Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Harris RJ, Hall JA, Zaidi A, et al. Effect of vaccination on household transmission of SARS-CoV-2 in England. N Engl J Med. DOI: 10.1056/NEJMc2107717

Contents

Methods

Dataset

The creation of the routine HOSTED dataset has been described in detail elsewhere(1). In brief, laboratory confirmed cases of COVID-19 in England which are reported to national laboratory surveillance systems(2) are linked to individuals who share the same address, using National Health Service (NHS) number and the Unique Property Reference Number (UPRN). Known institutional settings (using UPRN information) such as care homes, prisons, and households with more than 10 residents are excluded.

For this analysis, the HOSTED dataset was linked to data from the National Immunisation Management System (NIMS)(3) to obtain information such as the dates and types of COVID-19 vaccinations for all individuals vaccinated in England.

HOSTED includes individual-level socio-demographic data for cases and household contacts, including age, sex, and Index of Multiple Deprivation (IMD); information on property type, confirmed cases using PCR-based SARS-CoV-2 through national reporting systems, and linked information on hospitalisation and mortality. The vaccination data include date and type (ChAdOx1 nCoV-19 or BNT162b2) of first and second doses for all vaccinated individuals in the HOSTED dataset.

We defined *Index Cases* as the earliest case of laboratory-confirmed COVID-19, by diagnosis date, for a household. *Household Contacts* were defined as all individuals with the same address as the index cases of COVID-19; and *Secondary Cases* as a known household contact of an index case with a positive SARS-CoV-2 test that has a specimen date between two and 14 days after the specimen date of the index case.

The analysis cohort included households with an index case occurring between 4th January 2021 to 28th February 2021, with 14 days observable follow up for all contacts. Households in which *any* individual was vaccinated prior to the 4th January were excluded, so that our analysis would be as broadly generalizable as possible to the overall vaccination campaign. Households in which the index case was vaccinated 1-14 days after testing positive for COVID-19 were also excluded, as were all contacts who had been vaccinated prior to the index case testing positive. We excluded index cases tested under 'pillar 1' of the national testing strategy, which is a proxy for a case being either hospitalised or a health worker. This was because the household contacts of hospitalised cases are likely to have differential exposure profiles compared to contacts of non-hospitalised cases. Finally, we restricted analyses to households with a single index case age 16+, and no co-primary cases (any other cases on the same or next day as the index case).

Statistical analysis

We defined *vaccinated* index cases as having been vaccinated 21 days or more prior to testing positive for COVID-19 based on evidence of the time needed for the vaccine to provide a sufficient level of immunity(4). *Non-vaccinated* index cases were defined as not having received a vaccine prior to testing positive. Households where the index case received the vaccine less than 21 days before testing positive were excluded from this analysis.

We compared household contacts of index cases receiving either the ChAdOx1 nCoV-19 or BNT162b2 vaccines, with contacts of unvaccinated index cases, and the proportion of contacts who

tested positive within 2-14 days of the index case (*secondary cases*) in each group: unvaccinated (base group), BNT162b2, and ChAdOx1 nCoV-19.

The odds of being a secondary case was modelled using logistic regression, adjusted for the following covariates: age of index case and contact (ages grouped as 0-15, 16-29, 30-39, 40-49, 50-59, 60-69, 70-79, and 80+), sex of index case and contact, government office region (nine groups), calendar week of index case, index of multiple deprivation quintile (IMD) and household type. Household types were defined as in previous work(1) and further subdivided by size to give the following groups: pairs/couples of adults both age under 65; pairs/couples of adults with one or both 65+; households with children (<16) of size 2-4, 5-7 and 8-10; multi-generational households (at least one adult age 65+ and age gaps of at least 16 years between this adult and two younger household members) of size 3-4, 5-7 and 8-10; and 3+ adult-only households of size 3-4, 5-7 and 8-10.

We used robust standard errors in the final model to allow for dependence between individuals in the same household. Results are presented in terms of odds ratios (ORs) and 95% confidence intervals (CI). We also considered whether the effect of vaccination varied according to both the age of index case and contact, which was assessed via the inclusion of interaction terms. Models were compared via likelihood ratio tests.

In addition to a binary exposure (vaccinated 21+ days before becoming an index case vs. unvaccinated) we considered households in which the index case was vaccinated any time up to the date of testing positive, in whom there may be a partial effect on transmission. Timing of vaccination was grouped in two-to-three day intervals in the 0-24 days prior to testing positive, then at four, five, or more than six weeks prior to testing positive. The odds of being a secondary case according to vaccination timing (vs. unvaccinated index case) was modelled using the same logistic regression model described above, and we examined the timing according to calendar time of index case (weeks 1-2, 3-4, 5-6 and 7-8 in 2021).

Finally, we compared households in which the index case was vaccinated 21-35 days before testing positive with households in which the index case was vaccinated 1-10 days before testing positive. We used a similar multivariable logistic regression model to that used in the main analysis.

Summary Results

Description of the data set

Data were extracted on 23rd March 2021 and included individuals with at least 14 days of observable follow-up from the date of specimen collection from the index case (i.e. index cases between 4th January 2021 to 28th February 2021, secondary cases up to 14th March 2021). There were 552,984 residential households of two to 10 people where there was at least one case. Table S1 shows the exclusions applied to arrive at the final analysis cohort.

The exclusion criteria did not overlap substantially, although contacts in households where the index case was vaccinated prior to 4th January 2021 were more likely to have been vaccinated (28.6% vs. 3.7%). The final cohort consisted of 365,447 households with a single index case and 1,018,842 contacts. There were 4,107 households where the index case was vaccinated 21 days or more before testing positive (1.12%), and 20,110 where the index case was vaccinated less than 21 days before testing positive (5.51%). Table S2 shows characteristics of contacts according to vaccination status of the index case.

In all households, the majority of index cases and contacts were age <60, with a high proportion of individuals aged <40 in unvaccinated households. Over half of contacts were also aged <40 for all household groups. Male contacts were more common in households in which the index case had been vaccinated. As households where the index case had been vaccinated for 21+ days before being infected required a 3-week window, there are no households in this group with an index case between 4th – 24th January; otherwise, index cases approximately follow the overall case rate, as described above. Distributions of IMD, region and household size were generally similar between the vaccinated and unvaccinated groups.

In households where the index case received the vaccine more than 21 days before testing positive, nearly two-thirds received BNT162b2. Table S3 shows the proportions of each vaccine received by index cases over time: in January, less than 10% of those vaccinated received the ChAdOx1 nCoV-19 vaccine, but this increased to around 40% by mid-February.

Additional statistical analysis

Vaccine effect on transmission

In households where the index case was not vaccinated before testing positive, there were 96,898 secondary cases out of 960,765 household contacts (10.1%). There were 196 secondary cases in 3,424 contacts (5.72%) where the index case received the ChAdOx1 nCoV-19 vaccine 21 days or more before testing positive, and 371 secondary cases in 5,939 contacts (6.25%) where the index case received the BNT162b2 vaccine 21 days or more before testing positive.

The unadjusted odds ratio for being a secondary case if the index case was vaccinated with ChAdOx1 nCoV-19 21 days or more before testing positive (vs. index case not vaccinated) was 0.55 (95% CI 0.46, 0.67), and for BNT162b2, 0.57 (95% CI 0.49, 0.65). Full results from the multivariable logistic regression model are given in table S4. There was little evidence for effect modification by age of index case (likelihood ratio p-value=0.085) or age of contact (p=0.177).

Vaccination timing

Figure S1 shows odds ratios for contacts in which the index case had been vaccinated according to the timing of vaccination, compared to households where the index case was not vaccinated. This analysis also includes households where the index case was vaccinated between 0 and 20 days before testing positive, which due to the timing of the rollout and data under consideration, are a larger group than those vaccinated 21+days before testing positive.

The results show that contacts of vaccinated cases have lower odds of being secondary cases if the index case was vaccinated 14 days or more before testing positive after controlling for calendar week, but this protective effect diminishes sharply if vaccination occurs closer to the positive test date. Of note however is that estimates diverge for the two vaccines: where index cases are recently vaccinated (less than 10 days before testing positive), the odds for contacts being a secondary case are lower for ChAdOx1 nCoV-19, but higher for BNT162b2 (vs. contacts of an unvaccinated index case). The latter may be due to priority administration of BNT162b2 early in the vaccination campaign in those with high-risk social care occupations during a peak incidence period, whose contacts may also have higher risks.

Supplementary Figure S2 show odds ratios of becoming a secondary case according to vaccination timing for different time periods of index case. This suggests that odds of transmission were lower in

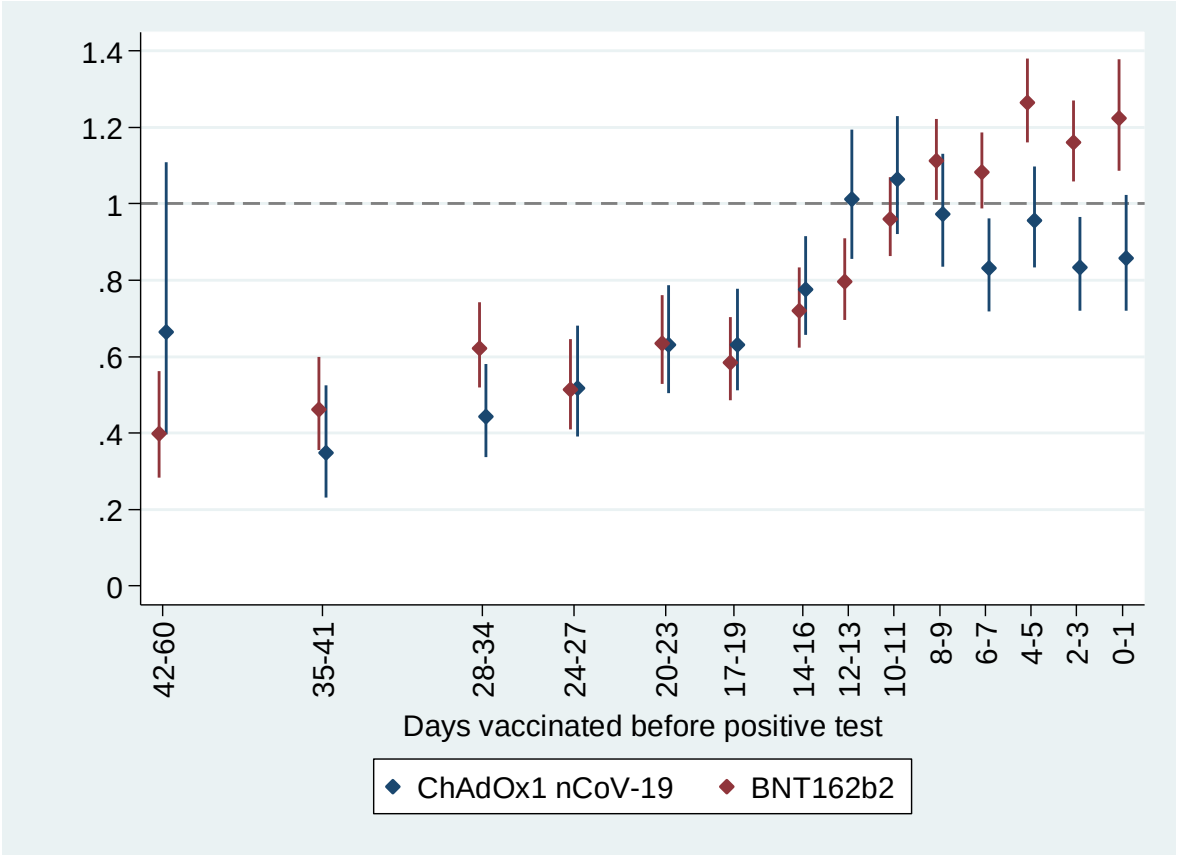
later calendar weeks for index cases receiving either vaccine compared to unvaccinated cases, suggesting an effect of timing as case rates declined. The difference between vaccines in ORs at less than 10 days is more pronounced in weeks 3-4 (and to some degree in weeks 1-2, although data are sparse). Again, this may suggest that those receiving BNT162b2 were targeted for early vaccine due to high exposure risk, whereas the difference is somewhat less pronounced in weeks 5-6 and 7-8.

Estimated effects based on vaccine timing

We compared household contacts of index cases vaccinated 21-35 days before testing positive with household contacts of index cases vaccinated 1-10 days before testing positive (with the same vaccine type). The adjusted ORs from multivariable logistic regression were 0.53 (95% CI: 0.44, 0.63) for ChAdOx1 nCoV-19 and 0.49 (95% CI: 0.44, 0.56) for BNT162b2, indicating a halving in the odds of contacts becoming secondary cases if the index case was vaccinated with either ChAdOx1 nCoV-19 or BNT162b2 21-35 days before testing positive.

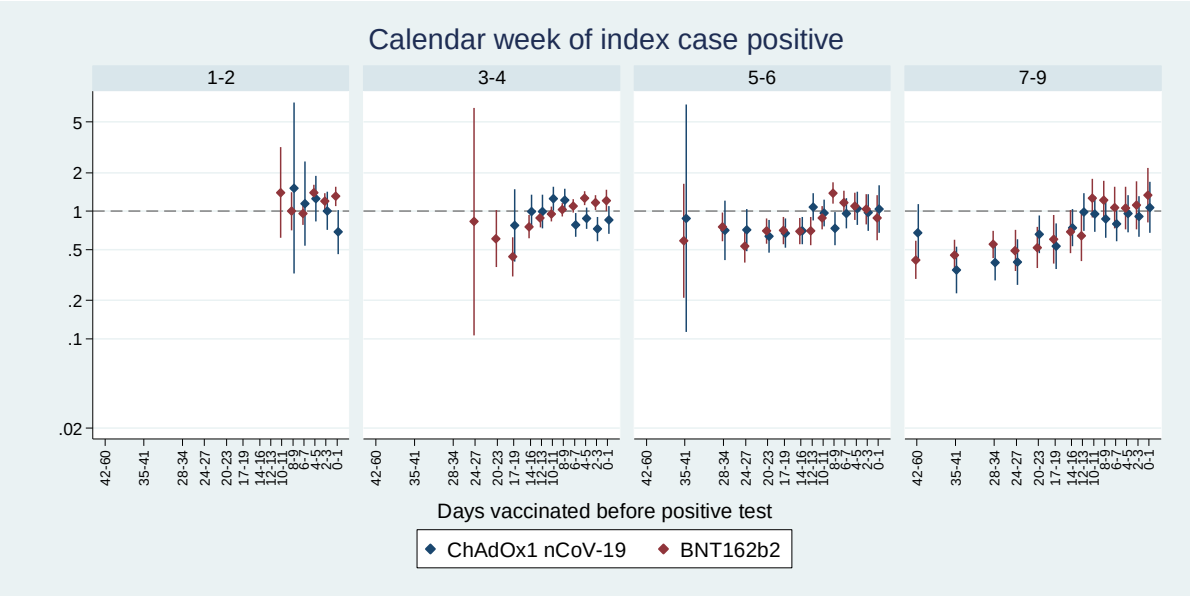
Figures

Figure S1*. Odds ratios for contacts becoming a secondary case according to vaccination timing of the index case (days before testing positive)*



By type of vaccination, vs. contacts where the index case was not vaccinated. Results from multivariable logistic regression.

Figure S2. Odds ratios for contacts becoming a secondary case according to vaccination timing of the index case, by calendar week of index case.



Odds ratios by type of vaccination, vs. contacts where the index case was not vaccinated, as in Figure S1. Results from the multivariable logistic regression.

Tables

Table S1. Numbers of households, contacts of index cases, and secondary cases in the data, numbers excluded according to different criteria, total exclusions, and final analysis cohort.

	Households	Contacts	Secondary cases
Starting cohort	552,984	1,449,427	147,109
Exclusions (overlapping criteria)			
Multiple index/co-primary cases	83,960	146,193	17,944
Index case age under 16	23,560	69,213	8,120
Index case tested in pillar 1	65,270	153,821	12,448
Index case vaccinated prior to 4th January 2021	21,475	54,904	4,991
Index case vaccinated 2-14 days after testing positive	3,849	8,283	516
Contacts vaccinated prior to index case date	59,161	67,622	8,152
Total exclusions (any criteria)	211,276	430,585	44,447
Households where no remaining contacts after exclusions	8,435	0	0
Final cohort	365,447	1,018,842	102,662

Table S2. Characteristics of household contacts according to vaccination status of index case.

	Unvaccinated index case		Index case vaccinated 21+ days before +ive		Index case vaccinated <21 days before +ive	
	N	col %	N	col %	N	col %
Age of index case						
16-29	302,600	31.5%	1,750	18.7%	9,436	19.7%
30-39	259,731	27.0%	2,269	24.2%	11,178	23.3%
40-49	196,898	20.5%	2,215	23.7%	11,277	23.5%
50-59	138,531	14.4%	2,074	22.2%	9,981	20.8%
60-69	50,538	5.3%	743	7.9%	4,157	8.7%
70-79	10,031	1.0%	181	1.9%	1,289	2.7%
80+	2,436	0.3%	131	1.4%	632	1.3%
Age of contact						
0-15	263,190	27.4%	2,969	31.7%	13,429	28.0%
16-29	221,231	23.0%	2,321	24.8%	11,630	24.3%
30-39	141,858	14.8%	1,216	13.0%	6,276	13.1%
40-49	117,352	12.2%	965	10.3%	5,258	11.0%
50-59	126,144	13.1%	1,210	12.9%	6,388	13.3%
60-69	64,166	6.7%	574	6.1%	3,679	7.7%
70-79	20,388	2.1%	77	0.8%	1,020	2.1%
80+	6,436	0.7%	31	0.3%	270	0.6%
Sex						
Male	503,099	52.4%	5,776	61.7%	28,460	59.4%
Female	457,666	47.6%	3,587	38.3%	19,490	40.6%
Index case date						
04 Jan - 10 Jan	293,672	30.6%	0	0.0%	1,618	3.4%
11 Jan - 17 Jan	219,283	22.8%	0	0.0%	8,242	17.2%
18 Jan - 24 Jan	151,170	15.7%	0	0.0%	12,180	25.4%
25 Jan - 31 Jan	101,926	10.6%	334	3.6%	10,182	21.2%
01 Feb - 07 Feb	73,531	7.7%	1,302	13.9%	6,908	14.4%
08 Feb - 14 Feb	49,759	5.2%	2,117	22.6%	3,887	8.1%
15 Feb - 21 Feb	42,509	4.4%	3,004	32.1%	2,786	5.8%
22 Feb - 28 Feb	28,915	3.0%	2,606	27.8%	2,147	4.5%
IMD quintile						
1	265,475	27.6%	2,486	26.6%	11,070	23.1%
2	239,131	24.9%	2,065	22.1%	10,850	22.6%
3	184,066	19.2%	1,930	20.6%	10,099	21.1%
4	149,319	15.5%	1,556	16.6%	8,720	18.2%
5	122,774	12.8%	1,326	14.2%	7,211	15.0%
Region						
East Midlands	79,301	8.3%	1,310	14.0%	5,180	10.8%
East of England	102,804	10.7%	931	9.9%	5,973	12.5%
London	229,362	23.9%	571	6.1%	5,442	11.3%
North East	32,640	3.4%	390	4.2%	1,795	3.7%
North West	132,466	13.8%	1,730	18.5%	6,991	14.6%
South East	139,455	14.5%	1,156	12.3%	8,440	17.6%
South West	58,015	6.0%	811	8.7%	4,433	9.2%

	Unvaccinated index case		Index case vaccinated 21+ days before +ive		Index case vaccinated <21 days before +ive	
	N	col %	N	col %	N	col %
West Midlands	122,359	12.7%	1,594	17.0%	6,453	13.5%
Yorkshire and The Humber	64,363	6.7%	870	9.3%	3,243	6.8%
Household type						
Adult pair/couple	51,802	5.4%	733	7.8%	3,547	7.4%
Older pair/couple	14,006	1.5%	143	1.5%	1,260	2.6%
HH with children, 2-4	200,338	20.9%	2,469	26.4%	11,322	23.6%
HH with children, 5-7	212,808	22.1%	1,931	20.6%	8,952	18.7%
HH with children, 8-10	57,688	6.0%	314	3.4%	1,449	3.0%
Multi-gen HH, 3-4	9,876	1.0%	142	1.5%	701	1.5%
Multi-gen HH, 5-7	53,814	5.6%	496	5.3%	2,728	5.7%
Multi-gen HH, 8-10	31,917	3.3%	139	1.5%	1,137	2.4%
Adult only, 3-4	208,556	21.7%	2,264	24.2%	12,093	25.2%
Adult only, 5-7	98,479	10.3%	632	6.7%	4,158	8.7%
Adult only, 8-10	21,481	2.2%	100	1.1%	603	1.3%

Table S3. Numbers and proportions of index cases that received each vaccine type 21 days or more prior to testing positive, by week of index case date.

Week	ChAdOx1 nCoV-19	BNT162b2
25 Jan - 31 Jan	9 (6.7%)	125 (93.3%)
01 Feb - 07 Feb	136 (23.4%)	445 (76.6%)
08 Feb - 14 Feb	326 (35.4%)	596 (64.6%)
15 Feb - 21 Feb	512 (39.0%)	802 (61.0%)
22 Feb - 28 Feb	481 (41.6%)	675 (58.4%)
Total	1464 (35.6%)	2643 (64.4%)

Table S4. Numbers and proportions of secondary cases and contacts, and odds ratios from multivariable logistic regression for household contact becoming a secondary case.

	Secondary cases/contacts	% secondary cases	Adjusted OR (95% CI*)
Age of index case			
16-29	22585/313920	7.2%	1 (base)
30-39	24179/273347	8.8%	1.33 (1.30, 1.37)
40-49	24565/210575	11.7%	1.72 (1.68, 1.77)
50-59	21055/150775	14.0%	1.98 (1.93, 2.04)
60-69	8146/55486	14.7%	2.09 (2.01, 2.16)
70-79	1677/11526	14.5%	2.00 (1.86, 2.14)
80+	455/3213	14.2%	2.00 (1.75, 2.28)
Age of contact			
0-15	20140/279815	7.2%	1 (base)
16-29	22187/235378	9.4%	1.67 (1.63, 1.71)
30-39	14804/149442	9.9%	1.76 (1.72, 1.81)
40-49	15312/123655	12.4%	2.13 (2.08, 2.18)
50-59	18036/133848	13.5%	2.45 (2.38, 2.51)
60-69	8904/68472	13.0%	2.21 (2.14, 2.29)
70-79	2563/21492	11.9%	1.82 (1.73, 1.92)
80+	716/6740	10.6%	1.67 (1.54, 1.82)
Sex of index case			
Male	49363/472744	10.4%	1 (base)
Female	53299/546098	9.8%	0.99 (0.97, 1.01)
Sex of contact			
Male	49068/537794	9.1%	1 (base)
Female	53594/481048	11.1%	1.27 (1.25, 1.28)
Index case date			
04 Jan - 10 Jan	31588/295290	10.7%	1 (base)
11 Jan - 17 Jan	23651/227525	10.4%	0.93 (0.91, 0.96)
18 Jan - 24 Jan	16052/163350	9.8%	0.87 (0.85, 0.89)
25 Jan - 31 Jan	10639/112615	9.4%	0.82 (0.80, 0.85)
01 Feb - 07 Feb	7828/81961	9.6%	0.85 (0.82, 0.88)
08 Feb - 14 Feb	5574/55952	10.0%	0.90 (0.86, 0.93)
15 Feb - 21 Feb	4482/48406	9.3%	0.83 (0.79, 0.87)
22 Feb - 28 Feb	2848/33743	8.4%	0.76 (0.72, 0.81)
IMD quintile			
1 (most deprived)	25422/279229	9.1%	1 (base)
2	23330/252232	9.2%	1.06 (1.03, 1.09)
3	20108/196272	10.2%	1.12 (1.09, 1.15)
4	18295/159688	11.5%	1.19 (1.16, 1.22)
5 (least deprived)	15507/131421	11.8%	1.18 (1.15, 1.22)
Region			
East Midlands	9324/85888	10.9%	1.09 (1.05, 1.13)
East of England	11864/109789	10.8%	1.04 (1.01, 1.08)
London	17744/235439	7.5%	0.78 (0.76, 0.80)
North East	4317/34860	12.4%	1.24 (1.18, 1.30)
North West	15900/141326	11.3%	1.13 (1.10, 1.17)
South East	15618/149158	10.5%	1 (base)
South West	6307/63333	10.0%	0.95 (0.91, 0.99)
West Midlands	14012/130523	10.7%	1.14 (1.10, 1.18)

	Secondary cases/contacts	% secondary cases	Adjusted OR (95% CI*)
Yorkshire and The Humber	7576/68526	11.1%	1.14 (1.09, 1.18)
Household type/size			
Adult pair/couple	9160/56138	16.3%	1 (base)
Older pair/couple	3028/15422	19.6%	1.04 (0.98, 1.10)
HH with children, 2-4	26766/214337	12.5%	1.21 (1.17, 1.25)
HH with children, 5-7	19144/223822	8.6%	0.79 (0.76, 0.81)
HH with children, 8-10	3090/59469	5.2%	0.48 (0.45, 0.51)
Multi-gen HH, 3-4	1114/10729	10.4%	0.79 (0.73, 0.86)
Multi-gen HH, 5-7	5194/57094	9.1%	0.71 (0.68, 0.75)
Multi-gen HH, 8-10	2221/33200	6.7%	0.54 (0.51, 0.58)
Adult only, 3-4	24771/223092	11.1%	0.72 (0.70, 0.74)
Adult only, 5-7	7468/103333	7.2%	0.53 (0.51, 0.55)
Adult only, 8-10	706/22206	3.2%	0.25 (0.22, 0.28)
Vaccination of index case			
Not vaccinated	96898/960765	10.1%	1 (base)
ChAdOx1 nCoV-19	196/3424	5.7%	0.52 (0.43, 0.62)
BNT162b2	371/5939	6.2%	0.54 (0.47, 0.62)

*Confidence intervals based on robust standard errors accounting for household clustering.

Acknowledgements, ethics and funding

Acknowledgements:

The authors wish to thank Heather Pinches, Phillip Bowker and Paul Ellingham as well as the rest of the NHS Digital team whose hard work and dedication enabled the creation of these data during the UK Covid19 pandemic response. The authors also wish to thank Sarah Woodhall for her integral role in developing the HOSTED project; Jamie Lopez-Bernal, Mary Ramsay, Liz Miller and the Vaccine Efficacy Working Group for their input and discussion on this project; and Jemma Walker for comments on an early version of the pre-print relating to this manuscript.

Ethics approval

The HOSTED surveillance system was reviewed and approved by the PHE Research Ethics Governance Group. The data was collected and linked by NHS Digital. The data was processed lawfully under GDPR Article 6(1)e and 9(2)i and shared under Regulation 3(4) of the Health Service (Control of Patient Information) Regulations 2002.

Funding

This work was undertaken as part of the core functions of Public Health England in relation to the surveillance of communicable diseases and outbreak response. Funding: None

References

1. Hall JA, Harris RJ, Zaidi A, Woodhall SC, Dabrera G, Dunbar JK. HOSTED—England's Household Transmission Evaluation Dataset: preliminary findings from a novel passive surveillance system of COVID-19. Int J Epidemiol. 2021 Apr;dyab057. Available from: https://doi.org/10.1093/ije/dyab057

2. Clare T, Twohig KA, O'Connell A-M, Dabrera G. Timeliness and completeness of laboratory-based surveillance of COVID-19 cases in England. Public Health. 2021 Apr;194:163–6. Available from: https://www.sciencedirect.com/science/article/pii/S0033350621001219

3. NHS Digital. The National Immunisation Management Service. Apr, 2021. Available from: https://www.england.nhs.uk/contact-us/privacy-notice/national-flu-vaccination-programme/

4. Hall VJ, Foulkes S, Saei A, Andrews N, Oguti B, Charlett A, et al. Effectiveness of BNT162b2 mRNA Vaccine Against Infection and COVID-19 Vaccine Coverage in Healthcare Workers in England, Multicentre Prospective Cohort Study (the SIREN Study). SSRN Electron J. 2021; Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3790399