Supplemental Material*

Hulme WJ, Williamson E, Horne EMF, et al. Challenges in estimating the effectiveness of COVID-19 vaccination using observational data. Ann Intern Med. 2 May 2023. [Epub ahead of print]. doi:10.7326/M21-4269

Contents

* This supplemental material was provided by the authors to give readers further details on their article. The material was not copyedited.

Additional Details of Analyses

OpenSAFELY

OpenSAFELY is a data analytics platform created on behalf of NHS England to address urgent COVID-19 research questions (https://opensafely.org). It provides a secure software interface allowing analysis of pseudonymized primary care patient records from England in near real-time within the EHR vendor's highly secure data centre. This, with other technical and organisational controls, minimizes the risk of re-identification. Primary care data were linked (via patients' NHS numbers) to hospital records via NHS Digital's Hospital Episode Statistics (HES), national coronavirus testing records via the Second Generation Surveillance System (SGSS), and national death registry records from the Office for National Statistics (ONS).

Exclusion criteria

People were excluded if:

1. They had unreliable vaccination data (vaccinated before eligible or unknown vaccine brand);
2. They had less than one year of continuous GP registration up to the study start date (to ensure reasonably complete data on covariates);
3. They were a health or social care worker (as vaccination was coordinated separately in these groups: this affects very few because eligible persons were aged 70 years or over);
4. They were known to be care- or nursing-home residents or medically housebound (because the vaccine rollout was organised separately in these groups);
5. They were on end-of-life care pathways on or before the study start date (as vaccination was expected to be unlikely in such people);
6. Their sex, ethnicity, deprivation, or geographical region was unknown;
7. They had evidence of SARS-CoV-2 infection (either a positive polymerase chain reaction (PCR) or lateral flow test or probable COVID-19 recorded in primary care records) before the start of the rollout in their eligibility group (as vaccine efficacy is likely to be attenuated in such people).

Baseline confounders

All baseline confounders except age were derived at each trial start date for the sequential trials approach and on 8 December 2020 for the single trial approach. Baseline confounders in the single trial approach were: age and age-squared; sex; ethnicity (Black, Mixed, South Asian, White, Other, as per the UK census); English Index of Multiple Deprivation (IMD; grouped by quintile); NHS region; severe obesity (Body Mass Index $<40 kg/m^2$ or not recorded, $\geq 40 kg/m^2$); chronic heart disease; chronic kidney disease; diabetes; chronic liver disease; chronic respiratory disease; immunosuppressed (chemo- or radio-therapy, solid organ transplantation, permanent or temporary immunosuppression, asplenia, or disease-modifying anti-rheumatic drugs); chronic neurological disease; learning disabilities, including Down's syndrome; serious mental illness (psychosis, schizophrenia or bipolar disorder); multi-morbidity (0, 1 or 2+ comorbid conditions in different organ systems); influenza vaccination within the last 5 years; UK government COVID-19 shielding criteria met. Calendar time in the single trial approach was modelled using region-specific restricted cubic splines, with knots at the 1^{st}, 2^{nd}, and 3^{rd} quartile plus 2 boundary knots.

Baseline confounders in the sequential trial approach were the same as for the single trial approach, but with the following matching variables omitted as they were balanced between groups: age, sex, NHS region, UK government COVID-19 shielding criteria met. Calendar time was not modelled in the sequential trial approach.

Outcomes
Positive SARS-CoV-2 swab test results were obtained via SGSS and based on swab date (we did not distinguish between symptomatic and asymptomatic infection, nor between PCR and lateral flow testing). COVID-19 hospital admissions were obtained via HES in-patient records with ICD-10 COVID-19 diagnosis codes (both primary and non-primary diagnosis codes). All-cause mortality was obtained via linked death registry records.

Missing data
After exclusions for missing sex, ethnicity, deprivation, or geographical region, there were no missing values because all other variables were defined by presence or absence of clinical codes or events.

Sequential trial emulation
In the sequential trial approach, already matched people were not selected in trials starting on subsequent days. Excluding people already selected in a previous trial is not strictly necessary, but it simplifies the statistical analysis because it avoids having multiple copies of data for the same person in the unvaccinated group, which violates the assumption of statistical independence between each observational unit. When, as in our study, the number of individuals and outcome events is large, there is little advantage of requiring the use of more complex standard error estimates that including multiple individuals would entail.

Single trial emulation using marginal structural models
Marginal structural models (MSMs) enable estimation of causal treatment effects in observational data where time-varying confounding of the treatment-outcome relationship may occur. In summary, the substantive regression model is weighted so that the probability of treatment at each unit of follow-up time for each individual is unrelated to their observed time-dependent covariates. The weights are derived from auxiliary models that estimate the time-dependent probability of vaccination for each individual.

Fitting MSMs was a multi-step process. First, the data were reshaped as one-row-per-person-per-day of follow-up time, to encode the time-varying variables and estimate the probability of vaccination for each day of follow-up. Second, the vaccination probabilities were estimated using pooled logistic regression. Third, the substantive model was fitted using pooled logistic regression, weighted by the inverse of the probability of each person's vaccination history on each day, and with robust standard errors allowing for clustering on individuals that is induced by the weighting. This method is referred to as inverse probability weighting (IPW). The estimated vaccination-outcome odds ratio from this approach is a good approximation of the hazard ratio assuming the risk of the outcome is low on each day of follow-up. The steps are outlined in more detail below.

First, a one-row-per-person-per-day dataset was created, from the study start date (8 December 2021) until the study end date or an earlier censoring date. Time-varying variables (vaccination status, covariates, outcomes) were updated each day, with value changes (for instance vaccination status) assumed to occur at the end of each day. For example, if a person receives their first vaccine dose on day 3 they are considered to be at risk of vaccination on days 1, 2, and 3, and no longer at risk from day 4 onwards.

Second, models to estimate the probability of vaccination were fitted (henceforth the IPW models). Two models are necessary for stabilised weights: a full model including both time-dependent and time-independent covariates; and a reduced model excluding time-dependent covariates except variables related to time itself (for instance calendar time). The weight at time t for person i derived from each IPW model is based on the probability of their vaccination history, conditional on their time-dependent covariates, i.e.,

$$w_{it} = \prod_{t=1}^{t} P(\text{vaccination status at time } t \mid \text{vaccination status at } t-1, \text{person characteristics at time } t)$$

The stabilised weight for each day of follow-up for each person is the ratio of the reduced model against the full model $\frac{{w_{it}}^{fxd}}{w_{it}}$.

Note that $P(\text{person } i \text{ vaccinated at time } t \mid \text{vaccinated at } t-1, \text{person characteristics at time } t) = 1$.

Finally, a logistic regression model is fit on the one-row-per-person-per-day data, weighted by the stabilized weight $\frac{{w_{it}}^{fxd}}{w_{it}}$, and adjusting for time-independent covariates. The vaccinated/unvaccinated odds-ratio is a good approximation for the hazard ratio if the event risk on each day t is small. Cluster-robust standard errors were used to account for person-level clustering induced by the weighting.

Data sharing and reproducibility

Analyses were conducted in Python 3.8 and R version 4.0.2. All analysis code is available for review and reuse at https://github.com/opensafely/covid-vaccine-effectiveness-sequential-vs-single. The downloaded repository is provided as supplementary material.

Supplement Table 1. Characteristics of People Included in the Single Trial and Sequential Trial Analyses

Data in each cell are the number of people (percent of population) unless otherwise stated.

Characteristic	**Level**	**Single trial cohort***	**Sequential trial arms****			
			BNT162b2		**ChAdOx1**	
			Vaccinated	**Unvaccinated**	**Vaccinated**	**Unvaccinated**
Total N		2780931	1245267	1245267	932901	932901
Age		76 (73, 82)	79 (74, 83)	78 (74, 84)	75 (73, 78)	75 (72, 78)
JCVI age band	70-74	1094373 (39.4%)	366501 (29.4%)	366501 (29.4%)	482715 (51.7%)	482715 (51.7%)
	75-79	756885 (27.2%)	326565 (26.2%)	326565 (26.2%)	297525 (31.9%)	297525 (31.9%)
	80+	929673 (33.4%)	552201 (44.3%)	552201 (44.3%)	152661 (16.4%)	152661 (16.4%)
Sex	Female	1501071 (54.0%)	679143 (54.5%)	679143 (54.5%)	501273 (53.7%)	501273 (53.7%)
	Male	1279863 (46.0%)	566121 (45.5%)	566121 (45.5%)	431625 (46.3%)	431625 (46.3%)
Ethnicity	White	2660415 (95.7%)	1201641 (96.5%)	1182465 (95.0%)	898617 (96.3%)	883311 (94.7%)
	Black	22653 (0.8%)	6801 (0.5%)	13623 (1.1%)	5625 (0.6%)	10965 (1.2%)
	South Asian	69393 (2.5%)	26769 (2.1%)	35007 (2.8%)	20175 (2.2%)	27063 (2.9%)
	Mixed	8661 (0.3%)	3153 (0.2%)	4467 (0.4%)	2481 (0.3%)	3591 (0.4%)
	Other	19815 (0.7%)	6903 (0.5%)	9705 (0.8%)	5997 (0.6%)	7959 (0.9%)
Deprivation	1 (most deprived)	355413 (12.8%)	145035 (11.7%)	169929 (13.7%)	113133 (12.1%)	136581 (14.6%)
	2	469863 (16.9%)	204795 (16.4%)	220449 (17.7%)	151347 (16.2%)	169029 (18.1%)
	3	639765 (23.0%)	284205 (22.8%)	292797 (23.5%)	214053 (22.9%)	218499 (23.4%)
	4	658845 (23.7%)	302475 (24.3%)	285339 (22.9%)	223119 (23.9%)	209985 (22.5%)
	5 (least deprived)	657045 (23.6%)	308757 (24.8%)	276747 (22.2%)	231249 (24.8%)	198807 (21.3%)

Region	Midlands	583359 (21.0%)	262833 (21.1%)	262833 (21.1%)	197547 (21.2%)	197547 (21.2%)
	North East and Yorkshire	498747 (17.9%)	213939 (17.2%)	213939 (17.2%)	171879 (18.4%)	171879 (18.4%)
	South East	204417 (7.3%)	89721 (7.2%)	89721 (7.2%)	71283 (7.6%)	71283 (7.6%)
	London	95019 (3.4%)	46641 (3.8%)	46641 (3.8%)	26637 (2.9%)	26637 (2.9%)
	East of England	659475 (23.7%)	307017 (24.6%)	307017 (24.6%)	210741 (22.6%)	210741 (22.6%)
	North West	256077 (9.2%)	107433 (8.6%)	107433 (8.6%)	94473 (10.1%)	94473 (10.1%)
	South West	483825 (17.4%)	217683 (17.5%)	217683 (17.5%)	160341 (17.2%)	160341 (17.2%)
Clinically vulnerable	Not clinically at-risk	1081155 (38.9%)	453879 (36.4%)	453879 (36.4%)	396771 (42.5%)	396771 (42.5%)
	Clinically at-risk	1378839 (49.6%)	641385 (51.5%)	641385 (51.5%)	435387 (46.7%)	435387 (46.7%)
	Clinically extremely vulnerable	320937 (11.5%)	150003 (12.0%)	150003 (12.0%)	100743 (10.8%)	100743 (10.8%)
Body Mass Index > 40 kg/m^2		60783 (2.2%)	23133 (1.9%)	27003 (2.2%)	22083 (2.4%)	25401 (2.7%)
Chronic heart disease		883857 (31.8%)	421947 (33.9%)	418731 (33.6%)	269241 (28.9%)	264147 (28.3%)
Chronic kidney disease		528063 (19.0%)	265629 (21.3%)	266073 (21.4%)	147657 (15.8%)	146661 (15.7%)
Diabetes		503247 (18.1%)	227217 (18.2%)	239523 (19.2%)	161901 (17.3%)	169371 (18.2%)
Chronic liver disease		68721 (2.5%)	29463 (2.4%)	30657 (2.5%)	24681 (2.6%)	25749 (2.8%)
Chronic respiratory disease		292083 (10.5%)	132981 (10.7%)	133041 (10.7%)	93255 (10.0%)	93957 (10.1%)
Chronic neurological disease		320649 (11.5%)	148173 (11.9%)	158313 (12.7%)	96069 (10.3%)	99945 (10.7%)

Morbidity count	0	1146813 (41.2%)	483513 (38.8%)	482169 (38.7%)	419139 (44.9%)	418611 (44.9%)
	1	904317 (32.5%)	413517 (33.2%)	404193 (32.5%)	297399 (31.9%)	293271 (31.4%)
	2+	729801 (26.2%)	348231 (28.0%)	358905 (28.8%)	216363 (23.2%)	221013 (23.7%)
Immunosuppressed		133599 (4.8%)	62433 (5.0%)	59415 (4.8%)	44259 (4.7%)	42057 (4.5%)
Learning disabilities		2745 (0.1%)	711 (0.1%)	1515 (0.1%)	861 (0.1%)	1539 (0.2%)
Serious mental illness		17523 (0.6%)	6099 (0.5%)	8643 (0.7%)	5697 (0.6%)	7611 (0.8%)
Influenza vaccination in previous 5 years		2463543 (88.6%)	1150845 (92.4%)	1065375 (85.6%)	844113 (90.5%)	768093 (82.3%)

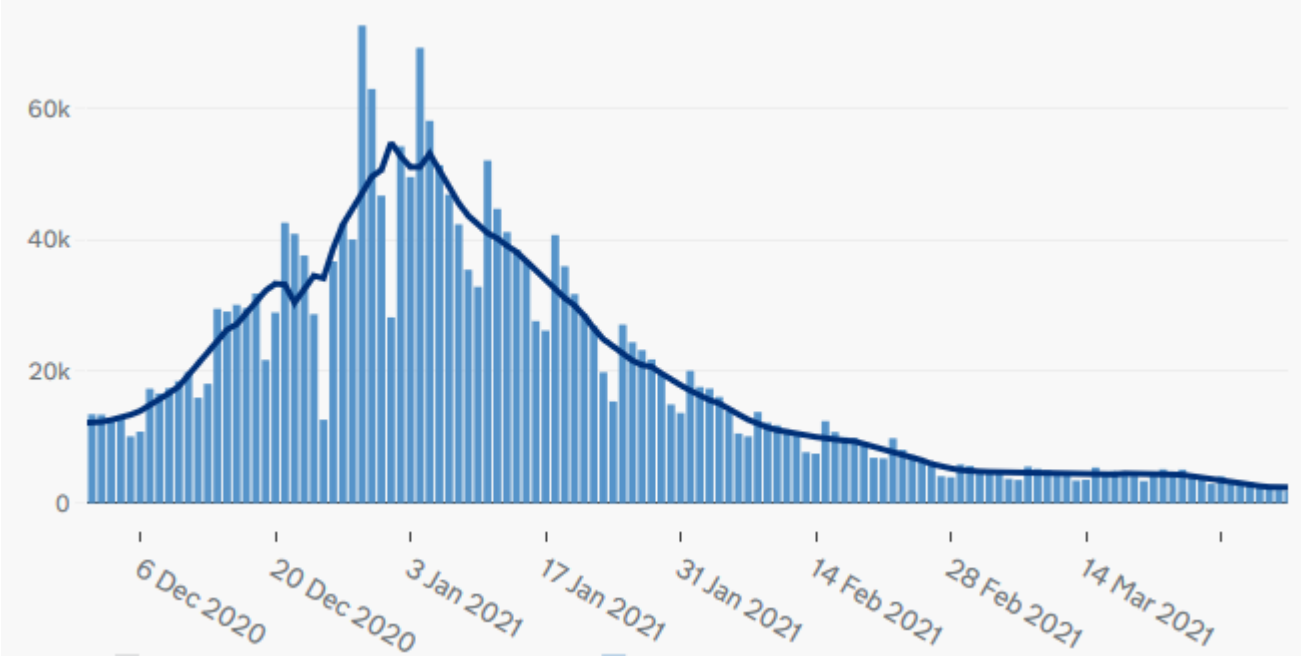
*The single trial cohort cannot be split into BNT162b2/ChAdOx1/unvaccinated arms as vaccination is a time-varying variable. Characteristics for the single trial cohort were defined on 8 December 2020.

**The same individuals may be present in the unvaccinated arms of the BNT162b2 and ChAdOx1 trials. Characteristics for the sequential trial arms were defined on the trial start date.

Supplement Table 2. Estimated Hazard Ratios for the Three Outcomes, Comparing First COVID-19 Vaccination With No Vaccination Over Time Periods Since Vaccination, for the Single Trial and Sequential Trial Approaches

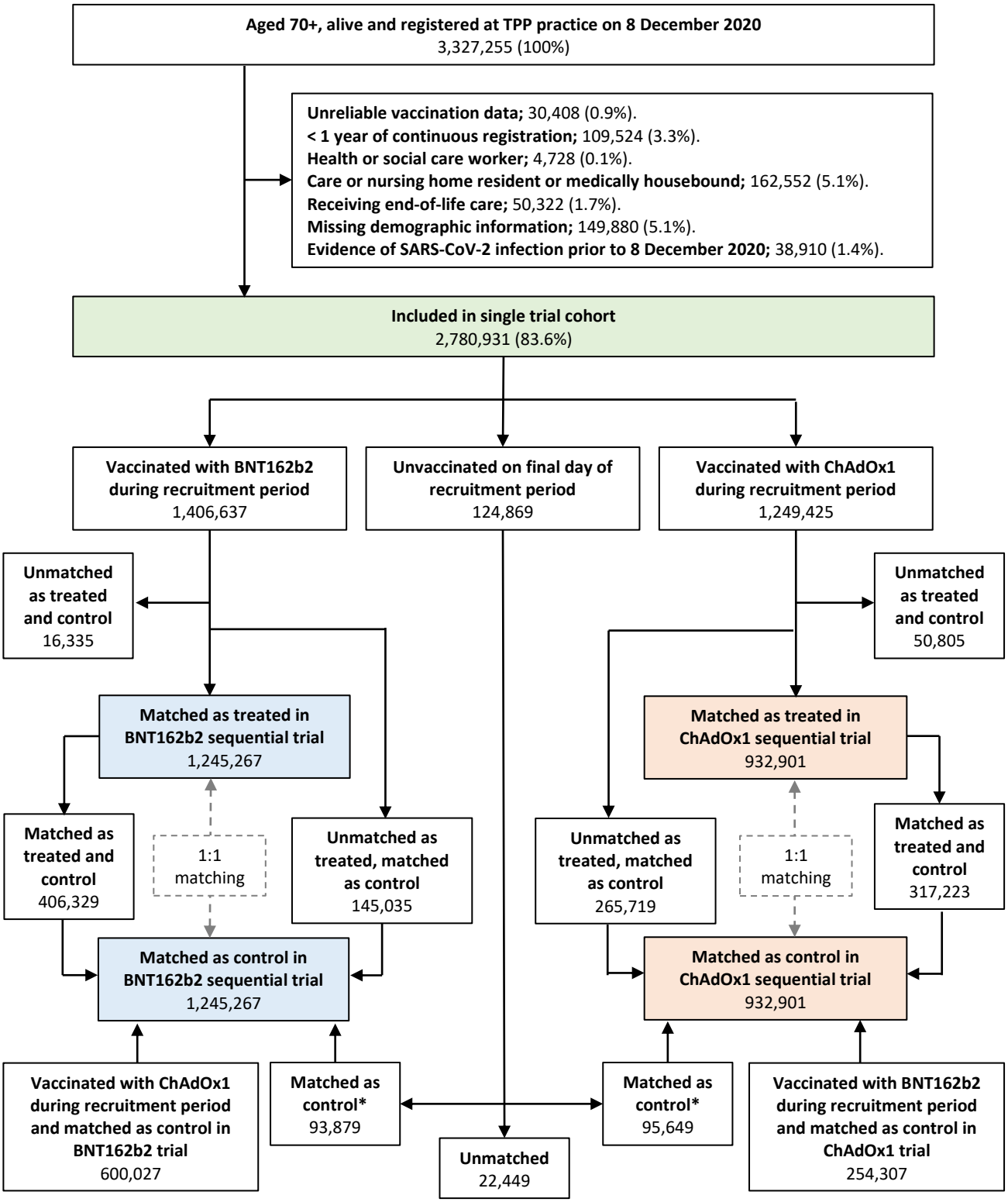
	Time (days) since first dose	Hazard ratio (95% CI) Sequential trial	Single trial
Positive SARS-CoV-2 test			
BNT162b2	1-3	0.19 (0.16, 0.22)	0.27 (0.24, 0.30)
	4-7	0.37 (0.32, 0.41)	0.42 (0.38, 0.46)
	8-14	0.50 (0.45, 0.54)	0.56 (0.52, 0.60)
	15-21	0.40 (0.36, 0.45)	0.49 (0.45, 0.54)
	22-28	0.31 (0.26, 0.36)	0.46 (0.41, 0.51)
	29-35	0.20 (0.16, 0.25)	0.44 (0.39, 0.50)
	36-70	0.14 (0.11, 0.17)	0.43 (0.37, 0.49)
ChAdOx1	1-3	0.25 (0.21, 0.30)	0.44 (0.38, 0.50)
	4-7	0.46 (0.39, 0.53)	0.66 (0.59, 0.74)
	8-14	0.50 (0.43, 0.58)	0.83 (0.75, 0.92)
	15-21	0.37 (0.30, 0.47)	0.79 (0.70, 0.90)
	22-28	0.28 (0.21, 0.39)	0.68 (0.58, 0.80)
	29-35	0.30 (0.20, 0.45)	0.61 (0.51, 0.74)
	36-70	0.29 (0.22, 0.39)	0.59 (0.49, 0.71)
COVID-19 Hospitalisation			
BNT162b2	1-3	0.19 (0.15, 0.24)	0.26 (0.22, 0.31)
	4-7	0.30 (0.26, 0.36)	0.35 (0.31, 0.40)
	8-14	0.41 (0.36, 0.46)	0.47 (0.42, 0.51)
	15-21	0.34 (0.29, 0.39)	0.36 (0.32, 0.40)
	22-28	0.28 (0.23, 0.35)	0.29 (0.25, 0.33)
	29-35	0.22 (0.16, 0.29)	0.26 (0.22, 0.31)
	36-70	0.10 (0.07, 0.14)	0.24 (0.20, 0.28)
ChAdOx1	1-3	0.30 (0.23, 0.39)	0.43 (0.37, 0.51)
	4-7	0.36 (0.29, 0.45)	0.41 (0.36, 0.48)
	8-14	0.33 (0.27, 0.41)	0.51 (0.44, 0.58)
	15-21	0.26 (0.19, 0.35)	0.47 (0.40, 0.55)
	22-28	0.16 (0.10, 0.25)	0.38 (0.31, 0.47)
	29-35	0.20 (0.11, 0.34)	0.33 (0.26, 0.41)
	36-70	0.28 (0.20, 0.41)	0.29 (0.23, 0.37)
Death			
BNT162b2	1-3	0.19 (0.13, 0.28)	0.10 (0.08, 0.12)
	4-7	0.23 (0.18, 0.30)	0.15 (0.13, 0.18)
	8-14	0.25 (0.21, 0.29)	0.19 (0.17, 0.21)
	15-21	0.21 (0.18, 0.25)	0.22 (0.20, 0.25)
	22-28	0.29 (0.25, 0.35)	0.25 (0.22, 0.28)
	29-35	0.19 (0.16, 0.24)	0.21 (0.19, 0.25)
	36-70	0.12 (0.10, 0.14)	0.23 (0.20, 0.27)
ChAdOx1	1-3	0.26 (0.17, 0.38)	0.16 (0.13, 0.20)
	4-7	0.24 (0.17, 0.33)	0.23 (0.20, 0.27)
	8-14	0.28 (0.22, 0.34)	0.29 (0.26, 0.33)
	15-21	0.27 (0.21, 0.33)	0.38 (0.33, 0.43)
	22-28	0.16 (0.12, 0.21)	0.33 (0.28, 0.38)
	29-35	0.20 (0.15, 0.27)	0.36 (0.31, 0.43)
	36-70	0.28 (0.24, 0.33)	0.38 (0.32, 0.46)

Supplement Figure 1. Number of positive tests for SARS-CoV-2 in the United Kingdom between December 2020 and April 2021.



Source: https://coronavirus.data.gov.uk/details/cases?areaType=nation&areaName=England

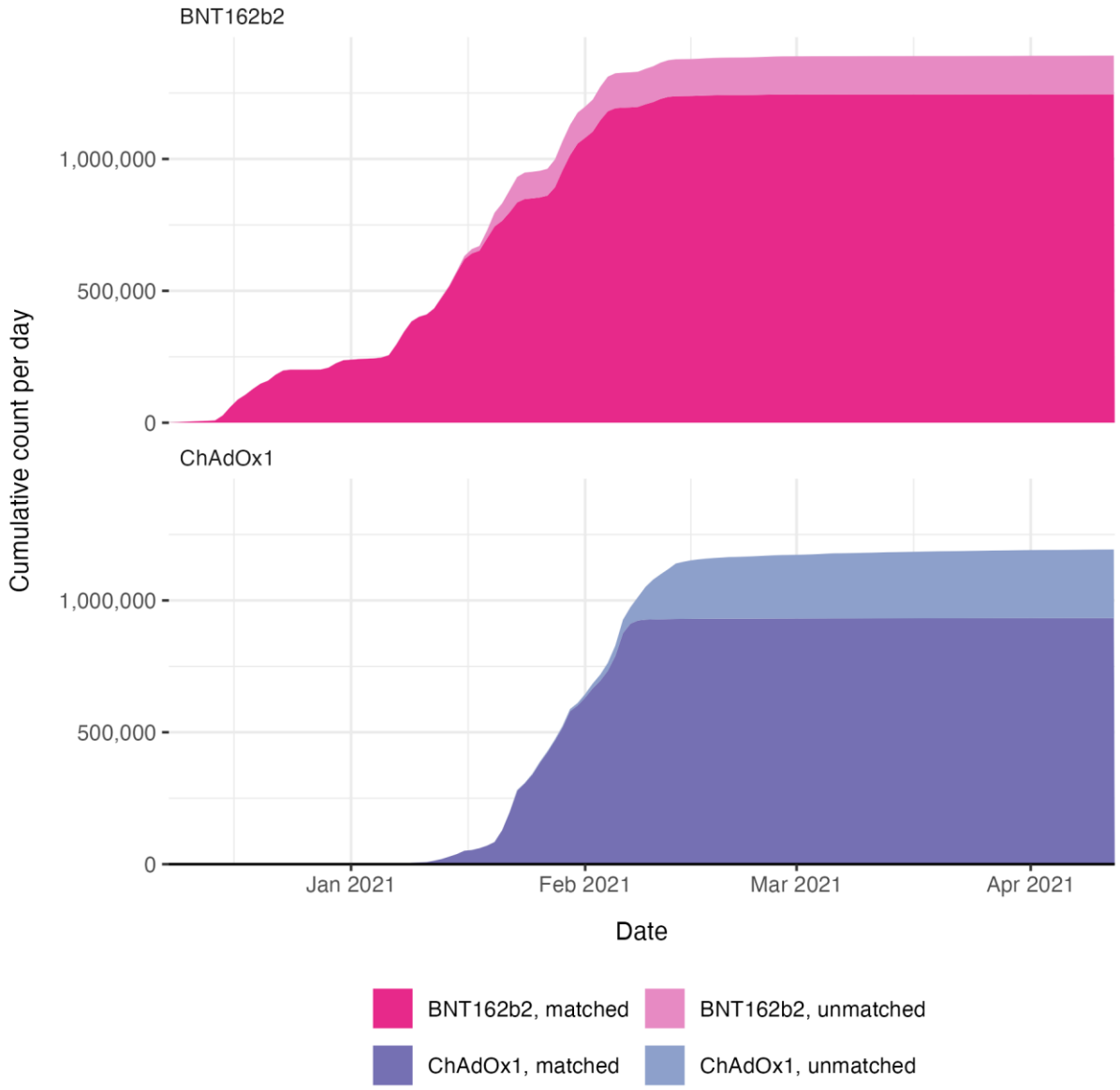
Supplement Figure 2. Flow chart showing application of the inclusion and exclusion criteria to select the single and sequential cohorts for analysis (single trial cohort shaded green, BNT162b2 and ChAdOx1 sequential trial cohorts shaded blue and orange, respectively).



TPP practice: a primary care surgery using TPP SystmOne software. Note that counts are rounded to mitigate the risk of statistical disclosure, so there may be some discrepancies between the total counts and sums of counts.

*the same unvaccinated individuals may be matched as controls in the BNT162b2 and ChAdOx1 trials

Supplement Figure 3. Cumulative number of vaccinated people matched and included in the sequential trials analysis, for each vaccine brand.



Supplement Figure 4. Kaplan-Meier estimates of the cumulative incidence of outcome events in vaccinated and unvaccinated people included in the sequential approach, by brand.

These comparisons do not account for the additional confounders modelled in the Cox regression analyses.