

Challenges in Estimating the Effectiveness of COVID-19 Vaccination Using Observational Data

William J. Hulme, PhD; Elizabeth Williamson, PhD; Elsie M.F. Horne, PhD; Amelia Green, PhD; Helen I. McDonald, PhD; Alex J. Walker, PhD; Helen J. Curtis, DPhil; Caroline E. Morton, MRCP; Brian MacKenna, MPharm; Richard Croker, MSc; Amir Mehrkar, MRCP; Seb Bacon, BA; David Evans, MPhil; Peter Inglesby, MPhil; Simon Davy, PhD; Krishnan Bhaskaran, PhD; Anna Schultze, PhD; Christopher T. Rentsch, PhD; Laurie Tomlinson, PhD; Ian J. Douglas, PhD; Stephen J.W. Evans, MSc; Liam Smeeth, PhD; Tom Palmer, PhD; Ben Goldacre, MRCPsych; Miguel A. Hernán, MD, DrPH; and Jonathan A.C. Sterne, PhD

The COVID-19 vaccines were developed and rigorously evaluated in randomized trials during 2020. However, important questions, such as the magnitude and duration of protection, their effectiveness against new virus variants, and the effectiveness of booster vaccination, could not be answered by randomized trials and have therefore been addressed in observational studies. Analyses of observational data can be biased because of confounding and because of inadequate design that does not consider the evolution of the pandemic over time and the rapid uptake of vaccination. Emulating a hypothetical “target trial” using observational data assembled during vaccine rollouts can help manage such potential sources of bias. This article describes 2 approaches to target trial emulation. In the sequential approach, on each day, eligible persons who have not yet been vaccinated are matched to a vaccinated person. The single-trial approach sets a single baseline at the start of

the rollout and considers vaccination as a time-varying variable. The nature of the confounding depends on the analysis strategy: Estimating “per-protocol” effects (accounting for vaccination of initially unvaccinated persons after baseline) may require adjustment for both baseline and “time-varying” confounders. These issues are illustrated by using observational data from 2 780 931 persons in the United Kingdom aged 70 years or older to estimate the effect of a first dose of a COVID-19 vaccine. Addressing the issues discussed in this article should help authors of observational studies provide robust evidence to guide clinical and policy decisions.

Ann Intern Med. doi:10.7326/M21-4269

Annals.org

For author, article, and disclosure information, see end of text.

This article was published at Annals.org on 2 May 2023.

The COVID-19 vaccines were developed and rigorously evaluated in randomized trials within a year of the first reports of COVID-19. The availability of effective vaccines has transformed the management of the pandemic. However, randomized trials were unable to address important questions about vaccine effectiveness because they were conducted before the Delta and Omicron variants emerged and the length of follow-up was insufficient to study the duration of protection or the benefit of booster vaccination. Clinicians, policymakers, and the public must therefore rely on evidence from observational studies. Thus, it is important to understand the challenges of using observational data to address these questions.

It is helpful to consider the hypothetical randomized trial, also known as the “target trial,” that an observational study aims to emulate (1). Consideration of the target trial helps to identify potential sources of bias in observational analyses that estimate vaccine effectiveness and to clarify analytic approaches to reduce bias. The lack of random assignment to vaccination necessitates adjustment for bias from confounding when factors influencing the outcome also influence receipt of vaccination.

In this article, we describe analytic issues that arise when data assembled during a rapid rollout of vaccines are used to estimate vaccine effectiveness. We describe 2 approaches to specifying target trials of interest and emulating them in observational data: a sequential approach based on matching vaccinated and unvaccinated persons on each day of vaccination, and a single-trial approach that splits follow-up time for each person into vaccinated and unvaccinated periods. We compare estimates of the effectiveness of a first dose of

the Pfizer-BioNTech BNT162b2 mRNA vaccine (BNT162b2) and the Oxford-AstraZeneca ChAdOx1 nCoV-19 AZD1222 vaccine (ChAdOx1) among persons in the United Kingdom aged 70 years or older using each approach.

OBSERVATIONAL DATA FROM THE U.K. COVID-19 VACCINE ROLLOUT

The United Kingdom’s COVID-19 vaccination program began in December 2020, with initial priority given to persons aged 80 years or older, health care workers, and care home residents, followed by persons aged 70 to 79 years and those with extreme clinical vulnerability (2). The BNT162b2 and ChAdOx1 vaccines were administered free of charge beginning on 8 December 2020 and 4 January 2021, respectively. The interval between the first and second doses was extended from 3 weeks (the interval used in market approval) to 12 weeks to expedite receipt of the first dose for more people. Observational studies of vaccine effectiveness using electronic health record data soon followed (3-5). The analyses reported here are based on primary care records linked to hospital, death registry, vaccination, and coronavirus testing surveillance data within the OpenSAFELY-TPP database (www.opensafely.org), which includes 24 million persons registered with English general practice primary care practices using TPP SystemOne software.

See also:

Web-Only
Supplement

THE CHALLENGES OF ESTIMATING VACCINE EFFECTIVENESS WITH OBSERVATIONAL DATA

Nonrandomized Assignment in Observational Data

The trials that established the effectiveness of COVID-19 vaccination randomly assigned persons without a prior diagnosis of COVID-19 to a vaccine or placebo and followed them until COVID-19 diagnosis, death, or the end of the study period. Although recruitment occurred across geographic locations with different and rapidly changing COVID-19 incidence, randomization balanced prognostic factors at the time of assignment ("baseline," or "time zero") between persons assigned to vaccination and those assigned to no vaccination. In the real world, persons who receive vaccination are likely to have different baseline prognostic factors for COVID-19 from those who do not. These potential baseline confounders include demographic, clinical, and behavioral characteristics that influence vaccine accessibility, acceptability, and hesitancy, as well as region and calendar period. Thus, in observational studies, unadjusted associations between vaccination and outcomes are subject to confounding bias.

As an example, consider the role of calendar period. Incidence of COVID-19 has varied dramatically during the pandemic. During December 2020 to April 2021, the number of positive SARS-CoV-2 test results in the United Kingdom peaked in early January 2021 and then decreased steadily after the U.K. lockdown was announced on 6 January 2021 (Supplement Figure 1, available at [Annals.org](#)). Rapid changes in incidence require that observational analyses account precisely for calendar time. Emulating a target trial using observational data requires adequate adjustment for calendar period and other potential baseline confounders through study design or data analysis.

Sequential Target Trial Emulation

Estimating the Intention-to-Treat Effect

In addition to adequate adjustment for confounding, emulating a target trial requires appropriate determination of time zero for each person. Time zero is easily defined as the date of vaccination for vaccinated persons, but unvaccinated persons are unvaccinated on a sequence of days until either they are vaccinated or follow-up ends. To ensure comparability of calendar time at time zero, observational data can be conceptualized as a sequence of target trials. In this sequence, on each day and in each region, eligible persons who have not yet been vaccinated are randomly assigned to receive immediate vaccination or to remain unvaccinated throughout follow-up and are then followed until COVID-19 diagnosis, death, or the end of the study period, whichever occurs first. Thus, a new trial starts on every day during the vaccination program. Each calendar day is considered as time zero for a new emulated trial, with persons who are determined to be eligible assigned to the vaccination group if they were vaccinated on that day or to the no-vaccination group if they were not vaccinated, then followed until occurrence of the outcome or the end of the

study. The trial-specific estimates of the effect of vaccination from each sequential trial can then be combined to estimate an observational analogue of the intention-to-treat effect of assignment to the intervention or, in our case, of receiving the first dose, ignoring whether a second dose is also received. Some observational analyses of the effect of COVID-19 vaccination have used this sequential-trial emulation approach (6).

There are various approaches to selecting unvaccinated persons to include in the comparison group. One or more persons can be selected at random from all eligible persons who were unvaccinated on that day and in that region, with exclusion of unvaccinated persons selected for a comparison made on a previous day. Another approach matches the vaccinated and unvaccinated person within each pair on baseline potential confounders in order to balance confounders so that the analyses need not adjust for them. It is desirable to closely match the vaccinated person and the unvaccinated person on characteristics that may reflect COVID-19 risk, such as neighborhood of residence or sociodemographic and clinical characteristics.

Estimating the Per-Protocol Effect

As vaccination programs roll out, the number of unvaccinated persons can decrease rapidly over time. This has 2 important implications for observational analyses. First, the pool of eligible unvaccinated persons will be smaller for comparisons that start later in the rollout. Second, many persons included in the no-vaccination group will be vaccinated soon afterward. As a result, the observational analogue of the intention-to-treat effect estimate may be uninformative because the no-vaccination group is contaminated by initially unvaccinated persons who deviated from the target trial protocol by being vaccinated.

To address vaccination during follow-up of persons included in the no-vaccination group, it is necessary to estimate the observational analogue of the "per-protocol effect" (the effect of receiving vaccination or no vaccination in accordance with the protocol of the target trial). This can be done by censoring persons in the no-vaccination group at the time they are vaccinated (7). Such censoring will be informative if factors that vary after baseline affect the rate at which persons in the no-vaccination group are vaccinated. In practice, however, over short timescales (weeks), the baseline and postbaseline values of most variables will be the same (8). Nonetheless, unmeasured time-varying factors (for example, respiratory symptoms) that influence both the probability of unvaccinated persons getting vaccinated and the risk of the outcome may introduce bias. Furthermore, when the time-varying confounders are themselves affected by vaccination, standard adjustment methods (for example, including the time-varying confounders in regression models) are inadequate and g methods, such as inverse probability weighting, are necessary (9, 10).

Censoring unvaccinated persons when they are vaccinated can lead to the follow-up time of vaccinated persons being longer and occurring in a different calendar period than for unvaccinated persons. Identifying pairs

of vaccinated and unvaccinated persons in each region and on each day and censoring follow-up of the vaccinated person on the day that the unvaccinated person in the pair is vaccinated can address this problem.

Single Target Trial Emulation

Rather than emulating a sequence of target trials, some observational studies have used a single-cohort approach, with follow-up starting at the beginning of the vaccine rollout and vaccination coded as a time-varying variable that switches from 0 (no vaccination) to 1 (vaccination) on the day of vaccination and stays as 1 thereafter (3, 5). Cox or Poisson models are then used to estimate a time-averaged hazard ratio for ever-vaccination versus no vaccination. This estimated hazard ratio approximates the per-protocol hazard ratio obtained by pooling sequential trials with censoring if there were no time-varying confounders and people remained eligible for all trials unless they developed COVID-19 or died. This approach can be seen as an attempt to emulate a target trial in which eligible persons are recruited at the start of the rollout (for example, 8 December 2020 for persons aged ≥ 80 years in the United Kingdom) and are randomly assigned to vaccination at different times during follow-up if they remain eligible at those times. However, this “single-trial” approach has significant disadvantages compared with the “sequential target trials” approach.

First, time-varying risk factors that are associated with vaccination after the start of rollout are “time-varying confounders” and must be adjusted for using g methods, such as inverse probability weighting of marginal structural models (as described later). In contrast, when sequential target trials are emulated, these factors are time-fixed confounders at each trial’s baseline. Many published analyses of observational data that used a single cohort with time-varying vaccination did not appear to address potential time-varying confounders (3, 5). Second, because vaccination starts at different times for different people, this approach does not naturally lead to estimation of absolute risks and cumulative incidence curves, so causal inferences are based on the time-averaged hazard ratio. Third, people may become ineligible for vaccination after the start of follow-up if, for example, they test positive for SARS-CoV-2. When sequential trials are emulated, persons with a prior positive test result are excluded at baseline. In the single-trial approach, however, persons who are vaccinated and those who are unvaccinated at a given time during follow-up can become increasingly noncomparable in terms of prior infection when the effect of vaccination on postinfection outcomes (such as hospitalization) is being estimated, and statistical methods cannot adequately adjust for this imbalance. Technically, we say that there is no “positivity” because the probability of vaccination soon after a documented infection is essentially zero, except in cases of data errors and in highly unusual circumstances. A way to manage this problem, at the expense of altering the original causal question, is to stop updating the time-varying vaccination variable and the time-varying weights after a positive test result.

ESTIMATING THE EFFECTIVENESS OF A FIRST DOSE OF BNT162B2 AND CHADOX1 AMONG PERSONS IN ENGLAND AGED 70 YEARS OR OLDER

To illustrate the aforementioned issues, we applied both the sequential approach and the single-trial approach to estimate the effectiveness of a first vaccine dose among persons aged 70 years or older, using the U.K. data described earlier.

There were 3 327 255 potentially eligible persons who were aged 70 years or older on 31 March 2020 (the date used to calculate priority group membership) and were alive on 8 December 2020. After exclusion of persons with unreliable vaccination data (0.9%), those with less than 1 year of continuous registration (3.3%), health or social care workers (0.1%), care or nursing home residents or persons who were housebound for medical reasons (5.1%), those receiving end-of-life care (1.7%), those with missing information on key demographic variables (5.1%), and those with evidence of prior SARS-CoV-2 infection (1.4%), 2 780 931 (83.6%) met eligibility criteria for subsequent analyses (Supplement Figure 2, available at [Annals.org](https://annals.org)).

We estimated per-protocol effects of a single dose of either BNT162b2 or ChAdOx1 compared with no vaccination on 3 outcomes: positive test result for SARS-CoV-2, COVID-19 hospitalization, and all-cause mortality. We estimated the effectiveness of the vaccines separately. Follow-up was censored if unvaccinated persons were vaccinated with the other vaccine and if vaccinated persons received a second dose. We conducted a sensitivity analysis in which vaccinated persons were not censored when they received a second dose. Follow-up ended on 12 April 2021, or earlier if the outcome occurred or participants de-registered from their primary care practice. During the study period, daily incidence rates of COVID-19 were 10 to 100 cases per 10 000 persons (11).

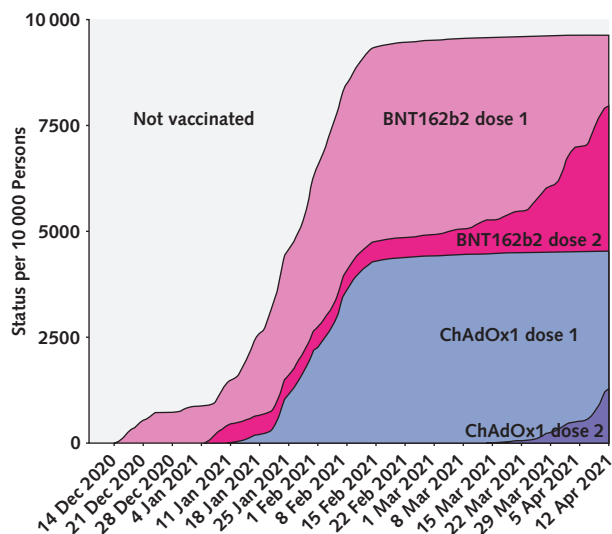
We identified potential confounders from variables used to define U.K. vaccination priority groups, government shielding guidance (12), and clinical expertise (Supplement Table 1, available at [Annals.org](https://annals.org)). We also considered 2 potential confounders that could vary after eligibility for vaccination: unplanned hospitalizations for infectious or noninfectious conditions (each categorized as not in the hospital, in the hospital, 1 to 21 days after discharge, and 22 to 28 days after discharge).

For each approach, we estimated hazard ratios for vaccination 1 to 3, 4 to 7, 8 to 14, 15 to 21, 22 to 28, 29 to 35, and 36 to 70 days after receipt of the first vaccine dose. Vaccine effectiveness was estimated as $100 \times (1 - \text{hazard ratio})$. The Supplement (available at [Annals.org](https://annals.org)) contains additional details on the OpenSAFELY data analytics platform, exclusion criteria, derivation of confounders and outcomes, the approach to dealing with missing data, and the analyses.

Characteristics of the Cohort

Vaccine coverage increased rapidly starting on 14 December 2020, initially with BNT162b2 and then with

Figure 1. Coverage of first and second dose of BNT162b2 and ChAdOx1 vaccination.



Coverage on each day was calculated as 10 000 times the number of persons in each status, divided by the number of persons alive and registered. BNT162b2 = Pfizer-BioNTech BNT162b2 mRNA vaccine; ChAdOx1 = Oxford-AstraZeneca ChAdOx1 nCoV-19 AZD1222 vaccine.

ChAdOx1 starting in early January 2021 (Figure 1). Second vaccinations with BNT162b2 began 3 weeks later, but on 31 December 2020, the U.K. Chief Medical Officers announced that the dosing interval would be increased to 12 weeks. Therefore, few second doses of ChAdOx1 were administered before March 2021. A total of 2 656 062 (96%) persons were vaccinated by the end of follow-up (1 406 637 [51%] with BNT162b2 and 1 249 425 [45%] with ChAdOx1). In the single-trial cohort, 54.0% were female, 95.7% were White, and 12.8% lived in one of the top 20% most deprived areas in England. Of 530 685 person-years of follow-up, 163 515 (30.8%) were after vaccination.

The Table shows the distribution of selected potential confounding factors together with hazard ratios for their association with vaccination, estimated using pooled logistic regression models (13–15). Supplement Table 1 shows the distribution of the full set of confounders considered in both the single- and sequential-trial cohorts. There were clear associations with vaccination, and many associations varied between BNT162b2 and ChAdOx1.

For both vaccines, vaccination rates were higher for persons in less deprived areas and lower for persons of non-White ethnicity and those with learning disabilities or a history of serious mental illness. Vaccination rates were markedly higher in persons who had received influenza vaccination in the previous 5 years.

Methods and Patient Characteristics

Sequential-Trial Approach

For the sequential approach, the first trial included persons vaccinated on the first day on which they were

eligible for vaccination. Variables used to apply the inclusion and exclusion criteria, matching variables, and baseline covariates were redefined on each trial start date, and the inclusion and exclusion criteria were reapplied on each trial start date. Matching was conducted independently for persons who received BNT162b2 and ChAdOx1, with the same unvaccinated persons available for matching in the BNT162b2 and ChAdOx1 analyses. Vaccinated persons were matched in a 1:1 ratio with persons who were not vaccinated on that day, using the following variables: age (within 3 years), Joint Committee on Vaccination and Immunisation age band (70 to 74, 75 to 79, and ≥ 80 years), sex, geographic region, and clinical vulnerability indicator. This matching was repeated on each subsequent day: Unvaccinated persons who had already been matched were no longer eligible to be unvaccinated controls on subsequent days, although they were eligible for subsequent inclusion in the vaccinated group. Time zero for each matched pair was the day of vaccination. Follow-up for each matched pair was censored if the unvaccinated person became vaccinated. We derived Kaplan-Meier estimates of the cumulative incidence of each outcome in vaccinated and unvaccinated persons included in the sequential-trials analysis. We fitted adjusted Cox models to estimate period-specific hazard ratios.

Matches were identified for 2 178 168 (82%) of 2 656 062 eligible vaccinations (1 245 267 [89%] of 1 406 637 for BNT162b2 and 932 901 [75%] of 1 249 425 for ChAdOx1) (Supplement Figure 2). Supplement Figure 3 (available at [Annals.org](https://annals.org)) shows the cumulative number of matches over time. Almost all persons vaccinated with BNT162b2 before mid-January 2021 were matched, but from mid-February 2021 there were few additional eligible vaccinations. Almost all persons vaccinated with ChAdOx1 by the end of January 2021 were matched, but few additional matches were identified after mid-February 2021.

In the BNT162b2 trials, there were 111 165 person-years of follow-up (57 045 in the unvaccinated group); during this time, there were 8337 positive test results (6297 in the unvaccinated group), 5235 COVID-19 hospitalizations (4101 in the unvaccinated group), and 5637 deaths (4731 in the unvaccinated group). In the ChAdOx1 trials, there were 71 121 person-years of follow-up (35 319 in the unvaccinated group); during this time, there were 3141 positive test results (2253 in the unvaccinated group), 1875 COVID-19 hospitalizations (1431 in the unvaccinated group), and 2703 deaths (2067 in the unvaccinated group). Supplement Table 1 shows baseline (day of vaccination) characteristics of the vaccinated and matched unvaccinated groups for each vaccine brand. As expected, the distributions of characteristics used for matching were identical in the vaccinated and unvaccinated groups. Other characteristics were also similar in the vaccinated and unvaccinated groups. The cumulative incidence of each outcome was markedly lower in vaccinated than unvaccinated persons for each vaccine brand (Supplement Figure 4, available at [Annals.org](https://annals.org)).

Table. Selected Baseline and Time-Varying Covariates for the Single-Trial Cohort, With Hazard Ratios (Estimated Using Pooled Logistic Regression) for Vaccination With BNT162b2 and ChAdOx1*

Characteristic	Participants, n (%)	Hazard Ratio (95% CI)	
		BNT162b2	ChAdOx1
Age			
Age	Median, 76 y	3.53 (3.42-3.65)	2.21 (2.12-2.30)
Age squared	(IQR, 73-82 y)	0.99 (0.99-0.99)	1.00 (1.00-1.00)
Sex			
Female	1 501 071 (54.0)	1 (reference)	1 (reference)
Male	1 279 863 (46.0)	1.06 (1.05-1.08)	0.99 (0.97-1.00)
Clinical vulnerability			
Not at risk	1 081 155 (38.9)	1 (reference)	1 (reference)
At risk	1 378 839 (49.6)	1.04 (0.99-1.09)	1.03 (0.98-1.08)
Extremely vulnerable	320 937 (11.5)	1.06 (1.01-1.12)	1.09 (1.04-1.15)
Deprivation			
1 (most deprived)	355 413 (12.8)	1 (reference)	1 (reference)
2	469 863 (16.9)	1.15 (1.12-1.18)	1.13 (1.10-1.16)
3	639 765 (23.0)	1.17 (1.14-1.20)	1.20 (1.17-1.24)
4	658 845 (23.7)	1.25 (1.22-1.28)	1.21 (1.18-1.24)
5 (least deprived)	657 045 (23.6)	1.41 (1.37-1.45)	1.21 (1.18-1.25)
Ethnicity			
White	2 660 415 (95.7)	1 (reference)	1 (reference)
Black	22 653 (0.8)	0.51 (0.44-0.58)	0.47 (0.41-0.54)
South Asian	69 393 (2.5)	0.50 (0.47-0.53)	0.54 (0.51-0.56)
Mixed	8661 (0.3)	0.63 (0.52-0.76)	0.66 (0.55-0.78)
Other	19 815 (0.7)	0.76 (0.67-0.86)	0.68 (0.60-0.76)
Morbidity count			
0	1 146 813 (41.2)	1 (reference)	1 (reference)
1	904 317 (32.5)	0.99 (0.94-1.05)	1.00 (0.95-1.06)
≥2	729 801 (26.2)	0.98 (0.91-1.05)	0.99 (0.92-1.07)
Immunosuppressed	133 599 (4.8)	1.03 (1.00-1.08)	1.05 (1.01-1.10)
Learning disabilities	2745 (0.1)	0.61 (0.48-0.78)	0.60 (0.48-0.74)
Serious mental illness	17 523 (0.6)	0.75 (0.67-0.84)	0.83 (0.76-0.92)
Influenza vaccination in previous 5 y	2 463 543 (88.6)	2.23 (2.16-2.29)	2.37 (2.30-2.43)
Time since discharge from hospitalization for infectious cause			
Not in hospital	NA (time-varying)	1 (reference)	1 (reference)
In hospital	NA (time-varying)	0.84 (0.81-0.87)	0.12 (0.10-0.16)
1-21 d	NA (time-varying)	0.40 (0.36-0.45)	0.66 (0.61-0.72)
22-28 d	NA (time-varying)	0.64 (0.53-0.78)	0.88 (0.75-1.03)
Time since discharge from hospitalization for noninfectious cause			
Not in hospital	NA (time-varying)	1 (reference)	1 (reference)
In hospital	NA (time-varying)	0.94 (0.90-0.98)	0.41 (0.31-0.55)
1-21 d	NA (time-varying)	0.43 (0.34-0.55)	0.87 (0.73-1.05)
22-28 d	NA (time-varying)	0.75 (0.54-1.05)	0.91 (0.66-1.27)

BNT162b2 = Pfizer-BioNTech BNT162b2 mRNA vaccine; ChAdOx1 = Oxford-AstraZeneca ChAdOx1 nCoV-19 AZD1222 vaccine; NA = not applicable.

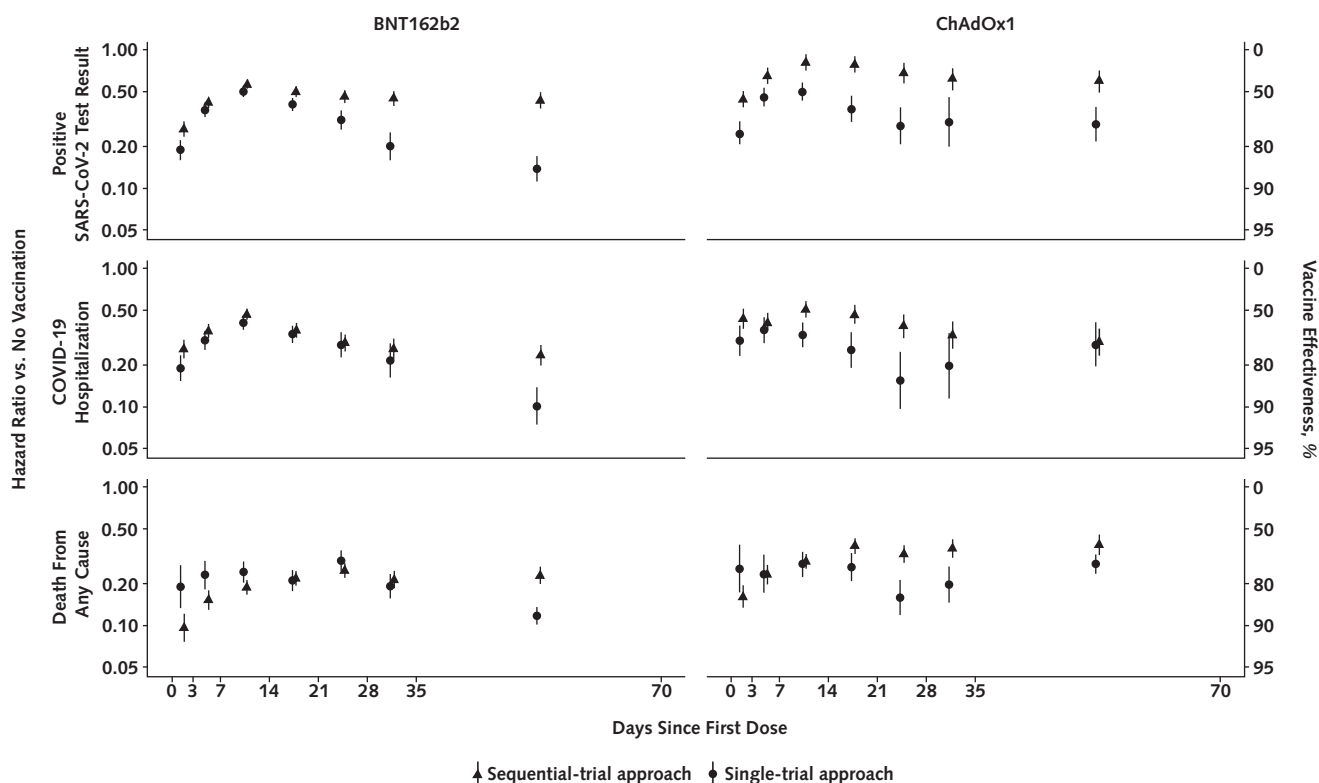
* Follow-up was censored on receipt of the other vaccine. Each pooled logistic regression model was based on a random sample of 150 000 persons because the data set contains a row for each day of follow-up for each person.

Single-Trial Approach

For the single-trial approach, follow-up started on 8 December 2020, and the baseline confounders were defined on that date. The data set had a row for each day of follow-up for each person. The time-varying vaccination status was not updated after a positive SARS-CoV-2 test result. To make computations feasible, we selected a random sample of 50 000 of the persons who did not experience the outcome and upweighted them by the reciprocal of their probability of being sampled in the analyses.

We fitted pooled logistic models (equivalent to Cox models [15]) with an indicator for vaccination and the baseline covariates shown in Supplement Table 1. We modeled calendar time using region-specific restricted cubic splines to account for rapid changes in outcome incidence rates over time and by geographic region. We accounted for time-varying confounding by the measured factors that varied after baseline by using stabilized inverse probability weights (9), which were derived from models predicting vaccination using measured baseline

Figure 2. Estimated vaccine effectiveness after ≥1 dose of BNT162b2 or ChAdOx1.



BNT162b2 = Pfizer-BioNTech BNT162b2 mRNA vaccine; ChAdOx1 = Oxford-AstraZeneca ChAdOx1 nCoV-19 AZD1222 vaccine.

and time-varying confounders. Inverse probability weights for censoring at the time of vaccination with the other brand were also derived. The probability of ChAdOx1 vaccination was zero until its first administration on 4 January 2021. The probability of vaccination in persons aged 70 to 79 years was zero until 5 January 2021, as this was the date on which they became eligible. Confidence intervals were derived using robust standard errors. Additional details on the analysis are provided on pages 2 to 4 of the Supplement.

In the single-trial approach, factors that varied after the vaccine eligibility date were considered as time-varying confounders. The Table shows associations of these factors with vaccination. Vaccination after a positive SARS-CoV-2 test result was rare, occurring in 4479 of 1 406 637 persons who received BNT162b2 and 10 713 of 1 249 425 who received ChAdOx1 (and in only 1437 and 2649, respectively, within 28 days after a positive result). Persons in or recently discharged from an unplanned hospitalization were less likely to be vaccinated (Table).

In the single-trial cohort, there were 530 682 person-years of follow-up (367 168 while unvaccinated); during this time, there were 38 853 positive SARS-CoV-2 test results (32 265 while unvaccinated), 19 821 COVID-19 hospitalizations (16 071 while unvaccinated), and 19 527 deaths (15 591 while unvaccinated).

Comparison of Estimated Vaccine Effectiveness Using Each Approach

Figure 2 and Supplement Table 2 (available at Annals.org) show estimated adjusted hazard ratios for vaccine effectiveness after 1 dose of BNT162b2 or ChAdOx1, within periods since vaccination and comparing results from the sequential- and single-trial approaches. In general, estimated vaccine effectiveness was greater with the sequential-trials approach than the single-trial approach. Hazard ratios were estimated less precisely for the sequential-trials approach because unmatched persons were excluded, and follow-up of each matched pair was censored when the unvaccinated control was vaccinated. Results from the sensitivity analysis without censoring of vaccinated follow-up time at the second dose were almost identical.

Vaccine effectiveness against a positive SARS-CoV-2 test result was estimated to be substantial immediately after vaccination and less substantial during the second week after vaccination; the respective adjusted hazard ratios during the second week after vaccination for the sequential- and single-trial approaches were 0.50 (95% CI, 0.45 to 0.54) and 0.56 (CI, 0.52 to 0.60) for BNT162b2 and 0.50 (CI, 0.43 to 0.58) and 0.83 (CI, 0.75 to 0.92) for ChAdOx1. Estimated vaccine effectiveness then increased over time: the respective adjusted hazard ratios 36 to 70 days after vaccination for the sequential- and single-trial

approaches were 0.14 (CI, 0.11 to 0.17) and 0.43 (CI, 0.37 to 0.49) for BNT162b2 and 0.29 (CI, 0.22 to 0.39) and 0.59 (CI, 0.49 to 0.71) for ChAdOx1.

Vaccine effectiveness against COVID-19 hospitalization was estimated to be greater than effectiveness against a positive SARS-CoV-2 test result and greater for BNT162b2 than for ChAdOx1. Thirty-six to 70 days after vaccination, the respective adjusted hazard ratios for the sequential- and single-trial approaches were 0.10 (CI, 0.07 to 0.14) and 0.24 (CI, 0.20 to 0.28) for BNT162b2 and 0.28 (CI, 0.20 to 0.41) and 0.29 (CI, 0.23 to 0.37) for ChAdOx1. Estimated vaccine effectiveness against death from any cause was estimated to be substantial throughout follow-up.

DISCUSSION

Observational studies of the effectiveness of COVID-19 vaccination were of crucial importance in documenting early evidence of substantial efficacy, informing policy on non-pharmaceutical interventions to reduce transmission of SARS-CoV-2, and addressing vaccine hesitancy by providing clear evidence that the benefits of vaccination outweigh its rare harms. They are now essential for understanding long-term effectiveness against emerging variants, examining evidence for waning efficacy, and studying the effectiveness of booster vaccination and novel vaccines. Potential biases can be identified and addressed through conceptualization of the target trial whose results an observational study aims to emulate.

The single-trial approach includes all persons eligible for vaccination at the start of the rollout for their eligibility group. Variables that change during follow-up and that predict both vaccination and outcomes (time-varying confounders) should be controlled for by, for example, using inverse probability weighting. However, very few persons were vaccinated in the week after a positive SARS-CoV-2 test result, so it was not possible to fully control for recent positive test results in the single-trial approach. We addressed this by changing the comparison to that being implicitly made in the context of the U.K. rollout by including all follow-up after a positive test result in the unvaccinated group (such vaccinations were not consistent with U.K. policy). In such an analysis, the causal contrast is being vaccinated without a positive test result versus being unvaccinated or vaccinated after a positive result. This issue is easily dealt with in the sequential-trials approach by excluding persons with a prior positive test result within a specified period before the trial start date from the matching process.

The substantial estimated vaccine effectiveness immediately after vaccination that was observed here and in other observational studies is inconsistent with the results of randomized trials (16–18) and is unexpected given the time required to develop an immune response to vaccination and the latent period for developing symptomatic COVID-19. This suggests that estimated effectiveness immediately after vaccination was biased due to unmeasured confounding (for example, postponement of vaccination when people presented with respiratory symptoms). Cancellation or postponement of

scheduled vaccination was not recorded, and symptoms consistent with COVID-19 were not recorded unless they led to a primary or secondary care consultation. Differential depletion of susceptible persons in the unvaccinated group over time may lead to attenuation of hazard ratios within periods defined by time since vaccination, even when true vaccine effectiveness does not change. However, such bias is likely to be minimal when effectiveness is high (19).

Electronic health records are designed to facilitate care and reimbursement; they are not designed for research. Important confounders may be incompletely recorded or not recorded. For example, mild respiratory symptoms may not be recorded, so observational analyses can only adjust for proxies, such as a recent contact with the health care system, or include sensitivity analyses that explore the potential magnitude of the bias, such as by censoring unvaccinated persons 7 days after vaccination instead of the day of vaccination (6). Interpretation of observational studies should carefully consider the potential for bias due to unmeasured confounders.

“Test-negative” designs are an alternative approach that was widely used to estimate vaccine effectiveness during the pandemic (4, 20). Studies using this design compare vaccine status between persons testing positive for the condition of interest and those testing negative, typically with additional adjustment for measured confounders. The test-negative approach is attractive because it depends less on identifying and controlling for or matching on potential confounding factors than the approaches described here (21). Although test-negative designs aim to reduce confounding due to health care access and seeking behavior, they are subject to selection biases (21–23) if characteristics that predispose people to be tested also affect the outcome. Other biases can arise; examples include if vaccination affects other conditions causing similar symptoms or if there is misclassification of vaccination or the outcome.

Policy-level and individual decisions about COVID-19 vaccination should be informed by evidence about the comparative effectiveness of different vaccination strategies. Specification of a target trial requires explicit description of the strategies being compared, which should assist decisions based on observational data analyses. The sequential approach described in this article can control for confounding factors, including those that vary during vaccine rollout, and address other biases that arise in observational data analyses. Further comparisons of estimated vaccine effectiveness based on different approaches to observational analyses may clarify the advantages and disadvantages of these approaches and facilitate rapid, robust analyses during future public health emergencies.

From The Bennett Institute for Applied Data Science, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom (W.J.H., A.G., A.J.W., H.J.C., C.E.M., B.M., R.C., A.M., S.B., D.E., P.I., S.D., B.G.); London School of Hygiene and Tropical Medicine, London, United Kingdom (E.W., H.I.M., K.B., A.S., C.T.R., L.T., I.J.D., S.J.W.E., L.S.); Population Health Sciences, University of Bristol, Bristol, United Kingdom (E.M.F.H., T.P.); Department of Epidemiology, Department of Biostatistics, and CAUSALab, Harvard T.H. Chan School of Public Health, Boston,

Massachusetts (M.A.H.); and Population Health Sciences, University of Bristol; NIHR Bristol Biomedical Research Centre; and Health Data Research UK South West Better Care Partnership, Bristol, United Kingdom (J.A.C.S.).

Note: This research was funded in part by Wellcome. NHS England is the data controller, TPP is the data processor, and the key researchers on OpenSAFELY are acting on behalf of NHS England. This implementation of OpenSAFELY is hosted within the TPP environment, which is accredited to the ISO 27001 information security standard and is NHS IG Toolkit-compliant (24); patient data have been pseudonymized for analysis and linkage using industry standard cryptographic hashing techniques; all pseudonymized data sets transmitted for linkage onto OpenSAFELY are encrypted; access to the platform is via a virtual private network connection and is restricted to a small group of researchers; the researchers hold contracts with NHS England and only access the platform to initiate database queries and statistical models; all database activity is logged; and only aggregate statistical outputs leave the platform environment, following best practice for anonymization of results, such as statistical disclosure control for low cell counts (25). The OpenSAFELY research platform adheres to the obligations of the U.K. General Data Protection Regulation and the Data Protection Act 2018. In March 2020, the Secretary of State for Health and Social Care used powers under the U.K. Health Service (Control of Patient Information) Regulations 2002 to require organizations to process confidential patient information for the purposes of protecting public health, providing health care services to the public, and monitoring and managing the COVID-19 outbreak and incidents of exposure; this sets aside the requirement for patient consent (26). This was extended in November 2022 for the NHS England OpenSAFELY COVID-19 research platform (27). In some cases of data sharing, the common law duty of confidence is met using, for example, patient consent or support from the Health Research Authority Confidentiality Advisory Group (28). Taken together, these provide the legal bases to link patient data sets on the OpenSAFELY platform. General practitioner practices, from which the primary care data are obtained, are required to share relevant health information to support the public health response to the pandemic and have been informed of the OpenSAFELY analytics platform. This study was approved by the Health Research Authority (REC reference 20/LO/0651) and by the London School of Hygiene and Tropical Medicine Ethics Board (reference 21863). Dr. Sterne is the guarantor of the study.

Disclaimer: The views expressed in this article are those of the authors and not necessarily those of the National Institute for Health and Care Research (NIHR), NHS England, Public Health England, or the Department of Health and Social Care.

Acknowledgment: The authors are grateful for the support received from the TPP Technical Operations team throughout this work and for the generous assistance from the information governance and database teams at NHS England and the NHS England Directorate.

Financial Support: This study was supported by the COVID-19 Longitudinal Health and Wellbeing National Core Study, funded by UK Research and Innovation (UKRI) Medical Research Council (MRC) (grant reference MC_PC_20059), and the COVID-19 Data and Connectivity National Core Study, led by Health Data Research

UK in partnership with the Office for National Statistics and funded by UKRI MRC (MC_PC_20058). The OpenSAFELY Platform is supported by grants from the Wellcome Trust (222097/Z/20/Z); UKRI MRC (MR/V015757/1, MR/W016729/1); National Institute for Health and Care Research (NIHR135559, COV-LT2-0073), and Health Data Research UK (HDRUK2021.000, 2021.0157). TPP provided technical expertise and infrastructure within its data center pro bono in the context of a national emergency. Drs. Horne and Sterne are funded by the NIHR Bristol Biomedical Research Centre. Dr. Sterne is funded by Health Data Research UK South-West. The funders had no role in the design of the study; collection, analysis, or interpretation of the data; writing of the report; or the decision to submit the manuscript for publication.

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M21-4269.

Reproducible Research Statement: *Study protocol, statistical code, and data set:* Access to the underlying identifiable and potentially reidentifiable pseudonymized electronic health record data is tightly governed by various legislative and regulatory frameworks and restricted by best practice. The data in OpenSAFELY are drawn from general practice data across England, where TPP is the data processor. TPP developers initiate an automated process to create pseudonymized records in the core OpenSAFELY database, which are copies of key structured data tables in the identifiable records. These pseudonymized records are linked onto key external data resources that have also been pseudonymized via SHA-512 1-way hashing of NHS numbers using a shared salt. Bennett Institute for Applied Data Science developers and principal investigators holding contracts with NHS England have access to the OpenSAFELY pseudonymized data tables as needed to develop the OpenSAFELY tools. These tools in turn enable researchers with OpenSAFELY data access agreements to write and execute code for data management and data analysis without direct access to the underlying raw pseudonymized patient data and to review the outputs of this code. All code for the full data management pipeline—from raw data to completed results for this analysis—and for the OpenSAFELY platform as a whole is available for review at <https://github.com/OpenSAFELY>. All study code is available for inspection and reuse under MIT license at <https://github.com/opensafely/covid-vaccine-effectiveness-sequential-vs-single>.

Corresponding Author: Jonathan A.C. Sterne, PhD, Population Health Sciences, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN United Kingdom; e-mail, jonathan.sterne@bristol.ac.uk.

Author contributions are available at Annals.org.

References

1. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183:758-764. [PMID: 26994063] doi:10.1093/aje/kww254
2. UK Health Security Agency. COVID-19: the green book, chapter 14a. 27 November 2020. Accessed at www.gov.uk/government/publications/covid-19-the-green-book-chapter-14a on 13 May 2021.
3. Hall VJ, Foulkes S, Saei A, et al; SIREN Study Group. COVID-19 vaccine coverage in health-care workers in England and effectiveness of BNT162b2 mRNA vaccine against infection (SIREN): a prospective, multicentre, cohort study. *Lancet*. 2021;397:1725-1735. [PMID: 33901423] doi:10.1016/S0140-6736(21)00790-X

4. Lopez Bernal J, Andrews N, Gower C, et al. Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on Covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ*. 2021;373:n1088. [PMID: 33985964] doi:10.1136/bmj.n1088
5. 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;397:1646-1657. [PMID: 33901420] doi:10.1016/S0140-6736(21)00677-2
6. Dagan N, Barda N, Kepten E, et al. BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. *N Engl J Med*. 2021;384:1412-1423. [PMID: 33626250] doi:10.1056/NEJMoa2101765
7. Hernán MA, Robins JM. Per-protocol analyses of pragmatic trials. *N Engl J Med*. 2017;377:1391-1398. [PMID: 28976864] doi:10.1056/NEJMs1605385
8. Mansournia MA, Etminan M, Danaei G, et al. Handling time varying confounding in observational research. *BMJ*. 2017;359:j4587. [PMID: 29038130] doi:10.1136/bmj.j4587
9. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168:656-664. [PMID: 18682488] doi:10.1093/aje/kwn164
10. Daniel RM, Cousens SN, De Stavola BL, et al. Methods for dealing with time-dependent confounding. *Stat Med*. 2013;32:1584-1618. [PMID: 23208861] doi:10.1002/sim.5686
11. UK Health Security Agency. Coronavirus (COVID-19) in the UK. Accessed at <https://coronavirus.data.gov.uk/details/cases> on 5 May 2021.
12. UK Health Security Agency; Department of Health and Social Care. COVID-19: guidance for people whose immune system means they are at higher risk. 24 December 2021. Accessed at www.gov.uk/government/publications/guidance-on-shielding-and-protecting-extremely-vulnerable-persons-from-covid-19/guidance-on-shielding-and-protecting-extremely-vulnerable-persons-from-covid-19 on 5 May 2021.
13. Efron B. Logistic regression, survival analysis, and the Kaplan-Meier curve. *J Am Stat Assoc*. 1988;83:414-425. [10.2307/2288857]
14. Ngwa JS, Cabral HJ, Cheng DM, et al. A comparison of time dependent Cox regression, pooled logistic regression and cross sectional pooling with simulations and an application to the Framingham Heart Study. *BMC Med Res Methodol*. 2016;16:148. [PMID: 27809784] doi:10.1186/s12874-016-0248-6
15. D'Agostino RB, Lee ML, Belanger AJ, et al. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med*. 1990;9:1501-1515. [PMID: 2281238] doi:10.1002/sim.4780091214
16. Polack FP, Thomas SJ, Kitchin N, et al; C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med*. 2020;383:2603-2615. [PMID: 33301246] doi:10.1056/NEJMoa2034577
17. Voysey M, Clemens SAC, Madhi SA, et al; Oxford COVID Vaccine Trial Group. 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. [PMID: 33306989] doi:10.1016/S0140-6736(20)32661-1
18. Baden LR, El Sahly HM, Essink B, et al; COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med*. 2021;384:403-416. [PMID: 33378609] doi:10.1056/NEJMoa2035389
19. Kahn R, Schrag SJ, Verani JR, et al. Identifying and alleviating bias due to differential depletion of susceptible people in postmarketing evaluations of COVID-19 vaccines. *Am J Epidemiol*. 2022;191:800-811. [PMID: 35081612] doi:10.1093/aje/kwac015
20. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine*. 2013;31:2165-2168. [PMID: 23499601] doi:10.1016/j.vaccine.2013.02.053
21. Sullivan SG, Tchetgen EJ, Cowling BJ. Theoretical basis of the test-negative study design for assessment of influenza vaccine effectiveness. *Am J Epidemiol*. 2016;184:345-353. [PMID: 27587721] doi:10.1093/aje/kww064
22. Westreich D, Hudgens MG. Invited commentary: beware the test-negative design. *Am J Epidemiol*. 2016;184:354-356. [PMID: 27587722] doi:10.1093/aje/kww063
23. Dean NE, Hogan JW, Schnitzer ME. Covid-19 vaccine effectiveness and the test-negative design. *N Engl J Med*. 2021;385:1431-1433. [PMID: 34496195] doi:10.1056/NEJMe2113151
24. NHS Digital. Data Security and Protection Toolkit. Accessed at <https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/data-security-and-protection-toolkit> on 18 May 2021.
25. NHS Digital. ISB1523: Anonymisation Standard for Publishing Health and Social Care Data. Accessed at <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/isb1523-anonymisation-standard-for-publishing-health-and-social-care-data> on 18 May 2021.
26. Secretary of State for Health and Social Care - UK Government. Coronavirus (COVID-19): notification to organisations to share information. 2020. Accessed at <https://web.archive.org/web/20200421171727/www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information> on 18 May 2021.
27. Secretary of State for Health and Social Care - UK Government. Coronavirus (COVID-19): notification to organisations to share information. 2022. Accessed at <https://www.gov.uk/government/publications/coronavirus-covid-19-notification-to-organisations-to-share-information/coronavirus-covid-19-notice-under-regulation-34-of-the-health-service-control-of-patient-information-regulations-2002> on 18 October 2022.
28. Health Research Authority. Confidentiality Advisory Group. Accessed at www.hra.nhs.uk/about-us/committees-and-services/confidentiality-advisory-group on 28 March 2023.

Author Contributions: Conception and design: W.J. Hulme, E. Williamson, H.I. McDonald, C.T. Rentsch, L. Tomlinson, S.J.W. Evans, L. Smeeth, M.A. Hernán, J.A.C. Sterne.

Analysis and interpretation of the data: W.J. Hulme, E. Williamson, E.M.F. Horne, A. Green, H.I. McDonald, C.T. Rentsch, L. Tomlinson, I.J. Douglas, S.J.W. Evans, L. Smeeth, T. Palmer, M.A. Hernán, J.A.C. Sterne.

Drafting of the article: W.J. Hulme, E.M.F. Horne, E. Williamson, T. Palmer, M.A. Hernán, J.A.C. Sterne.

Critical revision for important intellectual content: W.J. Hulme, E. Williamson, H.I. McDonald, K. Bhaskaran, C.T. Rentsch, S.J.W. Evans, L. Smeeth, T. Palmer, M.A. Hernán, J.A.C. Sterne.

Final approval of the article: W.J. Hulme, E. Williamson, E.M.F. Horne, A. Green, H.I. McDonald, A.J. Walker, H.J. Curtis, C.E. Morton, B. MacKenna, R. Croker, A. Mehrkar, S. Bacon, D. Evans, P. Inglesby, S. Davy, K. Bhaskaran, A. Schultze, C.T. Rentsch, L. Tomlinson, I.J. Douglas, S.J.W. Evans, L. Smeeth, T. Palmer, B. Goldacre, M.A. Hernán, J.A.C. Sterne.

Provision of study materials or patients: W.J. Hulme, A.J. Walker, C.E. Morton, R. Croker, S. Bacon, D. Evans, P. Inglesby, S. Davy.

Statistical expertise: W.J. Hulme, E. Williamson, E.M.F. Horne, K. Bhaskaran, C.T. Rentsch, S.J.W. Evans, T. Palmer, J.A.C. Sterne.

Obtaining of funding: L. Tomlinson, L. Smeeth, B. Goldacre, J.A.C. Sterne.

Administrative, technical, or logistic support: W.J. Hulme, A. Green, H.J. Curtis, C.E. Morton, B. MacKenna, A. Mehrkar, S. Bacon, D. Evans, P. Inglesby, S. Davy, A. Schultze.

Collection and assembly of data: W.J. Hulme, E. Williamson, E.M.F. Horne, A. Green, A.J. Walker, C.E. Morton, S. Bacon, D. Evans, P. Inglesby, S. Davy, I.J. Douglas