction in events among those who were undervaccinated.

A major strength of our study is that, to our knowledge, it is the first epidemiological study carried out using individual-level electronic health records covering the whole population of the UK (aged ≥5 years). Carrying out analyses in parallel across the four nations of the UK and combining them in a meta-analysis allows a consistent UK-wide picture to be developed, as well as providing country-level and regional-level information that can be used to tailor policy interventions. The study represents a notable step towards the goal of real-time pooled and federated health data analytics across the UK.

Our study also has some limitations. A relatively small number of people who were in high-risk categories, as defined by JCVI,[1] were offered more than the standard number of doses for their age group. However, we were not able to reliably identify these individuals in our data, and therefore, we assumed the standard recommended vaccine schedule by age group. We were not able to do subgroup analyses on immunocompromised individuals for the same reason. Ethnicity and QCovid risk groups were also not available in Northern Ireland, and the number of BNF chapters from which the individual had received prescriptions was used as a proxy for the latter. There was a high proportion of unknown values for ethnicity in England, Scotland, and Wales. Effect estimates related to ethnicity were imprecise in Scotland and Wales because non-White ethnicity was a small minority. A small number of QCovid risk groups were not available in each of England, Scotland, and Wales, and individuals with missing values were assigned the category corresponding to absence of the risk factor. Although the estimates from our common and extended adjustment analyses were broadly similar in all nations, there might still have been omitted variable bias and confounding that were not adequately accounted for. There was heterogeneity in the rates of undervaccination as well as the rate of severe COVID-19 events across the four UK nations. There was also some heterogeneity in the results, and heterogeneity statistics were high for some variables in the meta-analysis. This is mainly due to one of the included countries being very large relative to the others. In general, the trends in all four nations were similar, although the heterogeneity could reflect some differences between the UK nations. We did not do a more granular analysis of the effect of vaccination on severe COVID-19 outcomes that took into account different vaccine types, and what the dose number would have been if individuals had followed JCVI guidance; this could be an area for future analysis. We did not have complete vaccination data for those vaccinated outside the UK. We could not accurately identify individuals who were resident in care homes, who might have been less likely to be sent to hospital when ill. Area-based deprivation measures do not fully reflect individual-level measures of socioeconomic standing, such as education, income, and home-ownership. There was some evidence of violation of the proportional hazards assumption in the Cox models, particularly for the time-dependent vaccine deficit variable, which might have been due to waning in vaccine effectiveness.

The meta-analysis was dominated by England because of its size relative to the other countries. However, for this reason there is likely substantially more statistical

heterogeneity within England than any of the other three nations. In future work aiming to study inequalities or regional variations, stratifying analyses by regions within England could be of value.

There is now the need to build on this work by better understanding barriers to vaccination, particularly in the subpopulations identified as less likely to be fully vaccinated and to formulate health policy and public health interventions aiming to improve coverage. This could, for example, include the need to tackle vaccine misinformation in a more direct fashion, and to continue to diversify the use of champions to support public messaging and the range of community-based centres offering vaccinations. There is also an opportunity to build on this unique UK-wide whole-population data to answer other questions relating to endemic infectious diseases and major pressures facing the NHS, including, for example, the post-pandemic waiting list backlog and annual winter pressures, and to create physical infrastructure and data governance frameworks for pooled and federated health data analytics across the UK.

In conclusion, although there are significant challenges to carrying out health data analyses on harmonised datasets across the UK, there are great potential benefits in terms of understanding population health outcomes and designing policy interventions. Our analysis provides numerical estimates of the association between undervaccination and severe COVID-19 outcomes. Our analysis indicates that higher vaccination coverage would have been associated with considerable reduction in severe COVID-19 outcomes, particularly among at-risk subpopulations in the UK.

The HDR UK COALESCE Consortium

Steven Kerr*, Stuart Bedston*, Genevieve Cezard*, Alexia Sampri*, Siobhan Murphy*, Declan T Bradley*, Kirsty Morrison, Ashley Akbari, William Whiteley, Christopher Sullivan, Lynsey Patterson, Kamlesh Khunti, Spiros Denaxas, Thomas Bolton, Samaira Khan, Alan Keys, David Weatherill, Karen Mooney, Jan Davies, Lewis Ritchie, Jim McMenamin, Frank Kee, Angela Wood, Ronan A Lyons, Cathie Sudlow, Chris Robertson*, Aziz Sheikh. *Contributed equally

Contributors

ASh and CSud conceived this study. CR drafted the statistical analysis plan. CR and SK led on the analysis. The analysis in Scotland was carried out by SK. The analysis in Wales was carried out by SB. The analysis in Northern Ireland was carried out by SM, LP, and DTB. The analysis in England was carried out by GC, ASa, AW, and TB. The meta-analysis was carried out by CR. SK drafted the manuscript. CSud is Director of the BHF Data Science Centre and coordinated approvals for and access to data within NHS England's Secure Data Environment service for England for CVD-COVID-UK/COVID-IMPACT. All authors edited later versions of the manuscript. CR, SK, SB, GC, ASa, TB, SM, LP, and DTB accessed and verified the data. All analysts had full access to the data for their respective countries. ASh was responsible for the final decision to submit the manuscript.

Declaration of interests

ASh reports being a member of the Scottish Government Chief Medical Officer's COVID-19 Advisory Group, the Scottish Government's Standing Committee on Pandemic Preparedness, the UK Government's New and Emerging Respiratory Virus Threats Advisory Group Risk Stratification Subgroup, the Department of Health and Social Care's COVID-19 Therapeutics Modelling Group, and AstraZeneca's COVID-19 Strategic Thrombocytopenia Taskforce (all unfunded). CR reports being a member of the Scottish Government Chief Medical Officer's COVID-19 Advisory Group, the Scientific Pandemic Influenza Group on Modelling, and the Medicines and Healthcare Products Regulatory Agency's Vaccine Benefit and Risk Working Group. DTB reports being a member of the Scientific Pandemic Influenza Group on Modelling and the COVID-19 Scientific Advisory Group for Emergencies and its subgroups. KK reports being chair of the ethnicity subgroup of, and a member of, the UK Scientific Advisory Group for Emergencies. RAL reports being a member of the Welsh Government COVID-19 Technical Advisory Group. All other authors declare no competing interests.

Data sharing

The analysis code can be found at https://github.com/HDRUK/COALESCE. The phenotyping and analysis code in England can be found at https://github.com/BHFDSC/CCU051_01. The data that were used in this study are highly sensitive and are not available publicly.

Acknowledgments

We thank Dave Kelly from Albasoft for his support with making primary care data available, and Lynn Morrice, Wendy Inglis Humphrey, and Laura Gonzalez Rienda for their support with project management and administration. This work uses data provided by patients and collected by the NHS as part of their care and support. We would also like to acknowledge all data providers who make anonymised data available for research. We wish to acknowledge the collaborative partnership that enabled acquisition and access to the de-identified data in Wales, which led to this output. The collaboration was led by the Swansea University Health Data Research UK team under the direction of the Welsh Government Technical Advisory Cell and includes the following groups and organisations: the Secure Anonymised Information Linkage Databank, Administrative Data Research Wales, NHS Wales Informatics Service, Public Health Wales, NHS Shared Services Partnership, and the Welsh Ambulance Service Trust. All research conducted has been completed under the permission and approval of the Secure Anonymised Information Linkage Independent Information Governance Review Panel (project number: 0911). We acknowledge the help provided by the staff of the Honest Broker Service (HBS) within the Business Services Organisation Northern Ireland (BSO). The HBS is funded by the BSO and the Department of Health. The authors alone are responsible for the interpretation of the data and any views or opinions presented are solely those of the author and do not necessarily represent those of the BSO. Research in Northern Ireland was conducted under approval by the HBS Governance Board approval number 064. KK is supported by the National Institute for Health Research Applied Research Collaboration East Midlands and Leicester Biomedical Research Centre. This work was supported by the UK Research and Innovation National Core Studies: Data and Connectivity (MC_PC_20029 or MC_PC_20058), with further support from Health Data Research UK (HDR UK), an initiative funded by UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities. This work was also carried out with the support of the BHF Data Science Centre at HDR UK (SP/19/3/34678).This study makes use of de-identified data held in NHS England's Secure Data Environment service for England and made available via the BHF Data Science Centre's CVD-COVID-UK/COVID-IMPACT consortium. The BHF Data Science Centre (grant No SP/19/3/34678, awarded to HDR UK) funded co-development (with NHS England) of the trusted research environment, provision of linked datasets, data access, user software licences, computational usage, and data management and wrangling support, with additional contributions from the HDR UK Data and Connectivity component of the UK Government Chief Scientific Adviser's National Core Studies programme to coordinate national COVID-19 priority research. Consortium partner organisations funded the time of contributing data analysts, biostatisticians, epidemiologists, and clinicians.

References

1 UK Health Security Agency. COVID-19: the green book, chapter 14a. 2022. https://www.gov.uk/government/publications/covid-19-the-green-book-chapter-14a (accessed May 22, 2023).

2 UK Health Security Agency. JCVI advises a spring COVID-19 vaccine dose for the most vulnerable. 2022. https://www.gov.uk/government/news/jcvi-advises-a-spring-covid-19-vaccine-dose-for-the-most-vulnerable (accessed May 22, 2023).

3 UK Department of Health and Social Care. JCVI statement on the COVID-19 booster vaccination programme for autumn 2022: update 15 August 2022. 2022. https://www.gov.uk/government/publications/covid-19-vaccines-for-autumn-2022-jcvi-advice-15-august-2022/jcvi-statement-on-the-covid-19-booster-vaccination-programme-for-autumn-2022-update-15-august-2022 (accessed May 22, 2023).

4 UK Department of Health and Social Care. JCVI statement on spring 2023 COVID-19 vaccinations, 22 February 2023. 2023. https://www.gov.uk/government/publications/spring-2023-covid-19-vaccination-programme-jcvi-advice-22-february-2023/jcvi-statement-on-spring-2023-covid-19-vaccinations-22-february-2023 (accessed May 22, 2023).

5 Magesh S, John D, Li WT, et al. Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a systematic review and meta-analysis. *JAMA Netw Open* 2021; **4:** e2134147.

6 Tessier E, Rai Y, Clarke E, et al. Characteristics associated with COVID-19 vaccine uptake among adults aged 50 years and above in England (8 December 2020–17 May 2021): a population-level observational study. *BMJ Open* 2022; **12:** e055278.

7 Stock SJ, Carruthers J, Calvert C, et al. SARS-CoV-2 infection and COVID-19 vaccination rates in pregnant women in Scotland. *Nat Med* 2022; **28:** 504–12.

8 Polack FP, Thomas SJ, Kitchin N, et al. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. *N Engl J Med* 2020; **383:** 2603–15.

9 Voysey M, Clemens SAC, Madhi SA, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* 2021; **397:** 99–111.

10 Baden LR, El Sahly HM, Essink B, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med* 2021; **384:** 403–16.

11 Vasileiou E, Simpson CR, Shi T, et al. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: a national prospective cohort study. *Lancet* 2021; **397:** 1646–57.

12 Martínez-Baz I, Miqueleiz A, Casado I, et al. Effectiveness of COVID-19 vaccines in preventing SARS-CoV-2 infection and hospitalisation, Navarre, Spain, January to April 2021. *Euro Surveill* 2021; **26:** 2100438.

13 Haas EJ, Anguloa FJ, McLaughlin JM, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. *Lancet* 2021; **397:** 1819–29.

14 Andrews N, Stower J, Kiresbom F, et al. Effectiveness of COVID-19 booster vaccines against COVID-19-related symptoms, hospitalization and death in England. *Nat Med* 2022; **28:** 831–37.

15 Grewal R, Kitchen SA, Nguyen L, et al. Effectiveness of a fourth dose of COVID-19 mRNA vaccine against the omicron variant among long term care residents in Ontario, Canada: test negative design study. *BMJ* 2022; **378:** e071502.

16 UK National Health Service. Master Person Service. https://digital.nhs.uk/services/master-person-service (accessed May 22, 2023).

17 Wood A, Denholm R, Hollings S, et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. *BMJ* 2021; **373:** n826.

18 Lyons RA, Jones KH, John G, et al. The SAIL databank: linking multiple health and social care datasets. *BMC Med Inform Decis Mak* 2009; **9:** 3.

19 Ford DV, Jones KH, Verplancke JP, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. *BMC Health Serv Res* 2009; **9:** 157.

20 Mulholland RH, Vasileiou E, Simpson CR, et al. Cohort profile: Early Pandemic Evaluation and Enhanced Surveillance of COVID-19 (EAVE II) database. *Int J Epidemiol* 2021; **50:** 1064–74.

21 UK Ministry of Housing, Communities and Local Government. English indices of deprivation 2019. 2019. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 (accessed May 22, 2023).

22 Northern Ireland Statistics and Research Agency. Northern Ireland multiple deprivation measure 2017. https://www.nisra.gov.uk/statistics/deprivation/northern-ireland-multiple-deprivation-measure-2017-nimdm2017 (accessed May 22, 2023).

23 Scottish Government. Scottish Index of Multiple Deprivation 2020. 2020. https://www.gov.scot/publications/scottish-index-multiple-deprivation-2020/ (accessed May 22, 2023).

24 Welsh Government. Welsh Index of Multiple Deprivation. 2022. https://gov.wales/welsh-index-multiple-deprivation (accessed May 22, 2023).

25 Clift AK, Coupland CAC, Keogh RH, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *BMJ* 2020; **371:** m3731.

26 National Institute for Health and Care Excellence. British National Formulary. https://bnf.nice.org.uk/ (accessed May 22, 2023).

27 Henderson DAG, Atherton I, McCowan C, Mercer SW, Bailey N. Linkage of national health and social care data: a cross-sectional study of multimorbidity and social care use in people aged over 65 years in Scotland. *Age Ageing 2021;* **50:** 176–82.

28 Strengthening the Reporting of Observational Studies in Epidemiology. https://www.strobe-statement.org (accessed May 22, 2023).

29 Reporting of Studies Using Observational Routinely-Collected Data. https://www.record-statement.org/ (accessed May 22, 2023).