BMC Medical Research Methodology

RESEARCH

Open Access

Unmeasured confounding and misclassification in studies estimating vaccine effectiveness against hospitalisation and death using electronic health records (EHRs): an evaluation of a multi-country European retrospective cohort study

James Humphreys[1*], Nathalie Nicolay[2], Toon Braeye[3], Izaak Van Evercooren[3], Christian Holm Hansen[4], Ida Rask Moustsen-Helms[4], Chiara Sacco[5,6], Alberto Mateo-Urdiales[5], Jesús Castilla[7,8], Iván Martínez-Baz[7,8], Ausenda Machado[9], Patricia Soares[9], Brechje de Gier[10], Hinta Meijerink[11], Susana Monge[12,13], Sabrina Bacci[2], Baltazar Nunes[1] and VEBIS-EHR working group

Abstract

Background Electronic health record (EHR)-based observational studies can rapidly provide real-world data on vaccine effectiveness (VE), though EHR data may be prone to misclassification and unmeasured confounding.

Methods In VEBIS-EHR, a retrospective multi-country COVID-19 VE cohort study, we examined unmeasured confounding using a negative control outcome (death not related to COVID-19) and misclassification due to timing of data extraction. The evaluation spanned two periods (November–December 2023, January–February 2024), encompassing up to 18.7 million individuals across six EU/EEA countries. Vaccine confounding-adjusted hazard ratios (aHRs) were pooled using random-effects meta-analysis.

Results aHRs against non-COVID-19 mortality ranged from 0.35 (95% CI: 0.28–0.44) to 0.70 (0.66–0.73) when comparing vaccinated versus unvaccinated. Delaying EHR data extraction modestly increased the capture of outcome and exposure events, with some variation by vaccination status. Site-level fluctuations in aHRs did not meaningfully alter the overall pooled VE, suggesting stable estimates despite misclassification related to extraction timing.

Conclusions We observed some evidence of unmeasured confounding when using non-COVID-19 deaths as a negative outcome, though the specificity of our negative control must be considered. This result may suggest overestimation of VE, but also the need for further analysis with more specific negative control outcomes and confounding-adjustment techniques. Addressing such confounding using richer data sources and more refined

*Correspondence:
James Humphreys
j.humphreys@epiconcept.fr

Full list of author information is available at the end of the article

approaches remains critical to ensure accurate, timely VE estimates based on retrospective cohorts constructed using registry data. Extending the delay between the end of observation and data extraction modestly improves the completeness of exposure and outcome data, with limited effect on pooled VE estimates.

Keywords Electronic health records, Vaccine effectiveness, Methods, Bias, Evaluation, COVID-19, Healthy-vaccinee bias, Frailty bias

Background

Electronic health records (EHRs) have increasingly been used in epidemiological research to provide extensive real-world data for evaluating health interventions, including in studies estimating vaccine effectiveness (VE) [1]. EHRs capture routinely-collected clinical data such as diagnoses, laboratory results, and vaccination events from real-world healthcare settings, gradually proving their utility in large-scale observational epidemiological studies designed to provide timely evidence for public health policy. This factor was of particular importance during the COVID-19 pandemic [2–12].

Because EHRs are generally collected for purposes other than research, studies relying on EHR-derived data for secondary analysis must account for specific sources of potential bias—particularly information bias arising from misclassification of exposures or outcomes [13, 14]. This makes the original intent of data collection a key consideration; for example, databases designed for billing may be unique in their coding practices, which can render their records less suitable for use in research, surveillance and monitoring. Another potential source of misclassification is incomplete or delayed reporting of exposures (e.g., vaccination events), outcomes (e.g., cause of death) and confounders (e.g., comorbidity status) which, if not monitored and adjusted for, may bias effect estimates drawn from the abstracted data [15–18].

If adjustment for confounders is inadequate, the estimates may be biased whether data are based on EHR or not [19]. Investigators relying on the secondary usage of data from EHRs are not generally able to adapt their data scope (e.g., to include additional measures from patients) according to study requirements, which may make accurate effect estimation challenging if data necessary to perform key confounding bias adjustments are not available or linkable [20, 21]. Studies aiming to evaluate interventions using EHRs, including VE studies, must ensure that the presence and degree of confounding, in particular unmeasured differences in baseline risks between the exposed and unexposed, are sufficiently evaluated and described [22]. Considering these and other limitations, the internal validity of estimates based on EHR data has been discussed extensively to date [23, 24].

The focus of the present evaluation is an EHR-based VE monitoring platform, the VEBIS project (Vaccine Effectiveness, Burden and Impact Studies) funded by the European Centre for Disease Prevention and Control (ECDC) [25]. The VEBIS-EHR study aims to monitor COVID-19 VE in real time using data from population EHR across six EU/EEA countries against severe outcomes of COVID-19 including COVID-19 related hospitalizations and deaths in individuals aged 65 + or other population subgroups [25–30].

An evaluation of this multi-country monitoring platform was undertaken in June 2024. The present study concerns the internal validity component of this evaluation, which was designed to detect the presence and magnitude of bias in estimates produced by this monitoring platform. In order to assess the validity of VE estimates produced by the system, we evaluated i) unmeasured confounding in models estimating COVID-19 VE against COVID-19-related deaths using a negative control outcome to assess differences in baseline risks in the exposed versus unexposed, and ii) how the timing of EHR data extraction affects classification of exposure, outcome and measured confounders, and any impact thereof on estimates of COVID-19 VE against COVID-19 related hospitalisation and death.

Methods

VEBIS-EHR study

VEBIS-EHR is a retrospective multi-site cohort study (Belgium, Denmark, Italy, Norway, Portugal, and Spain (Navarre)) that monitors VE of the autumn 2024/25 dose against COVID-19-related hospitalisation and death. Detailed methods are published elsewhere [25–30]. Monthly VE estimates use events from rolling 8-week periods. Each site extracts electronic-health-record (EHR) data 4–6 weeks after a period ends to allow database consolidation. The methods to estimate VE according to the main protocol are briefly detailed below. All analyses were conducted in R (Version 4.4.0, http://www.r-project.org).

We analysed two 8-week periods (Nov-Dec 2023 and Jan-Feb 2024). Eligible individuals were ≥ 65 years, resident in a study area, had completed primary vaccination ≥ 180 days before their national 2023/24 booster campaign, and had neither vaccination nor documented infection in the preceding 90 days. The exposure of interest was receipt of an autumn 2023 COVID-19 booster dose. To account for the lag between vaccination and the development of protective immunity, person-time during days 1–14 post-vaccination was excluded.

Outcome events of interest were COVID-19 hospitalisation (hospital admission for severe acute respiratory infection with a positive SARS-CoV-2 test from 14 days before to 1 day after admission, or a primary COVID-19 diagnosis) and COVID-19-related death (COVID-19 recorded as cause of death, or unknown cause with death occurring within 30 days of a positive test).

Six sites extracted individual-level EHR data on exposure, outcomes and demographics, including available data to adjust for potential confounders. Covariates for adjustment included age, sex, region, comorbidities, prior number of booster doses received to date and, where available, nationality and socioeconomic status [25]. It was not possible to adjust for health-seeking-behaviours consistently across study sites due to data access constraints.

Adjusted hazard ratios (aHRs) with 95% CIs were estimated by Cox models with time-varying exposure (14–89, 90–179, ≥ 180 days since vaccination) and calendar time as the underlying scale, stratified by age 65–79 vs ≥ 80 years. Covariates used to adjust for potential confounding were measured at baseline and were included as strata in Cox models. Site-specific aHRs were pooled with random-effects meta-analysis (Paule-Mantel), and heterogeneity quantified by I^2 [31].

Among adults aged ≥ 65 years eligible for inclusion in this VEBIS-EHR study, the primary estimand is the conditional cause-specific hazard ratio, expressed as VE(t) = 1-HR(t), for COVID-19 hospitalisation (or COVID-19-related death) comparing being vaccinated (from day 14 after an autumn-2023 booster, with time-since-vaccination bands) versus unvaccinated at time t. The estimand is conditional on baseline covariates and calendar time, with subsequent doses handled by a treatment-policy strategy through time-varying exposure, and non–COVID-19 death treated as a competing event [32].

Unmeasured confounding using a negative control outcome

To detect unmeasured confounding in estimates of VE against COVID-19-related death produced according to the above-described methods, we applied the Cox regression model without pooling to a negative control outcome defined as death *not* related to COVID-19 intended to represent the inverse of the main mortality definition (COVID-19 not reported in cause of death and no record of a positive test indicating SARS-CoV-2 infection within the 30 days preceding death). The study period between 1 November to 25 December 2023, after the start of the autumn 2023 vaccination campaign, was selected to include the maximum number of study sites according to their data access, and included four VEBIS-EHR study sites which had access to death databases and could adapt death definitions at the time of the evaluation (Denmark, Navarra, Norway, and Portugal). Where estimates of aHR against non-COVID-19-related death meaningfully differ from the expected threshold of no difference in risk (aHR = 1), including upper 95% confidence intervals that do not cross this threshold, they should be interpreted as an indication of unmeasured confounding, as we do not expect COVID-19 vaccination status to influence the risk of death unrelated to COVID-19, assuming no substantial degree of misclassification in the negative control outcomes [33].

Underreporting of outcomes and misclassification of vaccine status due to delayed recording

In a separate analysis designed to evaluate the impact of extraction delay on data consolidation, we described the number of eligible participants and relevant outcome events (hospitalisations or deaths related to COVID-19) by vaccination status and study site using three different consolidation periods: 4–6 weeks post-observation period (the 'usual' delay [25]), 8–10 weeks post-observation period ('intermediate' delay), and 26–32 weeks post-observation period ('longest' delay) (Fig. 1). We calculated aHR for each extraction delay and age group according to the VEBIS-EHR methods described above, to assess the impact of these delays on estimates both at the site and pooled level [25]. The study period selected was 1 January to 25 February 2024, as it permitted the re-extraction and analysis of data up to 32 weeks later. All

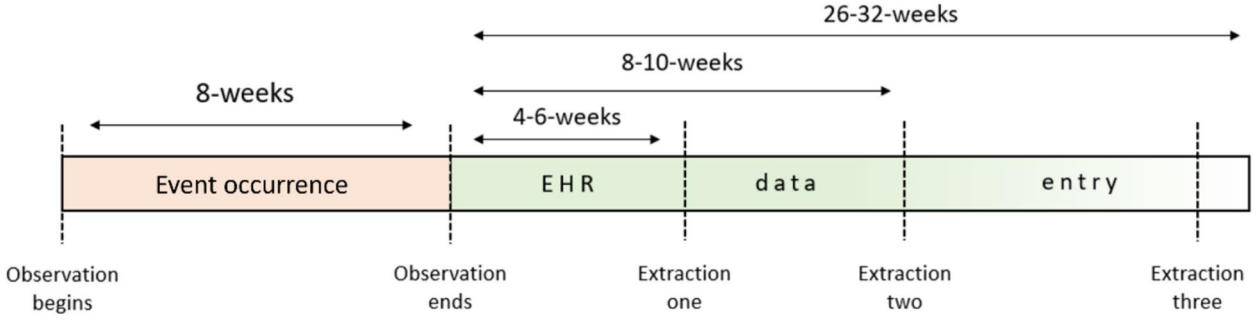

Fig. 1 Diagram of data extraction schema illustrating usual (4–6 week), intermediate (8–10 week) and late [26–32] extractions versus healthcare event occurrence and probable electronic health record data entry period

six of the VEBIS-EHR study sites (Belgium, Denmark, Navarra, Norway, Portugal, and Italy) who participated in the wider evaluation provided data for this analysis.

Results

Unmeasured confounding using a negative control outcome

Among the four study sites participating in this exercise, there were up to 4.3 million eligible individuals abstracted from source EHRs who contributed 7.7 million person-months at risk throughout the study period. A total of 8,811 relevant death events were identified, of which 291 were classified as related versus 8,520 unrelated to COVID-19 (Fig. 2). The majority (71%) of all events occurred among those aged 80 years plus.

The adjusted hazard ratio (aHR) of vaccination for non-COVID-19 mortality was consistently lower than the null effect threshold (aHR = 1.0) in all study sites, indicating that vaccinated individuals had a reduced risk of death unrelated to COVID-19 compared to the unvaccinated population. Among those aged 65–79 years, aHR ranged from 0.36 (95% CI: 0.25–0.51) to 0.58 (95% CI: 0.52–0.65), and from 0.35 (95% CI: 0.28–0.44) to 0.70 (95% CI: 0.66–0.73) among those aged ≥ 80 years-old (Fig. 2). Confidence intervals (95% CI) for estimates of aHR of vaccination against non-COVID-19-related death did not cross the highlighted threshold (aHR) in any study site or age group.

Underreporting of outcomes and misclassification of vaccine status due to delayed recording

Extracting data 4–6 weeks after the end of the observation period (25th February 2025) yielded a baseline total of 18.7 million eligible individuals, among whom 4,389 relevant COVID-19-related hospitalisations and 793 COVID-19-related deaths were identified and formed the reference for comparison with later data extractions (Annex, Tables 1, 2 and 3).

The total number of participants varied slightly with increasing delays to extraction. Decreases of 1,718 eligible participants (−0.01%), and 6,290 (−0.03%) were observed, respectively, between the intermediate and latest delay extractions versus the reference extract (Annex, Table 1). When stratified by vaccination status there was some evidence of a differential impact, as numbers of vaccinated individuals increased while unvaccinated counts declined.

Compared with the reference extraction and pooled across all study sites, there were an additional 144 (3.28%) and 148 (3.37%) COVID-19-related hospitalisations, and an additional 14 (1.77%) and 15 (1.89%) COVID-19 related deaths, respectively, in the intermediate and the latest extracts (Annex, Tables 2 and 3). When stratified by vaccination status, delayed extractions generally led to greater increases in the number of hospitalisations among vaccinated individuals (up to + 8.6%) compared with unvaccinated individuals (up to + 2.1%). Fluctuations in death counts had an opposing pattern, with a greater degree of increase among those who were unvaccinated. Most additional hospitalisations were observed in Portugal, where reported events increased by over 100% among vaccinated individuals and over 130% in the unvaccinated cohort in later extracts (Annex, Table 2). In contrast, increased COVID-19-related deaths were largely accounted for by Italy, which showed incremental increases of 2.69% to 4.18% in events recorded at the later extraction intervals (Annex, Table 3).

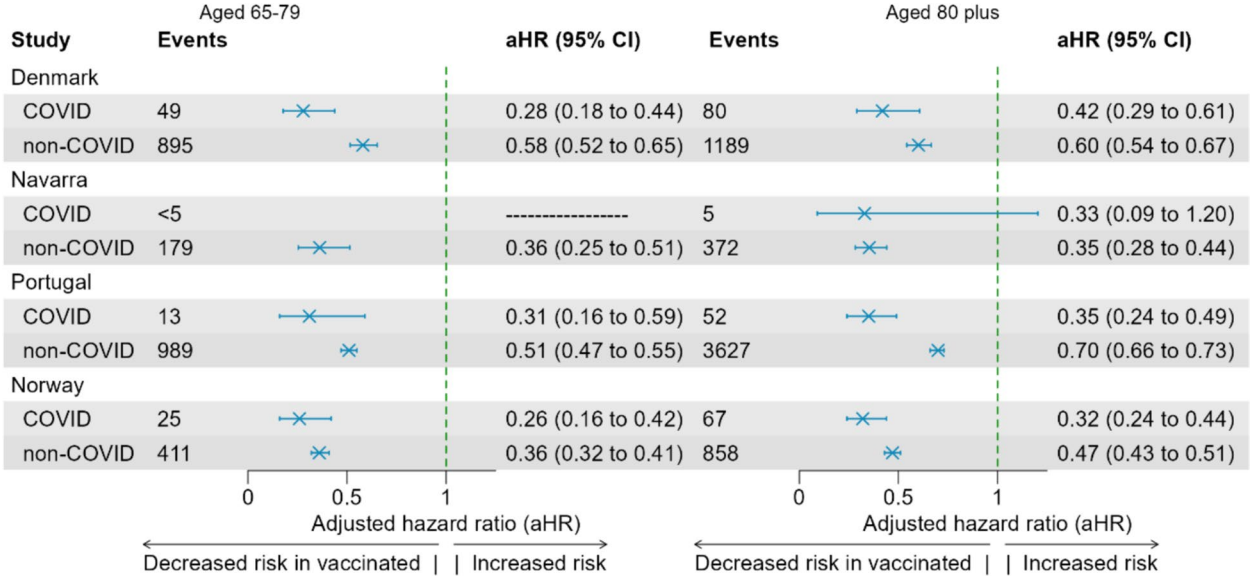

Fig. 2 Estimates of confounder-adjusted hazard ratios against COVID-19-related death (COVID) and the negative outcome control (non-COVID) among those aged 65–79 (left) and 80 plus (right)

Among individuals aged 65–79 years, the estimated aHR of vaccination against COVID-19-related hospitalisation remained stable across the usual (4–6 weeks), intermediate (8–10 weeks), and latest (26–32 weeks) data extracts for most study sites (Fig. 3). An exception was observed in Portugal, where vaccine aHR decreased from 0.56 (95% CI: 0.27–1.17) to 0.4 (95% CI: 0.23–0.7) between the usual and intermediate delays, then remained near this level based on the latest extract. After pooling estimates from all study sites using a random-effects model, pooled vaccine aHR slightly decreased from 0.53 (95% CI: 0.45–0.61) under the usual extraction to 0.52 (95% CI: 0.45–0.6) based on the intermediate extract, and further slightly still to 0.51 (95% CI: 0.44–0.59) based on the latest extract.

Among those ≥ 80-years, vaccine aHR against hospitalisation was also generally consistent across data extracts in most sites (Fig. 4). In Portugal, however, aHR fluctuated between 0.49 (95% CI: 0.3–0.78) based on the usual delay to 0.77 (95% CI: 0.56–1.06) in the intermediate delay, remaining stable thereafter based on the latest extract. When pooling estimates from all study sites, the overall vaccine aHR estimate increased slightly from 0.63 (95% CI: 0.55–0.71) to 0.65 (95% CI: 0.58–0.73) in the intermediate extract, remaining stable thereafter.

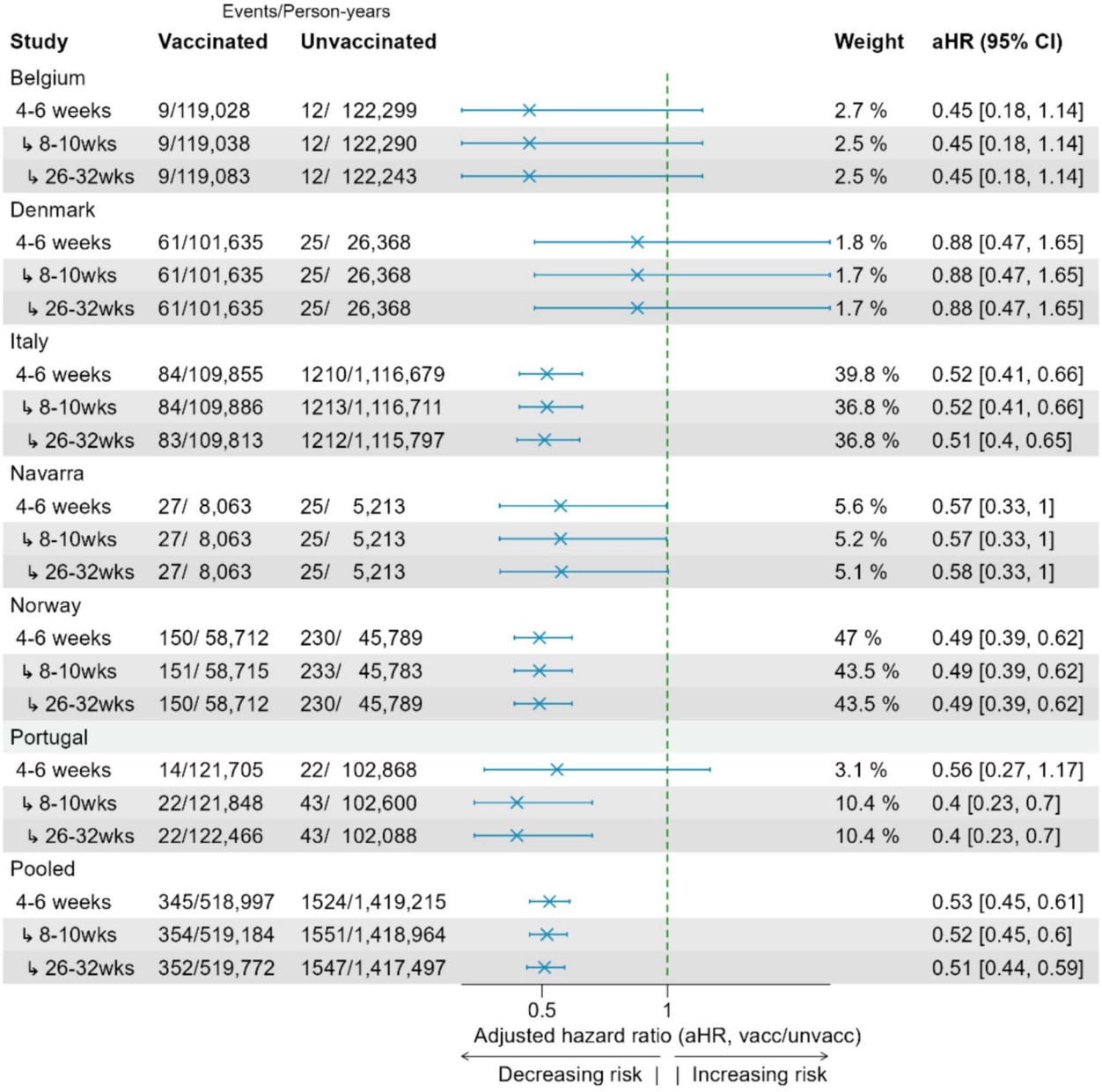

Fig. 3 Estimates of confounder-adjusted hazard ratios against hospitalisation due to COVID-19 stratified by delay group among those aged 65–79

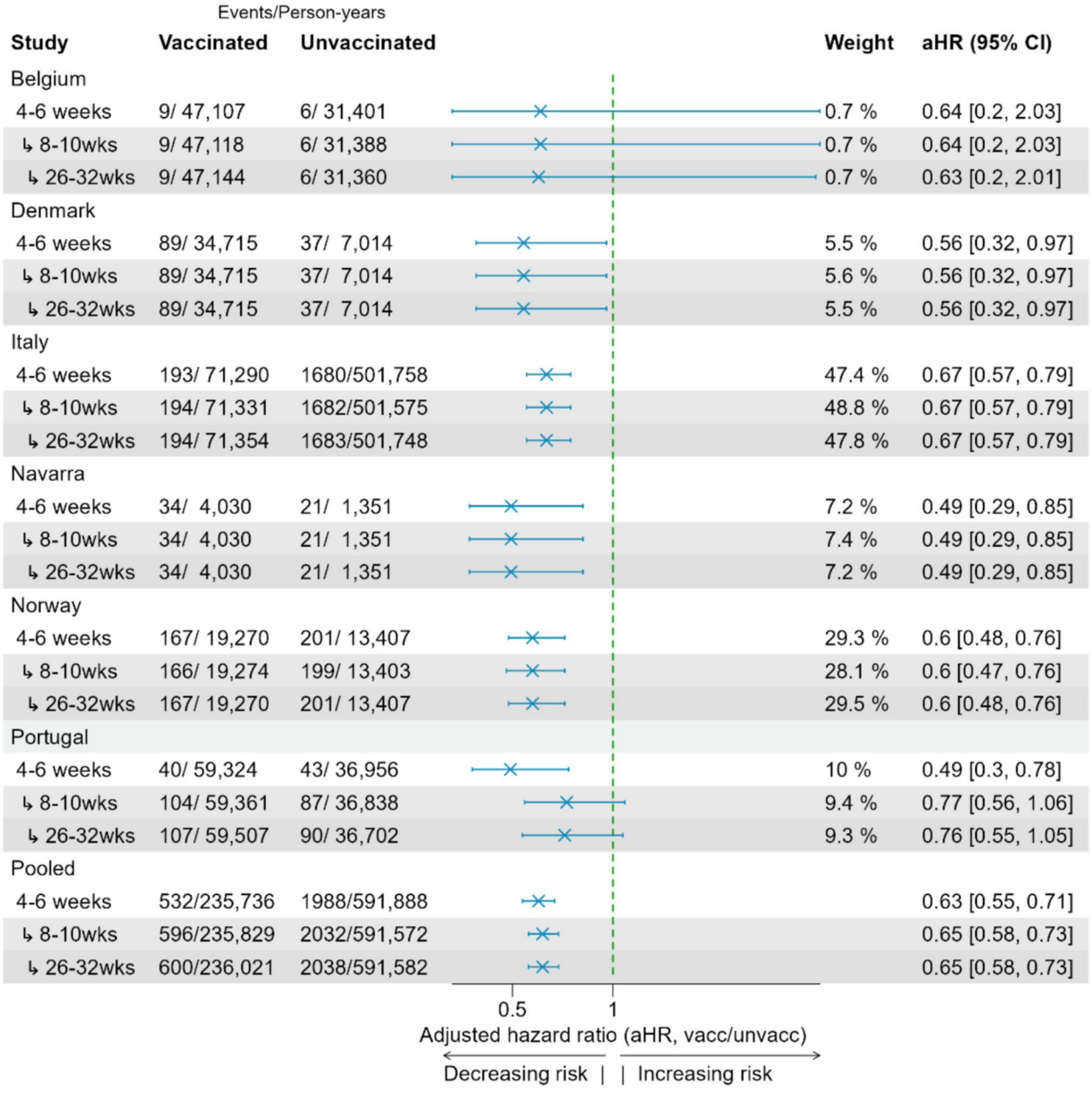

Fig. 4 Estimates of confounder-adjusted hazard ratios against hospitalisation due to COVID-19 stratified by delay group among those aged 80 plus

In the models estimating vaccine aHR against COVID-19-related death in participants aged 65–79 years, fluctuations were no greater than 0.02 in terms of absolute aHR, though estimates were often not at all changed, when compared across the three extracts. As a result, the pooled estimate for this age group and outcome remained essentially unchanged (Fig. 5).

For individuals aged 80 years and older, some study sites reported shifts in vaccine aHR for COVID-19-related death using longer extraction delays (Fig. 6). In Italy, estimates of vaccine aHR showed a steady decline from 0.60 (95% CI: 0.37–0.95) using data extracted with the usual delay (4–6 weeks), to 0.58 (95% CI: 0.36–0.92) based on an intermediate delay (8–10 weeks), and finally to 0.55 (95% CI: 0.35–0.88) based on the data extracted after the longest delay (26–32 weeks). In contrast, Navarra reported that vaccine aHRs that remained consistent between the usual and intermediate extracts at 0.53 (95% CI: 0.16–1.77), but which rose slightly in the latest extraction to 0.58 (95% CI: 0.19–1.77). Because Italy accounted for 27–30% of the overall pooled estimate weighting, reflecting its large number of events and person-months, the gradual decrease seen in Italy was similarly evident at the pooled level. The pooled aHR estimates for death in this older age group remained steady in subsequent extractions.

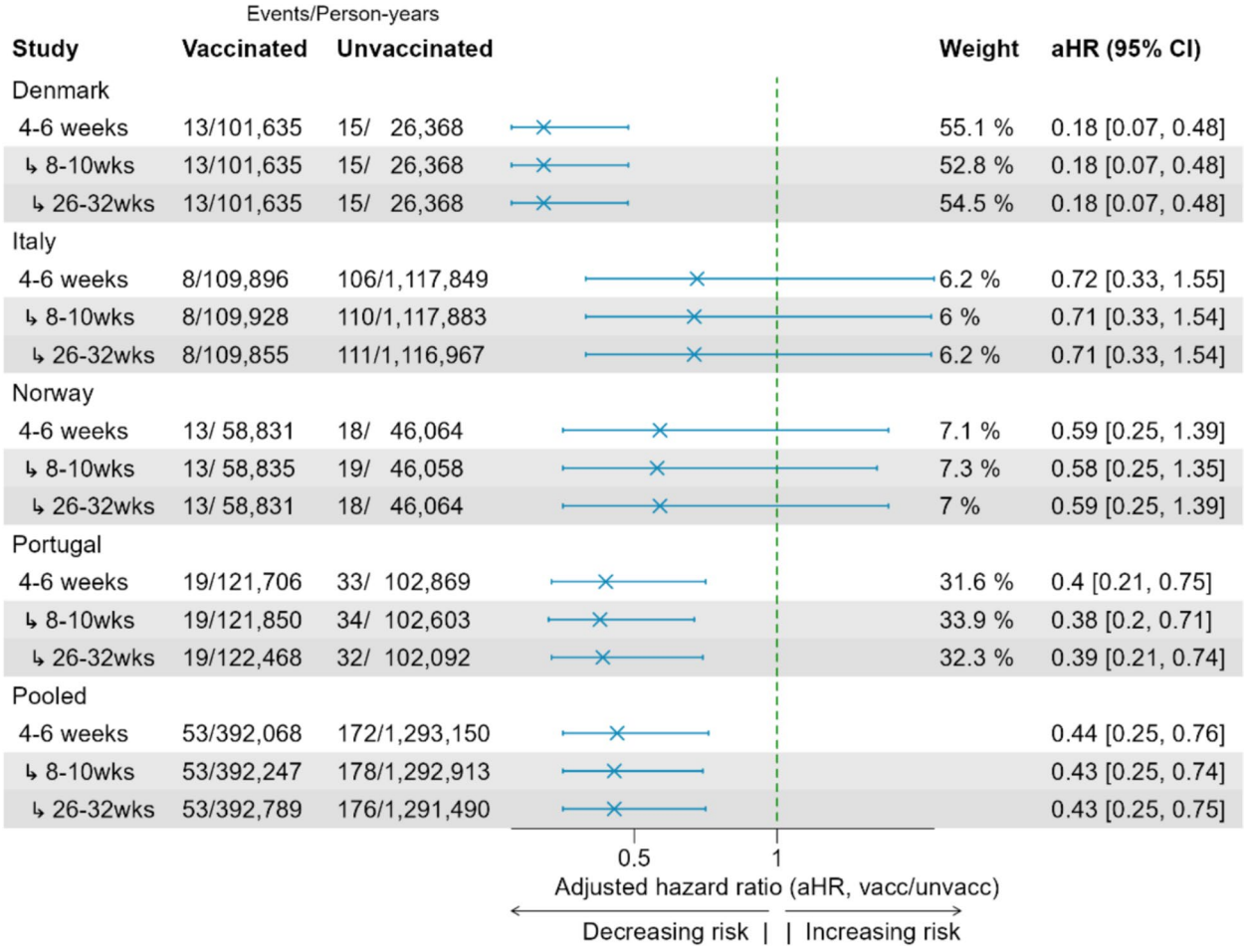

Fig. 5 Estimates of confounder-adjusted hazard ratios against death related to COVID-19 stratified by delay group among those aged 65–79

There were no substantial changes in heterogeneity (I^2) between study sites in any of the four models across both outcomes (hospitalisation or death related to COVID-19) and age groups (65–79 or 80 +). Where estimates at the study site level fluctuated, there was generally an increase in the precision of estimates based on data extracted later, likely due to a greater number of events in these extracts versus earlier delays.

Discussion

Using a negative control outcome, we found that unmeasured confounding may influence VEBIS-EHR's vaccine effectiveness (VE) estimates. Although data extraction timing can affect the completeness of records, especially for hospitalisations, though these effects appear limited to certain study sites and do not materially alter the pooled VE. Our findings are relevant to VEBIS-EHR but also to other observational, retrospective cohort COVID-19 VE studies with similar designs and covariates. Such studies are common, as shown in recent meta-analyses [34], and therefore the magnitude and direction of bias we identify are likely to apply more broadly to EHR- and registry-based COVID-19 VE evaluations that lack direct measures of health status or health-seeking behaviour. By quantifying this bias using a negative control outcome, our work contributes to ongoing efforts to standardise the assessment of unmeasured confounding in vaccine effectiveness research.

Our negative control outcome analysis demonstrated that adjusted hazard ratios for non-COVID-19 mortality consistently fell below the null, indicating a 30–65% reduced risk of non-COVID-19 deaths among vaccinated individuals. This finding suggests a "healthy-vaccinee effect" and/or the presence of extremely frail individuals in unvaccinated cohorts [35–41]. Such patterns, previously noted in influenza studies, imply that individuals who are severely frail or in end-of-life care often do not receive vaccinations and may die earlier, thereby artificially elevating VE estimates if not properly adjusted [36–39]. Approaches to mitigating this bias include measuring frailty or health-seeking behaviours more precisely, restricting cohorts to individuals with better prognoses, and accounting for baseline mortality risk [36–38]. Although VEBIS-EHR excludes long-term care

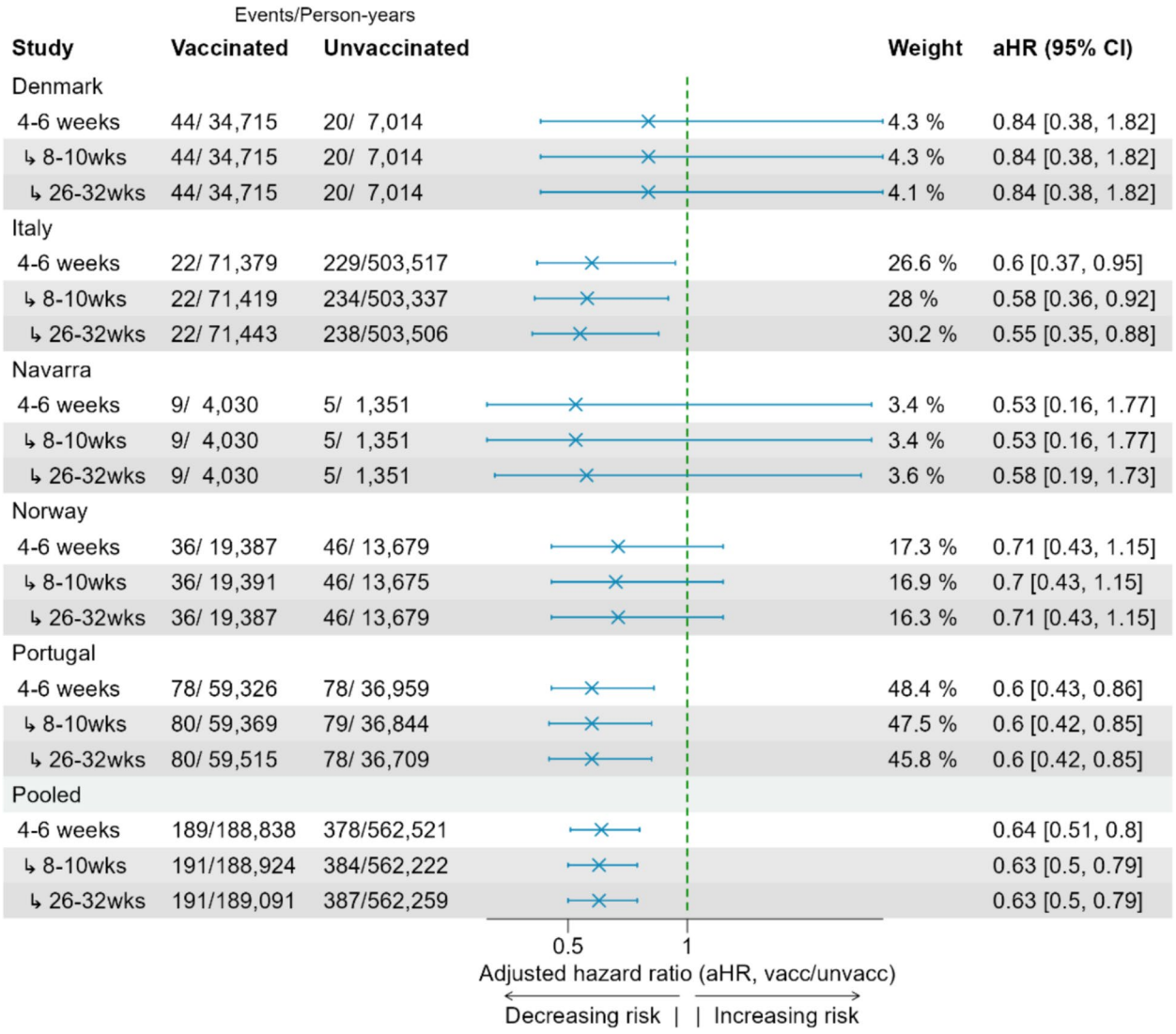

Fig. 6 Estimates of confounder-adjusted hazard ratios against death related to COVID-19 stratified by delay group among those aged 80 plus

residents and adjusts for comorbidities, our results suggest these strategies may not fully capture residual frailty. Future evaluations will therefore examine end-of-life status and health-seeking indicators to address confounding further.

Another potential explanation is the choice of negative control outcome. Although we defined non-COVID-19 deaths as having no mention of COVID-19 in the cause of death and no recent positive test, this may not be sufficiently specific (e.g., there could be unrecognised post-COVID-19 complications). Misclassification of COVID-19 deaths as non-COVID-19 deaths is a plausible concern. If vaccination reduces COVID-19 mortality, and some COVID-19 deaths are incorrectly classified as non-COVID-19, then fewer misclassified deaths will occur among vaccinated individuals, which would tend to push the hazard ratio for the negative control outcome modestly below 1. However, simple predictive value calculations indicate that the scale of this bias is limited. Even with conservative assumptions, for example assuming just 70% of all deaths are truly non-COVID-19 and that specificity of the non-COVID-19 definition is 0.70, the positive predictive value for non-COVID-19 death is around 0.85. Using more realistic European patterns for 2023, with at least 90% of deaths non-COVID-19 and specificity between 0.80 and 0.90, the positive predictive value exceeds 0.97 to 0.99, implying that only a very small fraction of negative control outcome events are misclassified COVID-19 deaths. Under an even more pessimistic scenario in which 30% of all deaths are due to COVID-19 and half of those are misclassified, the resulting contamination would shift the negative control outcome hazard ratio by only about 5–10% below 1 and therefore cannot account for the much larger apparent protective

associations that we observe. Empirical estimates of excess mortality in Europe are consistent with these calculations, as they show that non-COVID-19 deaths accounted for at least 70% of total mortality even during peaks in SARS-CoV-2 circulation and over 90% during most of 2023. Further, other investigations using highly specific outcomes (e.g., bone fractures) still report spurious protection estimates [41–43]. This strengthens the case that unmeasured confounding is not merely an artifact of choosing a less specific negative control outcome.

Addressing unmeasured confounding will likely require richer data sources, more thorough confounder evaluation, and additional analytic methods. Greater data sharing, for instance through the European Health Data Space, could facilitate linking healthcare records with other sectors, thereby capturing more detailed information on health and behaviours [44, 45]. Novel data-engineering approaches, including machine learning tools to extract relevant details from unstructured text, may also help identify and exclude individuals who are ineligible or exceptionally frail, thus improving the validity of VE estimates.

In examining data extraction timing, we observed that extending the delay between observation and extraction modestly increased the number of recorded outcomes in some study sites, especially Portugal. This suggests that records initially classified as unvaccinated or without hospitalisations may later be updated as vaccinated or hospitalised. Indeed, in Portugal, clinical coding is performed at discharge, which may lead to an increase in time between event occurrence and recording in registries. At the time of extraction, thirty days are generally allowed to have elapsed since the final event that may have occurred during the period, given the last date of follow-up. This is designed to allow for consolidation of data and recording of events, though our findings indicate that this delay is insufficient for some systems, such as those used in Portugal that rely on discharge coding. Nevertheless, a balance must be found between delays that allow for complete data consolidation and those that permit timely estimation of VE to support policy. In Italy and Portugal, unvaccinated cohorts shrank in tandem with increases in vaccinated cohorts, pointing to delayed vaccination data entry, although this pattern was not uniform across all sites. Site-specific variability in data completion implies that local validation and sensitivity analyses are critical when interpreting EHR-based VE studies.

Notably, despite these site-level differences, pooled VE estimates remained stable. This finding supports the current 4–6-week data extraction window as a reasonable compromise between timeliness and completeness. Shorter windows (e.g., under four weeks) would likely introduce more misclassification, though we did not formally investigate that scenario. Our results have prompted ongoing analyses within VEBIS-EHR to determine whether identifiable patterns of delayed reporting can guide adjustments to VE estimates.

Strengths and limitations

A major strength of our study is its multi-country design, which uses a common protocol and a large, representative dataset of routine healthcare interactions from multiple European nations—enhancing both comparability and external validity.

However, differences in data entry or coding practices across sites may limit uniformity, as suggested by our observations of delayed event reporting. Additionally, while negative control outcomes can signal unmeasured confounding, they cannot specify its exact sources. Non-COVID-19 mortality may also be insufficiently specific to serve as an ideal control. Our findings are likely to be impacted by some degree of misclassification, whereby COVID-19 deaths were misclassified as any other cause. In theory, this would bias our estimates of VE against non-COVID-19 mortality below 1, leading to an overestimation of the protective association between vaccination status and non-COVID mortality. That said, a basic simulation of the impact of such misclassification indicates that this phenomenon may not sufficiently explain the protected associations reported by the present study. Even pessimistic assumptions of the specificity of non-COVID mortality classification (50–70%) fail to simulate a large enough level of misclassification to explain the associations presently reported, especially given the low overall proportion of mortality attributable to COVID-19 during the study period [46].

The inclusion of other negative outcome controls such as non-COVID hospitalisation outcomes would have provided a more complete picture of unmeasured confounding and the impact of misclassification on mortality negative outcome control analyses, though data on non-mortality outcomes not related to COVID-19 were not sufficiently available across all study sites [47, 48]. Future analyses should test more-specific negative outcomes and consider negative exposures, or double negative control [49] as complementary methods.

The main study methods are limited by the lack of adjustment for confounding arising from underlying differences in health and health-seeking behaviours for those vaccinated versus those who do not vaccinate during autumn campaigns. This limitation is supported by the findings of our present study, which indicates unmeasured confounding. In addition, potential confounders were measured at baseline and were not treated as though they may vary over time. This could have key implications should a covariate in fact vary from this baseline measurement during the study period. None

of the included covariates (age, sex, region, socioeconomic index, and comorbidity status) were expected to vary substantially over the 8-week period being studied, however, and so the anticipated impact of this was low. We explicitly considered whether time-changing confounders could be material in our setting. For instance, immunity from prior infection can evolve over time, but (i) prior-infection dates were of variable quality across sites and (ii) we restricted to individuals ≥ 90 days since prior infection to mitigate short-term effects of recent infection.

Implications for COVID-19 vaccine monitoring and future research

Near real-time COVID-19 VE monitoring is increasingly valuable for public health decision-making, yet few multi-country EHR-based systems have been evaluated, and none, to our knowledge, have investigated how data extraction timing might affect VE. Our results underscore the importance of addressing unmeasured confounding in studies with similar designs and adjusted for similar confounders, by better quantifying the degree and direction of confounding due to underlying differences in the populations being compared. Our findings add to the current literature further evidence of the direction and magnitude of bias evident in observational studies based on EHR data without sufficient adjustment for underlying health and health-seeking-behaviours. Efforts to better measure this bias may require identification of more specific negative control outcomes, and techniques like negative control exposures or instrumental variables should be considered [49–53]. Once these sources of bias have been characterised, researchers should seek to expand data linkage to capture additional health determinants (e.g., health-seeking behaviours, socio-economic factors) to control for such bias. Alternative study designs such as self-controlled case series, which can reduce biases arising from inter-individual differences in baseline health, and g-methods, which may better account for time-varying confounding (e.g., hospital visit count as a proxy for healthcare-seeking-behaviours), also warrant further exploration [54].

As EHR-based monitoring systems mature, more granular data might become available, allowing enhanced confounder control and, by extension, more robust VE estimates. In the interim, the relative stability of pooled estimates despite delayed reporting is reassuring. A 4–6-week extraction window appears adequate, although site-specific patterns of data entry should be closely tracked and longer delays considered as part of sensitivity analyses—particularly in settings with known reporting lags.

Conclusions

These findings demonstrate that, in VEBIS-EHR, data extraction timing has a modest influence on classification of exposures and events but does not compromise pooled VE estimates. However, negative control outcome analysis revealed notable unmeasured confounding that likely inflates VE results; although the specificity of the chosen negative control outcome remains a factor. Future work should address this confounding through improved data collection, linkage, and analytic strategies, while remaining alert to site-specific differences in data capture.

Annex

Table 1 Total initial and net proportional change in record count stratified by vaccination status and study site calculated using each subsequent data extraction versus usual delay (4–6 weeks after study end on 25th February, between 24th March and 7th April 2024)

Study site	Vaccination status	4–6 weeks (Total eligible individuals, ref)	8–10 weeks (Net Δ, %)	26–32 weeks (Net Δ, %)
Belgium	Unvaccinated	1,026,123	−138 (−0.01%)	−638 (−0.06%)
	Vaccinated	1,108,665	147 (0.01%)	616 (0.05%)
Denmark	Unvaccinated	222,142	0 (-)	−1 (< 0.001%)
	Vaccinated	892,706	0 (-)	0 (-)
Italy	Unvaccinated	10,897,218	−856 (−0.01%)	−5,920 (−0.05%)
	Vaccinated	1,296,408	430 (0.03%)	118 (0.01%)
Navarra	Unvaccinated	43,598	0 (-)	0 (-)
	Vaccinated	80,357	0 (-)	0 (-)
Norway	Unvaccinated	399,417	−55 (−0.01%)	0 (-)
	Vaccinated	522,671	54 (0.01%)	0 (-)
Portugal	Unvaccinated	969,843	−2557 (−0.26%)	−6,839 (−0.7%)
	Vaccinated	1,239,445	1,257 (0.1%)	6,374 (0.51%)
Overall	Unvaccinated	13,558,341	−3,606 (−0.03%)	−13,398 (−0.1%)
	Vaccinated	5,140,252	1,888 (0.04%)	7,108 (0.14%)
	All	18,698,593	−1,718 (−0.01%)	−6,290 (−0.03%)

Table 2 Total initial and net proportional change in event (hospitalisations related to COVID-19) count stratified by vaccination status and study site, calculated using each subsequent data extraction versus usual delay (4–6 weeks after study end on 25th February, between 24th March and 7th April 2024)

Outcome	Site	Vaccination status	4–6 weeks (ref)	8–10 weeks (Net Δ, %)	26–32 weeks (Net Δ, %)
Hospitalisation due to COVID-19	Belgium	Unvaccinated	18	0 (-)	0 (-)
		Vaccinated	18	0 (-)	0 (-)
	Denmark	Unvaccinated	62	0 (-)	0 (-)
		Vaccinated	150	0 (-)	0 (-)
	Italy	Unvaccinated	2,890	5 (0.17%)	5 (0.17%)
		Vaccinated	277	1 (0.36%)	0 (-)
	Navarra	Unvaccinated	46	0 (-)	0 (-)
		Vaccinated	61	0 (-)	0 (-)
	Norway	Unvaccinated	431	1 (0.23%)	0 (-)
		Vaccinated	317	0 (-)	0 (-)
	Portugal	Unvaccinated	65	65 (100%)	68 (104.62%)
		Vaccinated	54	72 (133.33%)	75 (138.89%)
	Overall	Unvaccinated	3,512	71 (2.02%)	73 (2.08%)
		Vaccinated	877	73 (8.32%)	75 (8.55%)
		All	4,389	144 (3.28%)	148 (3.37%)

Table 3 Total initial and net proportional change in event (death related to COVID-19) count stratified by vaccination status and study site, calculated using each subsequent data extraction versus usual delay (4–6 weeks after study end on 25th February, between 24th March and 7th April 2024)

Outcome	Site	Vaccination status	Total (ref)	8–10 weeks (Net Δ, %)	26–32 weeks (Net Δ, %)
Death related to COVID-19	Belgium	Unvaccinated	*Not reported*		
		Vaccinated	Not reported		
	Denmark	Unvaccinated	35	0 (-)	0 (-)
		Vaccinated	57	0 (-)	0 (-)
	Italy	Unvaccinated	335	9 (2.69%)	14 (4.18%)
		Vaccinated	30	0 (-)	0 (-)
	Navarra	Unvaccinated	5	0 (-)	0 (-)
		Vaccinated	9	0 (-)	0 (-)
	Norway	Unvaccinated	64	1 (1.56%)	0 (-)
		Vaccinated	49	0 (-)	0 (-)
	Portugal	Unvaccinated	111	2 (1.8%)	−1 (−0.9%)
		Vaccinated	97	2 (2.06%)	2 (2.06%)
	Overall	Unvaccinated	551	12 (2.18%)	13 (2.36%)
		Vaccinated	242	2 (0.83%)	2 (0.83%)
		All	792	14 (1.77%)	15 (1.89%)

Acknowledgements
The authors would like to thank all of those participating in data collection and the production of estimates of VE as without their work these results wouldn't be available to the scientific community and public.

Authors' contributions
S Bacci, N Nicolay, J Humphreys, B Nunes, and S Monge conceived the study. J Humphreys, B Nunes, N Nicolay and S Monge conceived the methods. All authors from Public Health institutions at each study site were responsible for data management and analysis at the study site level. J Humphreys was responsible for pooling site level estimates. J Humphreys drafted the manuscript, with the help of B Nunes, A Nardone, and E Kissling. All authors contributed to the interpretation of the results and critically reviewed the manuscript. All authors approved the final version of this manuscript. All authors within the VEBIS-EHR working group made a substantial contribution to the conception or design of the work, critically revised the manuscript, provided their final approval of the version to be published, and agreed to be accountable for all aspects of the work.

Funding
All public health organizations involved received funding from the European Centre for Disease Prevention and Control (ECDC) implementing Framework Contract [ECDC/2021/018] 'Vaccine effectiveness and impact of COVID-19 vaccines through routinely collected exposure and outcome using health registries' [RS/2022/DTS/24104].

Data availability
The authors cannot share the data used for this study, access to which should be requested from the data owner institutions following their respective procedures.

Declarations

Ethics approval and consent to participate
All study sites participating in this study conformed with their respective national and EU ethical and data protection requirements. Ethical statements for each of the participating study sites:

Belgium: Data linkage and collection within the data-warehouse have been approved by the information security committee. The study was conducted in accordance with the Declaration of Helsinki. Ethical approval was granted for the gathering of data from hospitalised patients by the Committee for Medical Ethics from the Ghent University Hospital (reference number BC-07507) and authorisation for possible individual data linkage using the national register number from the Information Security Committee (ISC) Social Security and Health (reference number IVC/KSZG/20/384). Linkage of hospitalised patient data to vaccination and testing within the LINK-VACC project was approved by the Medical Ethics Committee UZ Brussels-VUB on 3 February 2021 (reference number 2020/523), and authorisation from the ISC Social Security and Health (reference number IVC/KSZG/21/034).
Denmark: Only administrative register data was used for the study. According to Danish law, ethics approval is exempt for such research, and the Danish Data Protection Agency, which is dedicated ethics and legal oversight body, thus waives ethical approval for the study of administrative register data when no individual contact of participants is necessary, and only aggregate results are included as findings. The study is, therefore, fully compliant with all legal and ethical requirements, and there are no further processes available regarding such studies.
Navarre (Spain): The study was approved by Navarre's Ethical Committee for Clinical Research, which waived the requirement of obtaining informed consent.
Norway: Ethical approval was granted by Regional Committees for Medical and Health Research Ethics (REC) Southeast (reference number 122745). The Norwegian Institute of Public Health has performed a Data Protection Impact Assessment (DPIA) for Beredt C19.
Portugal: The study received approval from the Ethical Committee and the Data Protection Officer of the Instituto Nacional de Saúde Doutor Ricardo Jorge. Given that data was irreversibly anonymised, the need for the participants' informed consent was waived by the Ethical Committee.
Italy: This study, based on routinely collected data, was not submitted for approval to an ethical committee because the dissemination of COVID-19 surveillance data was authorised by the Italian law N. 52 of 19 May 2022, following the law decree N. 24 of 24 March 2022 (Article n. 13). Based on the same acts, the information on COVID-19 vaccination was retrieved by the Italian National Institute of Health using data from the National Immunisation Information System of the Italian Ministry of Health. Because of the retrospective design and the large size of the population under study, in accordance with the Authorisation n. 9 released by the Italian data protection authority on 15 December 2016, the individual informed consent was not requested for the conduction of this study.

Consent for publication
All authors give their consent for publication.

Competing interests
The authors declare no competing interests.

Author details
[1]Epiconcept, Paris, France
[2]Vaccine Preventable Diseases and Immunisation, European Centre for Disease Prevention and Control (ECDC), Solna, Sweden
[3]Sciensano, Juliette Wytsmanstraat 14., Elsene 1050, Belgium
[4]Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut, Copenhagen, Denmark
[5]Infectious Diseases Department, Istituto Superiore Di Sanità, Rome, Italy
[6]European Programme On Intervention Epidemiology Training (EPIET), European Centre for Disease Prevention and Control, Stockholm, Sweden
[7]Instituto de Salud Pública de Navarra - IdiSNA, Pamplona, Spain
[8]CIBER of Epidemiology and Public Health (CIBERESP), Madrid, Spain
[9]Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisbon 1600-609, Portugal
[10]Center for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
[11]Norwegian Institute of Public Health (NIPH), Oslo, Norway
[12]Department of Communicable Diseases, National Centre of Epidemiology, Institute of Health Carlos III, Madrid, Spain
[13]CIBER On Infectious Diseases, Madrid, Spain

Received: 6 August 2025 / Accepted: 8 December 2025
Published online: 17 December 2025

References

1. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. Annu Rev Public Health. 2016;37:61–81.
2. Tartof SY, Slezak JM, Fischer H, Hong V, Ackerson BK, Ranasinghe ON, et al. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. The Lancet. 2021;398(10309):1407–16.
3. Barda N, Dagan N, Cohen C, Hernán MA, Lipsitch M, Kohane IS, et al. Effectiveness of a third dose of the BNT162b2 mRNA COVID-19 vaccine for preventing severe outcomes in Israel: an observational study. Lancet. 2021;398(10316):2093–100.
4. Tang P, Hasan MR, Chemaitelly H, Yassine HM, Benslimane FM, Al Khatib HA, et al. BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 Delta variant in Qatar. Nat Med. 2021;27(12):2136–43.
5. Vasileiou E, Simpson CR, Shi T, Kerr S, Agrawal U, Akbari A, 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(10285):1646–57.
6. Tessier E, Stowe J, Tsang C, Rai Y, Clarke E, Lakhani A, et al. Monitoring the COVID-19 immunisation programme through a National Immunisation Management System - England's experience. medRxiv; 2021:2021.09.14.21263578. Available from: https://www.medrxiv.org/content/https://doi.org/10.1101/2021.09.14.21263578v1. Cited 2025 Jan 7.
7. Nunes B, Rodrigues AP, Kislaya I, Cruz C, Peralta-Santos A, Lima J, et al. mRNA vaccine effectiveness against COVID-19-related hospitalisations and deaths in older adults: a cohort study based on data linkage of national health registries in Portugal, February to August 2021. Eurosurveillance. 2021;26(38):2100833.
8. Moustsen-Helms IR, Emborg HD, Nielsen J, Nielsen KF, Krause TG, Mølbak K, et al. Vaccine effectiveness after 1st and 2nd dose of the BNT162b2 mRNA Covid-19 Vaccine in long-term care facility residents and healthcare workers - a Danish cohort study. medRxiv; 2021:2021.03.08.21252200. Available from: https://www.medrxiv.org/content/https://doi.org/10.1101/2021.03.08.21252200v1. Cited 2025 Jan 7.
9. Fabiani M, Ramigni M, Gobbetto V, Mateo-Urdiales A, Pezzotti P, Piovesan C. Effectiveness of the Comirnaty (BNT162b2, BioNTech/Pfizer) vaccine in preventing SARS-CoV-2 infection among healthcare workers, Treviso province, Veneto region, Italy, 27 December 2020 to 24 March 2021. Euro Surveill. 2021. https://doi.org/10.2807/1560-7917.ES.2021.26.17.2100420.
10. Nordström P, Ballin M, Nordström A. Effectiveness of heterologous ChAdOx1 nCoV-19 and mRNA prime-boost vaccination against symptomatic Covid-19 infection in Sweden: A nationwide cohort study. Lancet Reg Health - Eur. 2021;11. Available from: https://www.thelancet.com/journals/lanepe/article/PIIS2666-77622100235-0/fulltext. Cited 2025 Jan 7.
11. Martínez-Baz I, Miqueleiz A, Casado I, Navascués A, Trobajo-Sanmartín C, Burgui C, et al. Effectiveness of COVID-19 vaccines in preventing SARS-CoV-2 infection and hospitalisation, Navarre, Spain, January to April 2021. Eurosurveillance. 2021;26(21):2100438.
12. de Gier B, Andeweg S, Joosten R, ter Schegget R, Smorenburg N, van de Kassteele J, et al. Vaccine effectiveness against SARS-CoV-2 transmission and infections among household and other close contacts of confirmed cases, the Netherlands, February to May 2021. Eurosurveillance. 2021;26(31):2100640.
13. Lanes S, Brown JS, Haynes K, Pollack MF, Walker AM. Identifying health outcomes in healthcare databases. Pharmacoepidemiol Drug Saf. 2015;24(10):1009–16.
14. Young JC, Conover MM, Funk MJ. Measurement error and misclassification in electronic medical records: methods to mitigate bias. Curr Epidemiol Rep. 2018;5(4):343–56.
15. Smedt TD, Merrall E, Macina D, Perez-Vilar S, Andrews N, Bollaerts K. Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. PLoS ONE. 2018;13(6):e0199180.
16. Bailie R, Bailie J, Chakraborty A, Swift K. Consistency of denominator data in electronic health records in Australian primary healthcare services: enhancing data quality. Aust J Prim Health. 2015;21(4):450–9.
17. Baum U, Kulathinal S, Auranen K. Exposure misclassification bias in the estimation of vaccine effectiveness. PLoS ONE. 2021;16(5):e0251622.
18 Sacco C, Manica M, Marziano V, Fabiani M, Mateo-Urdiales A, Guzzetta G, et al. The impact of underreported infections on vaccine effectiveness estimates derived from retrospective cohort studies. Int J Epidemiol. 2024;53(3):dyae077.

19. Groenwold RHH, Nelson DB, Nichol KL, Hoes AW, Hak E. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research. Int J Epidemiol. 2010;39(1):107–17.
20. Remschmidt C, Wichmann O, Harder T. Frequency and impact of confounding by indication and healthy vaccinee bias in observational studies assessing influenza vaccine effectiveness: a systematic review. BMC Infect Dis. 2015;15(1):429.
21. Doll MK, Pettigrew SM, Ma J, Verma A. Effects of Confounding Bias in COVID-19 and Influenza Vaccine Effectiveness Test-Negative Designs Due to Correlated Influenza and COVID-19 Vaccination Behaviors. medRxiv; 2021:2021.10.22.21265390. Available from: https://www.medrxiv.org/content/10.1101/2021.10.22.21265390v1. Cited 2024 Dec 12.
22. Loiacono MM, Van Aalst R, Pokutnaya D, Mahmud SM, Nealon J. Methods to account for measured and unmeasured confounders in influenza relative vaccine effectiveness studies: a brief review of the literature. Influenza Other Respir Viruses. 2022;16(5):846–50.
23. Gianfrancesco MA, Goldstein ND. A narrative review on the validity of electronic health record-based research in epidemiology. BMC Med Res Methodol. 2021;27(21):234.
24. Farmer R, Mathur R, Bhaskaran K, Eastwood SV, Chaturvedi N, Smeeth L. Promises and pitfalls of electronic health record analysis. Diabetologia. 2018;61(6):1241–8.
25. Protocol for a COVID-19 vaccine effectiveness estimation using health data registries, VEBIS multi-country study - Version 2.0. 2024. Available from: https://www.ecdc.europa.eu/en/publications-data/protocol-covid-19-vaccine-effectiveness-estimation-using-health-data-registries. Cited 2024 Dec 12.
26. Monge S, Humphreys J, Nicolay N, Braeye T, Van Evercooren I, Holm Hansen C, et al. Effectiveness of XBB.1.5 monovalent COVID-19 vaccines during a period of XBB.1.5 dominance in EU/EEA countries, october to november 2023: a VEBIS-EHR network study. Influenza Other Respir Viruses. 2024;18(4):e13292.
27. Kislaya I, Sentís A, Starrfelt J, Nunes B, Martínez-Baz I, Nielsen KF, et al. Monitoring COVID-19 vaccine effectiveness against COVID-19 hospitalisation and death using electronic health registries in ≥65 years old population in six European countries, October 2021 to November 2022. Influenza Other Respir Viruses. 2023;17(11):e13195.
28. Fontán-Vela M, Kissling E, Nicolay N, Braeye T, Van Evercooren I, Holm Hansen C, et al. Relative vaccine effectiveness against COVID-19 hospitalisation in persons aged ≥ 65 years: results from a VEBIS network, Europe, October 2021 to July 2023. Euro Surveill. 2024;29(1):2300670.
29 Nunes B, Humphreys J, Nicolay N, Braeye T, Van Evercooren I, Holm Hansen C, et al. Monovalent XBB.1.5 COVID-19 vaccine effectiveness against hospitalisations and deaths during the Omicron BA.2.86/JN.1 period among older adults in seven European countries: A VEBIS-EHR network study. Expert Rev Vaccines. 2024;23(1):1085–90.
30. Soares P, Machado A, Nicolay N, Monge S, Sacco C, Hansen CH, et al. COVID-19 vaccine effectiveness in the paediatric population aged 5–17 years: a multicentre cohort study using electronic health registries in six European countries, 2021 to 2022. Eurosurveillance. 2025;30(8):2400450.
31. Chapter 10: Analysing data and undertaking meta-analyses. Available from: https://training.cochrane.org/handbook/current/chapter-10. Cited 2025 Mar 3.
32. European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials: Step 5. European Medicines Agency, Committee for Medicinal Products for Human Use. 2020.
33. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 2010;21(3):383.
34. Chow JY, Goh ZJ, Li R, Lim DZY, Wee LE, Lye DCB, Tan KB, Lim JT. Approaches for estimating COVID-19 vaccine effectiveness using observational data in administrative databases: a systematic review. medRxiv. 2025 Sep 30;2025.09.29.25336864. https://doi.org/10.1101/2025.09.29.25336864.
35. Hak E, Verheij TJM, Grobbee DE, Nichol KL, Hoes AW. Confounding by indication in non-experimental evaluation of vaccine effectiveness: the example of prevention of influenza complications. J Epidemiol Community Health. 2002;56(12):951–5.
36. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. Int J Epidemiol. 2006;35(2):337–44.
37. Simonsen L, Taylor RJ, Viboud C, Miller MA, Jackson LA. Mortality benefits of influenza vaccination in elderly people: an ongoing controversy. Lancet Infect Dis. 2007;7(10):658–66.
38. Jackson ML, Yu O, Nelson JC, Naleway A, Belongia EA, Baxter R, et al. Further Evidence for Bias in Observational Studies of Influenza Vaccine Effectiveness: The 2009 Influenza A(H1N1) Pandemic. Am J Epidemiol. 2013;178(8):1327–36.
39. Castilla J, Guevara M, Martínez-Baz I, Ezpeleta C, Delfrade J, Irisarri F, et al. Enhanced estimates of the influenza vaccination effect in preventing mortality: a prospective cohort study. Medicine (Baltimore). 2015;94(30):e1240.
40. Chemaitelly H, Ayoub HH, Coyle P, Tang P, Hasan MR, Yassine HM, et al. Assessing Healthy Vaccinee Effect in COVID-19 Vaccine Effectiveness Studies: A National Cohort Study in Qatar. medRxiv; 2024:2024.07.28.24311115. Available from: https://www.medrxiv.org/content/10.1101/2024.07.28.24311115v2. Cited 2025 Jan 8.
41. Hansen CH, Moustsen-Helms IR, Rasmussen M, Søborg B, Ullum H, Valentiner-Branth P. Short-term effectiveness of the XBB.1.5 updated COVID-19 vaccine against hospitalisation in Denmark: a national cohort study. Lancet Infect Dis. 2024;24(2):e73–4.
42. Poukka E, Auranen K, Baum U. Non-COVID-19 hospitalisation as a negative control outcome in COVID-19 vaccine effectiveness studies. Lancet Infect Dis. 2024;24(5):e275.
43. Hansen CH, Moustsen-Helms IR, Rasmussen M, Søborg B, Ullum H, Valentiner-Branth P. Non-COVID-19 hospitalisation as a negative control outcome in COVID-19 vaccine effectiveness studies - Authors' reply. Lancet Infect Dis. 2024;24(5):e276.
44. European Health Data Space - European Commission. 2024. Available from: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en. Cited 2024 Oct 31.
45. Framework for financial data access - European Commission. Available from: https://finance.ec.europa.eu/digital-finance/framework-financial-data-access_en. Cited 2025 Mar 5.
46. Pizzato M, Gerli AG, La Vecchia C, Alicandro G. Impact of COVID-19 on total excess mortality and geographic disparities in Europe, 2020–2023: a spatio-temporal analysis. Lancet Reg Health. 2024;44:100996. https://doi.org/10.1016/j.lanepe.2024.100996.
47. Izurieta HS, Wu X, Lu Y, Chillarige Y, Wenecke M, Lindaas A, et al. Zostavax vaccine effectiveness among US elderly using real-world evidence: addressing unmeasured confounders by multiple imputation after linking beneficiary surveys with medicare claims. Pharmacoepidemiol Drug Saf. 2019;28:993–1001. https://doi.org/10.1002/pds.4801.
48. Izurieta HS, Lu M, Kelman J, Lu Y, Lindaas A, Loc J, et al. Comparative effectiveness of influenza vaccines among US medicare beneficiaries ages 65 years and older during the 2019–2020 season. Clin Infect Dis. 2021;73:e4251–9. https://doi.org/10.1093/cid/ciaa1727.
49. Double Negative Control Inference in Test-Negative Design Studies of Vaccine Effectiveness - PMC. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC8963685/. Cited 2025 Apr 30.
50. Martens EP, Pestman WR, de Boer A, Belitser SV, Klungel OH. Instrumental variables: application and limitations. Epidemiol Camb Mass. 2006;17(3):260–7.
51. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiol Drug Saf. 2010;19(6):537–54.
52. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. Biometrics. 2005;61(4):962–73.
53. Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly Robust Estimation of Causal Effects. Am J Epidemiol. 2011;173(7):761–7.
54. Self-controlled case series methods: an alternative to standard epidemiological study designs. The BMJ. Available from: https://www.bmj.com/content/354/bmj.i4515. Cited 2025 Mar 5.

Publisher's Note