Covid Vaccines Effectiveness (CoVE)

Effectiveness of heterologous and booster COVID-19 vaccination in 5 European countries, using a cohort approach in children and adults with a full primary COVID-19 vaccination regimen

Specific Contract No 01 implementing framework contract No EMA/2020/46/TDA/L5.06.

EU PAS Register No: EUPAS47725

Deliverable: D3 - Study report

Release date: Month 7 – 20 January 2023

Document version: 0.2

Disclaimer & acknowledgement

The research leading to these results was conducted as part of the activities of the EU PE&PV (Pharmacoepidemiology and Pharmacovigilance) Research Network (led by Utrecht University) with collaboration from the Vaccine Monitoring Collaboration for Europe network (VAC4EU).

This study was funded by the European Medicines Agency. The views expressed in this article are the personal views of the author(s) and may not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties. The project has received support from the European Medicines Agency under the Framework service contract nr EMA/2020/46/TDA/L5.06.

EU PE&PV research network

Document Information

Title	Effectiveness of heterologous and booster Covid-19 vaccination in 5 European countries, using a cohort approach in children and adults with a full primary Covid-19 vaccination regimen
Protocol version identifier	0.2
Date of last version of protocol	June 15, 2022
EU PAS register number	
Active substance	*NA*
Medicinal product	*NA*
Product reference	*NA*
Procedure number	*NA*
Marketing authorisation holder(s)	*NA*
Research question and objectives	The goal of this study is to assess the effectiveness and waning of immunity of primary Covid-19 vaccinations and the booster in preventing different Covid-19 outcomes.
Country(-ies) of study	Italy, Spain, the Netherlands and the United Kingdom
Main Authors of the study	Dr. Elisa Martin, AEMPS Dr. Fabio Riefolo, Teamit Institute Dr. Tiago Vaz, UMCU Dr. Rosa Gini, ARS
Contributors	Prof. Dr. Lamiae Grimaldi, APHP Dr. Satu Johanna Siiskonen, UU ROC12 Consortium

Table of Contents

1. Title

Effectiveness of heterologous and booster COVID-19 vaccination in 5 European countries, using a cohort approach in children and adults with a full primary COVID-19 vaccination regimen

2. Abstract

2.1. Title

Effectiveness of heterologous and booster COVID-19 vaccination in 5 European countries, using a cohort approach in children and adults with a full primary COVID-19 vaccination regimen

Date of the Abstract: 20th of January 2023

Main Author: Elisa Martin Merino (Agencia Española de Medicamentos y Productos Sanitarios-AEMPS)

2.2. Keywords

COVID-19; vaccines effectiveness; heterologous and homologous vaccine schedule; booster; real-word data.

2.3. Rationale and background

Real-world effectiveness data demonstrated that COVID-19 vaccines' protection against severe SARS-CoV-2 infection is high in the short term but wanes over time, also depending on the virus variants. The vaccine effectiveness (VE) can vary depending on the vaccine brand or type (e.g., among mRNA or adenoviral platforms). Mixing brands for the primary vaccination and/or boosters (heterologous vaccination schedules) has been applied in different countries or regions although the effectiveness of heterologous schedules was not fully understood beyond immunogenic clinical data. With its increased ability to elude immunity and cause reinfections, the SARS-CoV-2 Omicron variant became dominant worldwide and led to the highest ever COVID-19 incidence, also in countries with high vaccination coverage, increasing as well hospitalization and severe outcomes cases in paediatrics populations, particularly in the presence of comorbidities. Moreover, only limited real-life data information on VE for children and adolescents in the EU is available. Further evidence about the VE of homologous (use of the same COVID-19 vaccine for the primary vaccination and booster dose) and heterologous (use of different COVID-19 vaccines for the primary vaccination or booster dose) vaccination schedules are needed both in adult ad paediatrics populations to keep fueling regulatory authorities' preparedness in case of urgent decision-making situations.

2.4. Research questions and objectives

To investigate the VE and waning of immunity of diverse COVID-19 primary vaccination (1st and 2nd doses) and booster (3rd dose) schedules with Comirnaty (PF), Spikevax (MD), and Vaxzevria (AZ) vaccines in preventing different COVID-19-related disease outcomes.

Primary objectives

1) To estimate the VE, and its waning, in adults (>17 years old) and adolescents (12-17 years old), separately, between heterologous and homologous primary vaccinations.
2) To estimate the VE, and its waning, in children (5-14 years old) between homologous primary vaccinations and non-vaccination.
3) To estimate the VE, and its waning, in adults and adolescents with full homologous primary regimen between those with a homologous booster and heterologous booster, separately, compared to those without any booster.
4) To estimate the VE, and its waning, in adults and adolescents with full heterologous primary regimen between those with any booster and those without any booster.

For patients free of prior COVID-19 infection (all analysis), that VE was estimated:

- By vaccine brand of the primary homologous scheme, the combinations in the heterologous scheme, and booster dose (3rd dose)
- By age categories.
- By time since a complete primary vaccination regimen (2nd dose receipt) or booster among the compared groups.

- Among clinical subgroups associated with a high risk of severe COVID-19 (immunocompromised patients and patients with cancer, transplants, severe renal disease, and Down syndrome).

For patients with prior COVID-19, the overall VE of different vaccination schemes against severe COVID-19 and COVID-19-related death was estimated.

Secondary objective

To estimate the VE against all-cause mortality in ≥60 years old adults with a full primary regimen (homologous or heterologous) between those with any booster and those without any booster. This estimation complements the results of COVID-19-related death of the primary objective.

2.5. Study design

Herein, we present a retrospective cohort study to estimate the VE of different COVID-19 vaccines schemes, and their waning, using different SARS-CoV-2 infection-related outcomes: (i) non-severe COVID-19, (ii) severe COVID-19, and (ii) COVID-19 with death. The study used data from 6 European different data sources and focused on the period ranging from the beginning of the vaccination campaign (December 2020) to the last data available from the participating data sources (ranging from December 2021 to February 2022). Thus, it mainly covered the Delta-Omicron predominant SARS-CoV-2 virus variant periods within the full vaccination regimen (first vaccination scheme and booster doses).

2.6. Setting

We retrospectively used data from 6 electronic health care databases in Southern, Northern, and Western Europe: the Italian Caserta local health database (IT-INSPIRE srl), the Italian Societa Servizi Informatici (IT-PEDIANET) database, the Spanish Pharmacoepidemiological Research Database for Public Health System (ES-BIFAP), the Spanish Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària (ES-SIDIAP) database, the Dutch PHARMO Database Network (NL-PHARMO), and the British Clinical Practice Research Datalink (UK-CPRD) Aurum.

Participants were matched 1:1 on the calendar date of the vaccination of interest (2nd vaccination-time0, for the primary vaccination scheme analysis; 3rd vaccination-booster-time0, for the booster vaccination analysis), the calendar date of the 1st vaccine dose, the vaccine brand of the 1st dose (Comirnaty, Spikevax, Vaxzervria), vaccinees age, sex, geographical region, clinical subgroup, and SARS-CoV-2 infection prior 1st vaccination dose.

For primary vaccination schedules, the date of cohort entry (time0) was the date when the participant received the 2nd dose. For the booster vaccinations, the date of cohort entry (booster-time0) was the date when the participant received the 3rd dose. For the non-boosted comparators, the same calendar date as the corresponding boosted-matched individual (booster-time0) was used for comparison.

Children and pre-adolescents were defined as "not vaccinated" until the date of the receipt of the 1st COVID-19 vaccine dose, thus, potentially selected as unvaccinated control.

Participants were considered to have a complete primary vaccination regimen when the record of a 2nd COVID-19 vaccine dose existed after more than 19 days from the 1st dose. Individuals were defined as "boosted" (homologous or heterologous) from the date of the 3rd COVID-19 vaccine dose receipt if at least 28 days after the 2nd one. Participants were defined as "non-boosted" until the date of 3rd vaccine dose administration, thus, potentially selected as non-boosted control. Among individuals with complete vaccination schemes, both homologous/heterologous primary vaccinations and booster/non-boosted cohorts were identified separately.

Statistical data analyses

Inverse probability weighted (IPW) Cox models (CI, 95%) have been used to derive the average hazard ratio (HR) of COVID-19-related outcomes. The adjusted VE (%) values were estimated as 1 minus the adjusted HR multiplied by 100. Various covariates have been considered potential confounders for the IPW.

The VE for each matched cohort was estimated by (*i*) vaccine brands, (*ii*) time after vaccination, (*iii*) age categories, (*iv*) high-risk of severe COVID-19 clinical subpopulations (for the participant

databases able to identify them through diagnosis or medications prescriptions, (*v*) and SARS-CoV-2 infection before vaccination, and (*vi*) dominant SARS-CoV-2 variants (defined as the variant reaching 50% of the total sequenced specimens at (booster-)time 0). Three main SARS-CoV-2 variants' periods have been identified (pre-Delta, Delta, and Omicrons) specifically for each participating country.

Random-effects meta-analyses, using the main estimates against severe COVID-19 outcomes from each data source, were performed for clinical subgroups for both adult and children populations. Sensitivity analysis restricting to patients with prior SARS-CoV-2 negative tests was performed to control for surveillance bias.

2.7. Subjects and study size

Study Population

The source population comprises all children, adolescents, and adults registered in any of the data sources during the study period (December 2020 -February 2022 for ES-BIFAP and IT-INSPIRE and -December 2021 for the other data sources). Eligible study participants had at least 2 years of available healthcare data. The vaccinated study population includes all individuals with at least two recorded vaccinations since the start of the study period.

Study size

The study cohorts contained more than 20 million adolescents and adults with a complete primary vaccination scheme (1st and 2nd dose receipt) and 3 to 6 months of follow-up across all the participating data sources. We could match ≈24-51% of the population for the heterologous vaccination scheme (comparator) and ≈0.5-1.5% for the homologous scheme. The majority of adults and adolescents were free of SARS-CoV-2 infection prior to vaccination (58-95% across the total matched population for all the data sources).

During the study period, approximately 308,000 children with around 3 months of follow-up were vaccinated with two doses and included in the study across all the participating data sources. Based on the matching criteria, a total of 295,573 children were matched for the primary vaccination schedule (95% of the total children). Of the matched children, the main population (97%, 287,050 children) did not have encountered prior SARS-CoV-2 infection.

2.8. Variables and data sources

Data sources captured vaccination and outcomes from hospitalization and/or general practice and/or test registries. This study considered different outcomes related to COVID-19: non-severe SARS-CoV-2 infection, hospitalized COVID-19 (severe), COVID-19 with death, and all-cause mortality (secondary objective). The main exposure of interest was the receipt of a different primary regimen or booster COVID-19 vaccine (receiving the same 1st dose brand), the dose, and its brand. Selection of covariates, primary care physician' visits, clinical conditions, and medication use (including influenza vaccination among others), were collected up to 2 years before (booster-)time 0 (or 7 days before, for visits to control by healthy vaccinees effect) and considered as potential confounders for the IPW.

2.9. Results

The shown VE percentages are statistically significant and reported in ranges of values across the data sources, otherwise, the mention of non-statistically significance is specified.

Primary Objectives

Adults, Primary Vaccination

Among 89,528 adults' matched pairs, overall, homologous primary vaccinations showed a slightly decreased VE (-27% to –36% by data source) compared to heterologous regimens against **non-severe COVID-19**, mostly receiving AZ as the first dose (VE ranged from -43% to –27% across data sources). By time since vaccination, or, by age categories, no clear patterns were found. In ES-BIFAP, the lower VE with homologous was more marked during the Delta (-39%) than the Omicron (–24%) periods. No differences between homologous and heterologous regimens were observed for **severe COVID-19** (estimated in Spanish data sources) and no sufficient cases of **death with COVID-19** were found to analyse.

Adolescents, Primary Vaccination

We matched 1,329 pairs among adolescents. Considering non-severe COVID-19, no differences were found in the VE estimates comparing homologous versus heterologous primary vaccinations. The small sample size hampered the VE estimation related to the other severe outcomes.

Children, Primary Vaccination

287,000 children without prior COVID-19 were matched. Considering **non-severe COVID-19**, homologous primary two doses of both the mRNA vaccines (PF or MD) showed VE, varying from 29% to 77% across data sources, during the Delta predominant variant period, when compared to unvaccinated individuals. VE remained 4-5 months. During the Omicron predominance, VE decreased from 77% to 42% in IT-INSPIRE and reverted to an increased risk of non-severe SARS-CoV-2 infection, from 29 to -44%, in ES-BIFAP. Estimates did not show protection among children with prior SARS-CoV-2 infection. The protection against **severe COVID-19** was >90% in ES-SIDIAP (during the Delta period) and ≈50% in ES-BIFAP for PF vaccine versus unvaccinated individuals. In ES-BIFAP, VE during the Delta period was 61%, but non-statistically significant, and 50% during the Omicron period. No data for waning of immunity, from other data sources, vaccine brands and COVID-19 with death were available.

Adults, Booster Vaccination

5.6 million adults without prior SARS-CoV-2 infection were matched. 79,076 cases of **non-severe COVID-19** among boosted versus 138,638 among unboosted adults were captured in 5 data sources. VE against non-severe COVID-19 ranged 31-69% for homologous boosters and 42-70% for heterologous boosters across data sources, independently from the vaccine brand. Considering **severe COVID-19**, 1,015 cases among boosted versus 3,362 among comparators were captured in 3 data sources with hospitalization information (mostly in ES-BIFAP and ES-SIDIAP and a few in IT-INSPIRE). Against severe COVID-19, heterologous boosters (homologous doses 1 and 2), independently from the vaccine brand, showed a VE of 73-81% across data sources whereas homologous boosters have a VE of 42-67%, compared to their respective unboosted controls. Considering **death with COVID-19**, 313 cases of death with COVID-19 among adults who received any booster versus 1,367 among comparators were captured (mostly in ES-BIFAP and ES-SIDIAP and a few in UK-CPRD). Protection against death with COVID-19 was similar among homologous and heterologous schemes (70-88% across schemes and data sources). Duration of immunization varied from 1 to 6 months across data sources and events, independently of the booster schedule. Considering both, severe COVID-19 and death outcomes, no clear VE differences were identified during both the Delta and Omicron periods across data sources. In patients with **cancer**, effectiveness against severe COVID-19 ranged from 54% to 77% with heterologous and from 49% to 61% with homologous boosters across data sources, whereas in patients with **immunodeficiency** VE was between 60-78% with any scheme. Among patients with prior SARS-Cov-2 infection, VE against severe COVID-19 was 69% (ES-BIFAP) and 71% (ES-SIDIAP) for heterologous and 43% (ES-SIDIAP) for homologous 3 doses schemes. VE against death with COVID-19 was 92% (ES-BIFAP) for heterologous and 68% (ES-SIDIAP) for homologous boosters in these patients.

Adolescents, Booster Vaccination

We matched 17,652 adolescent pairs of which 408 cases of **non-severe COVID-19** among boosted versus 936 cases among unboosted individuals were captured in 5 data sources (ES-BIFAP, ES-SIDIAP, IT-INSPIRE, UK-CPRD and NL-PHARMO). Among adolescents with a homologous primary vaccination, the VE of homologous booster doses against non-severe COVID-19 varied from 35-67% across vaccine brands and data sources, whereas VE of heterologous boosters was 48% in ES-BIFAP (the only data source in which heterologous boosters were found) for PF as 1st and 2nd dose, and MD as 3rd dose, when compared to the respective unboosted controls. During the Delta predominance period, VE was only observed in Italy (69%) for homologous boosters. During the Omicron predominance, VE varied from 67% (IT-INSPIRE) to 44% (ES-BIFAP) for homologous boosters whereas, for the heterologous ones, was 51% (ES-BIFAP). VE for the homologous boosters lasted up to one month in Italy (75%; later on, <5 cases occurred) and two months in Spain (45%). Heterologous primary vaccinations (in UK-CPRD) were not sufficient to be analysed. No VE was estimated for **severe COVID-19 and death with COVID-19** outcomes due to the low number of cases (<5).

Sensitivity Analysis

Balancing by any prior testing, the VE against **non-severe** infection remained 57-59% for primary vaccination among children and 30-55% for boosters in adults. Against **severe** COVID-19, VE remained moderate (55-59%) for homologous boosters and high (70-81%) for heterologous boosters. Against **death with COVID-19,** VE was 67-79% for homologous and 77-81% for heterologous boosters among adults.

Meta Analysis

In adults, the pooled VE of **homologous boosters against severe COVID-19** was 62% (95% CI: 57 to 67%; I^2=0%) among subjects with immunodeficiency, 54% (95% CI: 41 to 64%; I^2=18%) among patients with cancer, 24% (95% CI: -54 to 63%; I^2=0%) among patients with a transplant and 57% (95% CI: -20 to 84%; I^2=65%) among those with severe renal disease. Additionally, the pooled VE of **homologous booster against death with COVID-19** was 73% (95% CI: 63 to 80%; I^2=15%) among immunocompromised patients, 75% (95% CI: 65 to 82%; I^2=0%) among patients with cancer, and 75% (95% CI: -38 to 96%; I^2=63%) for those with severe renal disease.

The pooled VE of **heterologous boosters against severe COVID-19** (homologous primary vaccination) was 72% (95% CI: 66 to 77%; I^2 =0%) among adults with immunodeficiency and 68% (95% CI: 36 to 84%; I^2 =77%) for adults with cancer. In addition, the pooled VE of **heterologous boosters against death with COVID-19** was 80% (95% CI: 70 to 86%; I^2 =0%) among immunocompromised patients and 81% (CI=95% CI: 70 to 89%; I^2 =0%) among those with cancer.

Among children, the pooled VE of **primary vaccination against severe COVID-19** in the Delta predominance period, was 82% (95% CI: -10 to 97%; I^2=62%) in Spain.

Secondary Objective (any cause of death)

In patients aged ≥60 years old, the VE against any cause of death for any booster dose (whether homologous or heterologous), following a homologous primary vaccination schedule, ranged between 72% and 96% across 10-by-10 age categories and data sources (ES, UK and NL). VE was higher during the Delta (75% for homologous and 83% for heterologous boosters) than the Omicron variant period (67% for both booster types). A few heterologous primary vaccinations were used, showing effectiveness (74%) only in 80+ years old in UK.

2.10. 1.1. Discussion

Using real-world data from 6 different databases in Europe, various schemes of COVID-19 vaccination (primary vaccination and booster doses) were identified among adults, adolescents, and children (only in 3 countries) in Spain, Italy, the Netherlands, and the UK from the beginning of the vaccination campaign to December 2021 or February 2022 (in Italy and Spain).

Most of the **adults** participating in our study were fully vaccinated from mid of 2021 and received booster doses in the last months of 2021. Under the scenario of a reported very high level of protection with the homologous COVID-19 vaccination during Delta predominance in the EU/EEA countries, in the current project, slightly higher effectiveness of heterologous versus homologous primary vaccinations against **non-severe** COVID-19 was observed in patients who received an AZ first dose and a 2nd mRNA dose in comparison to those receiving two AZ doses. VE was not different among people who mixed mRNA doses. This is in accordance with previous clinical trial studies and supporting the public health recommendation to switch to mRNA platforms after AZ vaccines in some countries. Against **severe COVID-19**, VE was not different between heterologous and homologous primary vaccinations administrated in the same calendar moments, and not enough cases of COVID-19 with **death** occurred for VE estimation. No clear diverse patterns of the waning of immunity for homologous versus heterologous primary schemes, or changes in the comparative effectiveness among age groups or variant predominance periods were observed. This is also due to the small sample size that contributed to this primary analysis.

The protection with a **booster** against **non-severe COVID-19** showed **important variations** (28-74%) by brand and/or country. Effectiveness was high for AZ boosters in the UK, medium for PF or MD in Italy, the UK, and The Netherlands, and very low for all the brands in Spain (that did not improve after balancing testing in sensitivity analyses). Variations can result from different clinical characteristics of the populations at the analysed moment: the regions (with different virus prevalence, predominance, and potential public health recommendation, health assistance or people habits, etc.) among others. **mRNA-boosted** adults with PF and MD showed high effectiveness **against severe COVID-19 and death with COVID-19** regardless of the brand of the two doses administrated as primary vaccination. Those results were more solid in Spanish cohorts probably due to the higher use of those mixed schemes. The booster doses were observed effective during both Delta (Spain and UK) and Omicron (Spain) periods supporting the public health recommendations for subsequent boosters during the Omicron era. VE remains apparently longer for the homologous booster (until the 2nd or 5th month) than the heterologous (until the 1st or 3rd month) that started to be administrated approximately one month later. This reduction of the VE could be triggered by both the waning of immunity and lower effectiveness against the more aggressive Omicron variant. The effectiveness against death with COVID-19 remained until enough number of cases were detected, with a maximum duration of 5-6 months in a Spanish data source. Although the reason to die during the study period was unknown and many could have died by other reason than COVID-19, all-cause mortality analyses' results complemented the effectiveness of booster doses found against death with COVID-19 that would include missing cases without SARS-COV2 test in the main analysis.

A good sample size of **patients with immunosuppression, cancer, and severe renal disease** was achieved among those who received homologous primary schemes. According to pooled estimates, patients with immunocompromised or cancer benefited substantially from a booster (homologous or heterologous) considering both the severe outcomes (VE ranged from 54% to 75%, low or no heterogeneity). Also, high VE of homologous boosters was observed against severe outcomes in adult patients with severe renal disease in one (ES-BIFAP) out of three data sources tested. These findings add values for understanding the vaccination outcomes in those clinical subgroups that are less represented in randomized control trials (RCT) and small studies.

Most of the **adolescent** individuals completed the primary vaccination during the Autumn of 2021 with homologous mRNA vaccines, the only approved for such young populations. Thus, the matched heterologous cohorts were insufficient to estimate compared VE. Then, three doses of PF showed moderate VE against non-severe COVID-19 in Italy (VE: 67%) both during Delta and Omicron, as well as three homologous Moderna doses in Spain (VE: 64%) in comparison with similar two doses without a booster. In Spain, the observed moderate effectiveness was only significant during the Omicron period, where Moderna could have been more used. Early boosted adolescents at high risk of infection or related comorbidities, inferior infection notification rates during the last months studied, frequent patients with prior infection along the time and/or more underlined prior infections in the control groups could be playing a crucial role in such reduced estimates. Also, the VE of a 3rd homologous dose was moderate among Italian adolescents who were immunocompromised, adding value to the vaccination recommendation at that moment. No severe COVID-19 or death with COVID-19 cases occurred for permitting the VE estimation among adolescents.

For **children**, data from Spain and Italy showed that most of them completed their primary vaccination scheme in February 2022. Among them, varying effectiveness was shown (from low to high) against non-severe SARS-CoV-2 infection during the **Delta variant** predominance period. **During the Omicron period**, VE decreased in Italy and Spain (the last with an even higher risk of infection among vaccinated children as consequence of differential surveillance/testing between vaccinated and unvaccinated children as observed in the sensitivity analysis). Overall, effectiveness decreased in magnitude over time. The increased incidence of the infection in all age groups from December 2021, the consequence of the onset of the Omicron variant, the relaxation in Public Health measures as it was considered a less severe variant, the potential underestimation of the infection incidence due to non-reported positive self-COVID-19-tests and other factors aforementioned for adolescents, could have also triggered the reduced VE among the youngest. A complete primary vaccination regimen offered a moderate-high level of protection for **severe COVID-19** in children (12-14 years old, as no cases occurred in those 5-11) in Spain. The database producing that moderate effectiveness showed differential COVID-19 severity misclassification in a manual review independent of the current project. Hospitalisations among non-vaccinated were more often confirmed due to COVID-19 (63%) than among vaccinated individuals (24%). This suggests that the

actual effectiveness would be higher than estimated. In the Omicron period, vaccination showed moderate protection against severe disease, as reported (49%; 95% CI: 26-64%) among people aged 12-17 years in February 2022 in Spain after suffering a decreasing trend from November 2021 (previously, up to October 2022, VE was kept ≥90%) (isciii report). There were no deaths related to COVID-19 disease among children, as in previous studies.

2.11. 1.2. Conclusions

The real-world evidence effectiveness of EU-approved PF, MD, and AZ COVID-19 vaccines and their combination in different schemes was assessed in this study during the Delta and the early Omicron circulating variants. Our results support health institutions' recommendation of mixing different vaccine brands for the second dose, particularly for those vaccinated with AZ first, as no threatening VE differences were found across homologous and heterologous primary vaccination schemes against severe and non-severe events. This is in line with previously small published clinical trials. Independently of the vaccination history, mRNA 3rd doses offered clear protection against severe post-infection events, both, in the general adult population and more vulnerable subjects, adding evidence for these subjects at higher risk and supporting previous observations in clinical studies for adults. We observed a wane of immunity from the early months post-vaccination. Our data infer that full primary vaccination protected 12-14-year-old children against hospitalized COVID-19 disease during the Delta and Omicron variant periods in Spain, including immunocompromised children. No clear conclusions are available for 5-11 years old vaccinees as no severe cases were encountered. Thus, broader and ad-hoc studies for children are further needed. Beyond the evidence of the COVID-19 vaccine herein reported, useful for the benefit-risk assessment, a cautious interpretation of these results should consider a broad reduction in infection notifications from the last months of 2021 due to self-testing at home, the vaccination campaign differences across countries and aforementioned limitations. These conclusions cannot fully address the current virus epidemiological situation with updated PF and MD vaccines with Omicron subvariants. Therefore, updated effectiveness estimations in all age groups on severe reinfections are required.

2.12. Marketing Authorisation Holder

N/A

2.13. Principal Investigators

Elisa Martin Merino Mar Martin Belen Castillo	Agencia Española de Medicamentos y Productos Sanitarios (AEMPS)
Fabio Riefolo	Teamit Institute, S.L.
Rosa Gini	Agenzia Regionale di Sanità (ARS)
Lamiae Grimaldi	l'Assistance Publique-Hôpitaux de Paris (APHP)
Tiago Vaz	University Medical Center Utrecht (UMCU), The Netherlands
Olaf Klungel Satu Siiskonen	Utrecht University, Pharmacoepidemiology & Clinical Pharmacology

3. List of abbreviations

AEMPS	Spanish Agency for Medicines and Medical Devices
BIFAP	Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria
CI	confidence interval
COVID-19	illness caused by the SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) virus
eHR	electronic health records
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU PAS Register	European Union Electronic Register of Post-Authorisation Studies
GPP	Good Pharmacoepidemiology Practices
GVP	Good Pharmacovigilance Practices
HCU	health care utilisation
HR	hazard ratio
ICPC-2	*International Classification of Primary Care, 2nd Edition*
ICD-9	*International Classification of Diseases, 9th Revision*
ICD-10	*International Classification of Diseases, 10th Revision*
ICU	intensive care unit
IRR	incidence rate ratio
ISPE	International Society for Pharmacoepidemiology
PASS	Post-authorisation safety study
RCT	Randomised clinical trial
RD	risk difference
RR	risk ratio
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
VE	Vaccine Effectiveness
PF	Pfizer Covid-19 vaccine
MD	Moderna Covid-19 vaccine
AZ	AstraZeneca Covid-19 vaccine
IR	Incidence Rates
CI	Confidence Interval

4. Investigators

All main responsible parties are presented below.

Responsible parties

Organisation	Key person	Role(s)
Agencia Española de Medicamentos y Productos Sanitarios (AEMPS)	Elisa Martin Merino, PI	Core Scientific team, WP1 lead
l'Assistance Publique-Hôpitaux de Paris (APHP)	Lamiae Grimaldi	Core Scientific team WP1 and Data Access Provider SNDS
University Medical Center Utrecht (UMCU), The Netherlands	Tiago Vaz	Core Scientific team, WP2 lead
Utrecht University, Pharmacoepidemiology & Clinical Pharmacology	1.Olaf Klungel 2.Satu Siiskonen	1.lead EU PE&PV research network 2.Consortium Management and liaison EMA
Teamit Institute, S.L.	Fabio Riefolo	Core Scientific team, Project Management
Agenzia Regionale di Sanità (ARS)	Rosa Gini	Core Scientific team

Key collaborators and roles

Organisation	Key person	Role(s)
Utrecht University, Pharmacoepidemiology & Clinical Pharmacology	Patrick Souverein Helga Gardarsdottir	Data Access Provider (ETL) & expertise
VAC4EU	1.Daniel Weibel, secretary general 2.Patrick Mahy, treasurer	1.Provide access and use of tools, coordination 2.subcontracting
Agencia Española de Medicamentos y Productos Sanitarios (AEMPS)	Mar Martin Patricia García Belen Castillo Miguel-Angel Macia	Data Access Provider BIFAP Epidemiologists Statistician
UMC Utrecht (UMCU).	1. Albert Royo 2. Vjola Hoxhaj 3. Judit Riera 4. Ivonne Martin 5. Dorieke Brink-Kwakke 6. Maria Schipper 7. Hae Won 8. Carlos Duran Salinas	1. programmer R 2. programmer & review data quality checks 3., 8. Review data quality checks 4., 5., 6.,7. Statistician
PHARMO Institute for Drug Outcomes Research.	Ronald Herings Karin Swart-Polinder Jetty Overbeek Jesse van den Berg	Data Access Provider PHARMO & expertise
Agenzia Regionale di Sanità (ARS)	Rosa Gini Davide Messina Olga Paoletti	Data Access Provider ARS & Programming expertise
Società Servizi Telematici - Pedianet.	Carlo Giaquinto Elisa Barbieri Luca Stona	Data Access Provider Pedianet, expertise
IDIAP JGol	Felipe Villalobos Carlo Alberto Bissacco	Data Access Provider SIDIAP, expertise
Academic Spin-Off "INSPIRE" srl	Gianluca Trifiro Ylenia Ingrasciotta Valentina Ientile Salvatore Crisafulli	Data Access Provider, expertise
Instituto Aragonés de Ciencias de la Salud.	Antonio Gimeno Beatriz Poblador Mercedes Aza Aida Moreno Jonás Carmona	Scientific support protocol, statistics, report
Navarre Health Service.	Juan Erviti Leire Leache Luis Carlos Saiz Marta Gutiérrez Valencia	Scientific support protocol, statistics, report

Organisation	Key person	Role(s)
Drug Safety Research Unit	Liz Lynn Dr. Debabrata Roy	Scientific support protocol, statistics, report
RESEARCH TRIANGLE INSTITUTE	Bradley Layton Xabier Garcia de Albeniz Susana Perez-Gutthann Rachel Weinrib	Scientific support protocol, statistics, report
Institute of Public Health, Riga Stradins University	Anda Kivite-Urtane	Scientific support protocol, statistics, report
Democritus University of Thrace	Christos Kontogiorgis Foteini Dermiki-Gkana	Scientific support protocol, statistics, report
National Public Health Agency (RIVM)	Susan Hahne	Scientific support protocol, statistics, report
Teamit Institute, S.L.	Eva Molero Gianmarco Di Mauro	Project Management

5. Milestones

Deliverable	Planned date of delivery	Actual date of delivery
D1 Study Plan	4 May 2022	4 May 2022
D2 Study Protocol	25 May 2022	31 May 2022
Registration in the EUPAS register		17 June 2022
Start of data collection	January 2021	
End of data collection	30 April 2022	
D3 Study report & annexes	20 September 2022	07 November 2021 (descripting data) 20 November 2022 (whole report) 20 January 2023 (re-submission after EMA review)
D4 Manuscript	20 September 2022	20 November 2022 20 January 2023 (re-submission after EMA review)

6. Rationale and Background

Real-world effectiveness data has demonstrated high levels of short-term protection by Covid-19 vaccines against clinical disease and, more so, against severe outcomes including hospitalization and death (1–9). Unfortunately, evidence shows that protection against symptomatic disease wanes over time and depends on circulating variants (1–9). Andrews et al. showed the effectiveness of booster vaccination in the UK until December 2021, just prior to the spread of Omicron (9) while Moge et al. (10) showed moderate effectiveness (51.3%) of mRNA vaccines boosters in comparison with no booster in preventing infection with the Omicron variant during the first month of 2022. Unfortunately, evidence shows that protection against symptomatic disease wanes over time and depends on circulating variants (1–9). Andrews et al. showed the effectiveness of booster vaccination in the UK until December 2021, just prior to the spread of Omicron (9) while Moge et al. showed moderate effectiveness (51.3%) of mRNA vaccines boosters in comparison with no booster in preventing infection with the omicron variant during the first month of 2022.

Booster/additional doses were implemented in many countries to combat the serious effects of the Delta variant and the highly infectious Omicron variant. As ECDC reports (10): "Results from observational studies show that the vaccines authorised in the EU/EEA are currently highly protective against COVID-19 related severe disease, hospitalisation and death caused by the Delta variant of concern. Although overall high effectiveness can vary, depending on the population groups (e.g., among elderly populations) and the vaccine.".

Following safety concerns related to thrombotic events, some European countries (e.g., Spain and Italy) recommended a second mRNA for people vaccinated with AZ in young adults < 60 (10) or minimize its use (i.e. heterologous scheme recommendation) after a phase II clinical trial (11) in Spain that showed robust immune response and mild reactogenicity of a heterologous vaccine regimen (ChAdOx1-S - BNT162b2) as compared with no 2nd vaccination (12). Apart from booster doses, it was also disseminated that mixing different vaccine platforms would raise advantages in terms of protection effectiveness against the new SARS-CoV-2 variants.

A literature study conducted by EMA and based on published studies suggested that the heterologous combination of mRNA and viral vector vaccines produced good levels of SARS-CoV-2 antibodies and a higher T-cell response compared to homologous vaccination, whether in a primary or booster regimen (13). Monge et al. (14) observed that effectiveness was significantly higher among patients boosted with Moderna in comparison with Pfizer and varied slightly by type of primary schedule. However, the use of a viral vector vaccine as the second dose in primary vaccination schemes, or the use of two different mRNA vaccines, was less well studied. Therefore, additional real-world evidence was needed on the effectiveness of heterologous and booster vaccination on a larger population scale.

There were also still insufficient real-life data on the effectiveness of the vaccines authorised in the EU against the Omicron variant in children and adolescents apart from adults and special population at higher risk of severe prognosis. Protection also appeared to wane over time. Giving an additional or booster COVID-19 vaccine dose following a full primary vaccination course was critical to ensure a higher level of protection, particularly in the face of emerging variants such as Omicron or subvariants.

Waning of immunity and appearance or new variants led to booster doses were authorized by the EMA and MHRA and rolled out throughout Europe. Table 1 shows the information on the uptake of first vaccination, full regimens, and additional doses per total population and brand administered. The use of adenovirus platforms for primary regimens has been undertaken with heterogeneity among European countries, and these differences in exposure (see Table 1) allowed us to investigate differences in effectiveness based on primary schemes and booster.

In July 2022, after the period studied in the current project, recommendations on additional booster doses of mRNA COVID-19 vaccines for people at high risk of severe outcomes were updated by EMA and ECDC. In September 2022, booster vaccinations with Omicron adapted Covid-19 vaccines were authorised (two bivalent Original/Omicron BA.1 and one Original/Omicron BA.4-5) (https://www.ema.europa.eu/en/documents/public-statement/ecdc-ema-statement-booster-vaccination-omicron-adapted-bivalent-covid-19-vaccines_.pdf) (15).

Table 1. Information on the uptake of COVID-19 vaccines in the participating countries and administered brands (as per ECDC Vaccine tracker 19-01-2022)

Country	First doses	Full primary regimen	Additional doses
Spain	85.4%	75.6%	37.4%
Children below 18 years	54.6%	27.7%	
Comirnaty	61,227,693		
Spikevax	18,931,042		
Vaxzevria	9,769,679		
Janssen	1,982,383		
Population size	47.4 million		
Italy	80.5%	74.3%	32.2%
Children below 18 years	41.2%	29.2%	
Comirnaty	75,357,550		
Spikevax	21,190,469		
Vaxzevria	11,975,044		
Janssen	1,478,485		
Population size 2022	60.3 million		
Netherlands	76.7%	70.6%	44.5%
Children below 18 years	24.5%	21.0%	
Comirnaty	22,646,011		
Spikevax	6,211,691		
Vaxzevria	2,800,808		
Janssen	867,708		
Population size 2022	17.2 million		
UK	76.2%	70.2%	53.5%
Children below 18 years			
Comirnaty	NA		
Spikevax	NA		
Vaxzevria	NA		
Janssen	NA		
Population size	68.4 million		

7. Research Question and objectives

7.1. Research Question

The aim of this study was to assess the effectiveness and waning of immunity of primary COVID-19 vaccinations and the booster in preventing different COVID-19 outcomes. Figure 1 provides an overview of the study design.

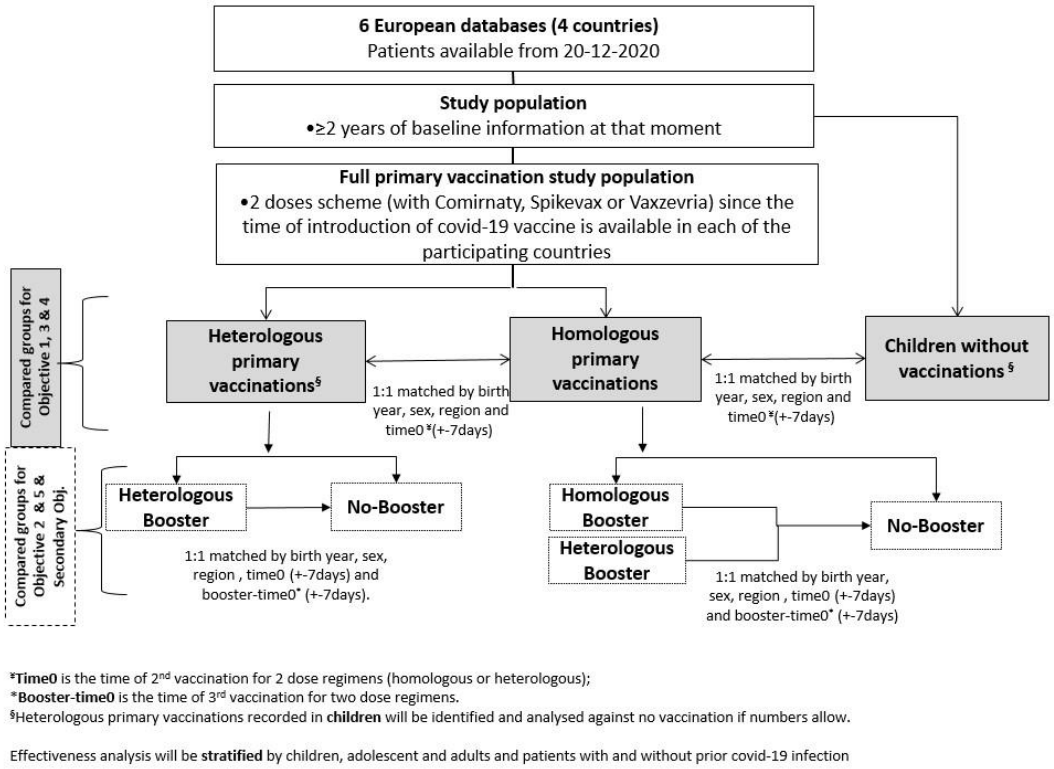

Figure 1. *Study flow chart.*

7.2. Primary Objectives

7.2.1. Primary objective 1 (adults and adolescents), 3 (children), and 4 (waning of immunity)

- To estimate the effectiveness and waning of effectiveness in adults and adolescents, separately, between heterologous and homologous primary vaccinations.
- To estimate the effectiveness and waning of effectiveness in children between homologous primary vaccinations and non-vaccination.

Up to March 2022, only Comirnaty vaccine was approved for children aged 5-11 years old. On 2 March 2022, the EU commission agreed to extend the marketing authorization of Moderna Covid-19 vaccine to be used in children aged 6 years and older. No heterologous schemes have been expected for the time of the study in this population according to the latest updates of health component authorities. Thus, the effectiveness of homologous vaccinations of those two vaccines have been estimated in comparison with non-vaccination among children. Therefore, the methodological approach for evaluating the effectiveness of COVID-19 vaccination in children differs from the one used in adults and adolescents. When heterologous vaccinations have been found among children, matched non vaccinated children are selected for comparisons.

7.2.2. Primary objective 2 (boosting), and exploratory objective 5 (waning of immunity after booster)

5) To estimate the effectiveness and waning of effectiveness in adults with full homologous primary regimen between those with a homologous booster and heterologous booster, separately, compared to those without any booster.

6) To estimate the effectiveness and waning of effectiveness in adults with full heterologous primary regimen between those with any booster and those without any booster.

On 24/02/2022, EMA recommended authorisation of booster doses of Comirnaty from 12 years of age. Therefore, even if these data are available in the participating databases, numbers are too small to evaluate booster doses in adolescents.

Vaccine effectiveness has been estimated:

- By vaccine brand of the primary homologous scheme, the combinations in the heterologous scheme and booster
- By age category (children aged 5-11 years; adolescents aged 12-17 years; adults aged 18-29; 30-49; 50-69; 70-79; ≥80; separate groups of adults were tested for heterogeneity).
- By time since a complete primary vaccination regimen or booster among the compared groups.
- Among clinically meaningful subgroups such as associated with a high risk of severe COVID-19 (patients with immunocompromise, cancer, transplants, severe renal disease, and Down syndrome) among DAPs allowed to identify them (through diagnosis or medication).
- For patients free of previous COVID-19 infection (all analysis) and for patients with previous COVID-19 infection (analysis for severe COVID-19 and COVID-related death), which have been matched by time since previous COVID-19 infection (i.e., within 2 months).

7.3. Secondary Objective (effectiveness of booster against all-cause mortality)

Information on the infection status of some patients may be missing in the electronic health records before death (12), due for instance to the periods of disrupted or overwhelmed health system, results of self-diagnosis test not recorded, or among patients not being attended in healthcare centers. Thus, secondary objective was to estimate the effectiveness against all-cause mortality in adults aged 60+ with a full primary regimen (whether homologous or heterologous) between those with any booster and those without any booster, to complement the results of Covid-19 with death in the primary objective.

Analysis of all-cause mortality were stratified by the following age groups (60-69; 70-79; ≥80). All-cause mortality analyses have been performed to incorporate additional informative details for the significant number of events encountered in age groups below 60 years old (<60) in this study.

8. Amendments and updates

Number	Date	Section of study protocol	Amendment or update	Reason
#001	*07/11/2022*	*D3 Study report & annexes* *D4 Manuscript*	*Amended delivery data*	*Postponed delivery date to 07/11/2022*

9. Research Methods

To meet the proposed goal and objectives, we have capitalized on the experience of the EU PE&PV and VAC4EU research network in the creation of readiness, governance, processes, data, people, methods, and tools developed in the ACCESS, Early-Covid-Vaccine-Monitor, and Covid-Vaccine-Monitor studies.

9.1. Study Design

The study design was a retrospective multi-database cohort study that complements the published and ongoing test negative case control studies by other groups. We leveraged the ability to follow exposure cohorts and the size of the populations, which is important because some of the heterologous primary schemes were low in frequency. Moreover, in many data sources, negative tests were not recorded. A restriction to people with negative tests was conducted as sensitivity analysis in BIFAP, SIDIAP and PEDIANET, where negative covid tests were available.

We used restriction and maximized matching to ensure comparability, adjusting for additional confounders to ensure the removal of confounding.

Matching on vaccination date of the second dose (i.e., time0, for the primary vaccination analysis) and on the date of the booster (i.e., booster-time0, for the booster analysis) plus region was crucial for controlling for confounding by calendar time and circulating strains as well as by vaccination prioritization (and consequently by risk of infection and prognosis).

Matching on birth year deals with the age-related roll-out of vaccination, risk of severe COVID-19 infection, and the vaccinations options (date, brand, homo- or heterologous schemes, and booster date).

Individuals were recommended to postpone the vaccination if they did not feel well or had a recent covid-infection, thus a healthy vaccinee effect may occur when comparing boosted patients with non-boosted people on the date the first one received the booster. Boosted and non-boosted comparators were required to have no contact with the healthcare system in the week prior to time 0 (the date the boosted pair received the booster) in main analysis. The strategies were tailored according to the data sources possibilities (for instance by mean of excluding patients with visits to primary care, or to hospital). This healthy vaccinee effect was not present in objective 1 since patients were compared at the moment they received the 2nd dose (when heterologous vs homologous were compared). Any other differential access to COVID-19 testing was investigated in a sensitivity analysis restricting to patients with a negative test as mentioned above. Other measured confounders were controlled for in the analysis through adjustment or stratification.

9.2. Setting

6 electronic health care databases in Southern, Northern, and Western Europe have been used to address the comparative effectiveness objectives during the study period 20/12/2020 to 28/02/2022 depending on the DAPs. The data sources that were included are those who have been working in prior studies (EU PE&PV or VAC4EU).

Italy

- Pedianet (Societa Servizi Informatici)
- Caserta local health database (INSPIRE srl)

Spain

- SIDIAP (Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària)
- BIFAP (Base de Datos para la Investigación Farmacoepidemiológica en el Ámbito Público)

The Netherlands

- PHARMO Database Network (PHARMO Institute for Drug Outcomes Research) (NL)

The United Kingdom

- CPRD (Clinical Practice Research Datalink)

Further information on the data sources used in this study can be found in Section 9.5.

These countries provided an outstanding setting for evaluating the COVID-19 vaccination effectiveness. They allowed for a study cohort with a size of 19 million vaccinated individuals. Pandemic emergency management, COVID-19 testing, and vaccination strategy significantly varied among countries, permitting the study of COVID-19 vaccine effectiveness in its whole heterogenicity.

9.3. Subjects

The source population comprises all persons registered in any of the data sources during the study period (December 2020-latest update [i.e., February 2022 for BIFAP, February 2022 for INSPIRE and December 2021 for remaining DAPs]).

The study population includes all individuals with:

- At least 2 years of baseline information at that moment ensuring that information on covariates was available (i.e., at least one year of data available before the start of the pandemic in 2020)

Then, the full**-COVID-19 vaccinated** study population includes all individuals with at least two recorded vaccinations since start of study period in each of the participating countries

- A person was classified to have a full primary vaccination regimen when a record of a second COVID-19 vaccine for persons with a first dose with Comirnaty, Spikevax or Vaxzevria (two doses primary vaccination schemes) was administered >19 days after the first dose (since it was difficult to distinguish between potential data entry errors, this is, two vaccination records too closed or 3 doses recorded, which may indicate double recording of the same vaccination, or failed/weak effect.)

9.3.1. Matched Populations

9.3.2. Matched population for the effectiveness of primary vaccination

From the full-Covid-19 vaccinated study population:

A) adolescents or adults with a homologous vaccination were matched 1:1 to persons with heterologous vaccination, or

B) children with full vaccination were matched 1:1 to children with no vaccination from the study population,

based on year of birth (to fall into the ranges 5-11; 12-17; 18-29; 30-49; 50-69; 70-79; ≥80) and opening by each birth year into those age bands, sex, region, any Covid-19 before time0, calendar

date of time zero (+/7 days) and 1st dose, brand of the 1st dose (in heterologous versus homologous analysis), and by whether the patient had any of the following conditions: immunodeficiency or Immunosuppressant treatment, cancer or malignant tumor, transplant, severe renal disease or Down Syndrome.

9.3.3. Matched populations for the effectiveness of booster

- Among individuals with a homologous primary vaccination:
    a. those with homologous booster were matched to persons without booster (1:1) based on year of birth (to fall into the ranges 18-29; 30-49; 50-69; 70-79; ≥80) and opening by each birth year into those age bands, sex, region, any COVID-19 before time0, calendar time0 (+/- 7 days) and booster-time0 (+/- 7 days).
    b. those with heterologous booster have been matched to persons with no booster (1:1) based on year of birth (to fall into the ranges 18-29; 30-49; 50-69; 70-79; ≥80) and opening by each birth year into those age bands, sex, region, any COVID-19 before time0 and calendar time0 (+/- 7 days), booster-time0 (+/- 7 days) and 1st dose.
    c. Among individuals with a heterologous primary vaccination, those with booster have been matched to persons without booster (1:1) based on year of birth (to fall into the ranges 18-29; 30-49; 50-69; 70-79; ≥80) and opening by each birth year into those age bands, sex, region, any COVID-19 before time0 and calendar time0 (+/- 7 days), booster-time0 (+/- 7 days) and 1st dose.

9.3.4. Follow-up

Persons without a matching pair were excluded from the analysis for that objective.

For primary objectives, follow-up started at time 0 (or booster time 0) and continue until the earliest of the following dates: COVID-19 disease/infection, death, last date of data extraction, or moving out of the data source. For primary objectives 1 and 4, follow-up is additionally censored at the date of booster vaccine administration in the non-boosted cohort or any extra dose recorded in the booster cohort. At this censoring date, follow-up for the matched boosted persons is also censored.

For children analysis, follow-up is additionally censored at the date of vaccine administration in the non-vaccinated cohort of children. At this censoring date, follow-up for the matched vaccinated children was also censored.

For secondary objective, follow-up starts at booster time 0 and continue until the earliest of the following dates: death, last date of data extraction, or moving out of the data source.

9.4. Variables

9.4.1. Definition of Time Zero (time 0)

Aligning the evaluation of eligibility criteria, covariate assessment, exposure assignment, and beginning of follow-up (time 0) avoids selection bias and immortal person-time bias. Time zero is when the vaccination status is assigned; all eligibility criteria must be fulfilled, and COVID-19 infection-related outcomes must start to be followed.

For objectives 1, 3 and 4, time0 is the date when the 2nd dose of COVID-19 vaccine is administered (i.e. recorded). That date has been used to match homologous to heterologous pairs or to non-vaccinated children.

For objectives 2 and 5 and secondary objective, the date when the booster vaccination is recorded (among patients with a booster dose) was defined as time 0 (booster-time 0). The booster-time 0 has been assigned also as time 0 for the non-boosted matched person.

9.4.2. Exposure information

Vaccination information is based on recorded prescription, dispensing, or administration of the COVID-19 vaccines. Vaccine receipt and date of vaccination have been obtained from all sources that can capture COVID-19 vaccination, such as pharmacy dispensing records, general practice records, immunization registers, vaccination records, medical records, or other secondary data sources. The main exposure of interest is the receipt of a primary regimen or booster COVID-19 vaccine, the dose, and its brand/manufacturer.

9.4.2.1. Exposure assignment

Exposure assignment is based on the brands used in the primary COVID-19 vaccination scheme and the booster vaccines. Table 3 below describes the different options.

Table 2. Classification of heterologous and homologous primary schemes and heterologous /homologous booster.

A full primary homologous regimen means:	**Booster classifications**		
	Non-boosted	Homologous	Heterologous
2 doses of Pfizer vaccine as primary regimen	no booster	Pfizer	Non-Pfizer
2 doses of Moderna vaccine as primary regimen	no booster	Moderna	Non-Moderna
2 doses of AstraZeneca vaccine as primary regimen	no booster	AstraZeneca*	Non-AZ
A full primary heterologous regimen means:			
1 st dose Pfizer, 2 nd dose Moderna or AstraZeneca	no booster		any
1 st dose Moderna, 2 nd dose Pfizer or AstraZeneca	no booster		any
1 st dose AstraZeneca, 2 nd dose Pfizer or Moderna	no booster		any

*Occur infrequently

9.4.2.2. Exposure assessment

9.4.2.2.1. Primary objective 1, 3, and 4: effectiveness of heterologous versus homologous primary vaccinations (or homologous vaccination versus no-vaccination among children)

Individuals have been assigned to the primary heterologous COVID-19 vaccination group if they receive a different COVID-19 vaccine brand for the 1st and 2nd doses of a two-dose primary (initial) course (see table 3 for the different options).

Heterologous primary vaccination is defined as the receipt of a different COVID-19 vaccine for the 1st and 2nd doses of a two-dose primary (initial) course.

Homologous primary vaccination is defined as the receipt of the same COVID-19 vaccine brand for the 1st and 2nd of a two-dose primary course.

Among children, non-vaccination is defined as no receipt of a COVID-19 vaccine at the time0.

A person is classified to be a boosted individual when any COVID-19 dose is administered at least 28 days after the 2nd dose (AstraZeneca, Pfizer, or Moderna). Unboosted individuals have been selected among those with at least 28 days of follow-up and with no third dose in that period. Until the time a third dose is received, the person has been considered non-boosted.

9.4.2.2.2. Primary objective 2 and exploratory objective 5: effectiveness of booster versus non-booster vaccinations

Persons receiving a dose ≥28 days after the 2nd dose (AstraZeneca, Pfizer, or Moderna) have been assigned as heterologous or homologous boosting, until the time a third dose is received, a person is considered non-boosted.

9.4.2.2.3. Induction time and effectiveness over time as a proxy of the waning of immunity

Effectiveness of heterologous versus homologous vaccinations or booster versus no-booster has been assessed over time in the first 0-6, 7-13, 14-30 days followed by monthly post (booster-)time0 intervals. Among children, effectiveness has been assessed for the main period of interest, i.e., after 2nd dose and over time. The date, dose, and type of vaccine administrated has been collected as reported in the previous section for each data source.

9.4.3. COVID-19 outcomes

This study considers different COVID-19 outcomes: severe COVID-19, COVID-19 with death and all COVID-19 infections.

9.4.3.1. Severe COVID-19 outcomes

A person has been considered to have severe COVID-19 disease when:

- COVID-19 was recorded as the cause of hospital or ICU admissions (if available in the data sources; see Data Sources Section)
- or, the hospital or ICU admission occurred within 30 days of a COVID-19 disease or positive test (If reason for admission is not available in a particular database; see Data Sources Section).

Cases of death with COVID-19 are included in this outcome.

9.4.3.2. Death with COVID-19

The administrative death (linked to the eHR, for instance, from the information on the expiring reason of the Health Identity Card), including the death date, was available in all participating databases but INSPIRE (Table 2), however the cause of death is missing in most of them. In those situations, algorithms (i.e., deaths with covid infection recorded in the previous 56 days) were utilised instead to identify death with COVID-19 for the primary analysis. A person was considered to have death with COVID-19:

- COVID-19 was recorded as the cause of death (if available in the data sources; see Data sources section).
- Or, the death occurred within 8 weeks of a COVID-19 disease or positive test (if the cause of death is not available in the database; see Data Sources section).

9.4.3.3. All COVID-19 infections

COVID-19 infection is defined as a positive test (PCR or antigens) or a COVID-19 diagnosis (depending on the database algorithm) regardless of the prognosis.

9.4.3.4. COVID-19 Information by data sources

Information on availability of tests and/or diagnosis for COVID-19 is listed in Table 4.

9.4.4. All-cause mortality

Deaths of any cause are included as available in all but one (IT-INSPIRE) data sources.

9.4.5. Covariates

To control for measurable confounders in the analysis, the following factors have been considered: lifestyle characteristics (BMI, smoking alcohol abuse), comorbidities, comedications, and health care utilization prior to or at time zero or booster-time zero. Lookback period covered 2018-time 0/booster-time 0 with a minimum of 2 years for each patient. The listed covariates are available in each database.

The following comorbidities (that may be shown to be associated with COVID-19 prognosis) have been assessed ever before time 0 and booster time 0:

- Diabetes mellitus (types 1 and 2)
- Hypertension
- Coronary artery disease
- Cerebrovascular disease
- Chronic respiratory disease
- Chronic liver disease
- Autoimmune disorders
- Parkinson disease
- Dementia, sepsis
- Heart failure
- Bladder incontinence
- Arthritis
- Coagulation deficiencies.

Comedication use has been assessed as proxies for comorbidities. The following comedications have been assessed during the year before time 0 and booster-time 0:

- Antibiotics
- Antiviral
- Corticosteroids
- Non-steroidal anti-inflammatory drugs
- Other analgesic
- Psychotropics
- Statins
- Influenza vaccine.
- >5 drugs (as a proxy of high level of morbidity)

Antibiotic and antiviral prescriptions have also been assessed in the month prior to time 0 and booster-time 0 as a marker of acute illness.

Health care utilization in the year before time 0 and booster time 0 have been evaluated as a proxy measure of health care–seeking behavior, overall health status, unmeasured confounders and access to health care. Additionally, short-term health care utilization in the week before time 0 and booster-time 0, have been assessed as short-term markers of current health status that may influence individuals' vaccination decisions (to minimize healthy vaccine effect). Considered variables included the following:

- Health care utilization year before (number of visits to PC or claims the year before time 0 or booster-time 0)
- Health care utilization week before (yes/no contact with the healthcare system in the week before time 0 or booster-time 0)
- Influenza vaccination (yes/no in the previous years, till 2018)
- Other non-childhood vaccinations (number in the previous years, till 2018)
- COVID-19 tests (total number, including positive and negative test if available in the data source)

Information on institutionalization or residency in a care home and on being a high-risk professional has been explored in the data sources and collected for confusion, controlling if available.

9.4.5.1. Genetic variant of SARS-CoV-2 virus

The presence of dominant genetic variants of SARS-CoV-2 (reaching 50% of total sequenced specimens) at time 0 and booster time 0 is defined based on three periods of dominant circulation (pre-delta, delta and omicron) obtained from surveillance data in the respective countries (according to SARS-CoV-2 variants dashboard in the European Centre for Disease Prevention and Control and other sources*) as follows:

Table 3. presence of dominant genetic variants of SARS-CoV-2 (reaching 50% of total sequenced specimens).

Country	Delta dominant start date	Omicron dominant start date
Spain	04/07/2021	03/01/2022
Italy		
Netherlands		
France		
UK	24/05/2021	

*ECDC: https://gis**.ecdc.**europa.eu/portal/apps/opsdashboard/index.html#/25b6e879c076412aaa9ae7adb78d3241
*France website: https://www.**santepubliquefrance.**fr/recherche/#search=Coronavirus%20:%20circulation%20des%20variants%20du%20SARS-CoV-2
*UK: https://www.**gov.uk**/government/publications/covid-19-variants-genomically-confirmed-case-numbers
Coronavirus Pandemic (COVID-19) - Our World in Data: https://ourworldindata.org/coronavirus#explore-the-global-situation

9.5. Data Sources and Measurement

The study has used data from secondary population-based electronic health care databases. All data sources provided high-quality data on COVID-19 vaccines (product types and dates), COVID-19 outcomes (test results, diagnoses in different settings, procedures, and treatments), and important covariates. Summary information about the data sources is provided in Table 4.

Table 4. Included VAC4EU and EU PE&PV data sources type of data and anticipated COVID-19 outcomes and vaccinations characteristics (for final data tables in results should be consulted)

Country	Data source	DAP & relation	HC Setting	Type of data	CDM	No. Active subjects (children, adolescents and adults)	Lag time	COVID-19 data	COVID-19 vaccine data (approximate N of vaccinations provided by DAPs on the 27/02/2022)	Death information (death date, cause of death)	Reference ENCePP data source Register (no. and link) and experience in COVID-19 studies
IT	CASERTA	Academic Spin-Off "INSPIRE srl"	Local health Unit	Hospital/emergency discharge, Immunization register, COVID-register, population file	CCDM	0.9 million	1-3 months	positive PCR tests from COVID register	COVID-19 vaccine administration date, brand and dose available. Homologous vacc. 1,744 aged 5-11; 22,791 aged 12-17 and 334120 adults. Heterologous: 212307 adults. Booster: Adolescent 12-17: 629; Adults: 27,406. until January 25th, 2022	Administrative death date	45094 Participation in CVM studies on COVID-19 disease and vaccines
IT	PEDIANET	SoSeTe	Pediatric primary care with linkage to the hospitalization database	Primary care medical records Immunization register Hospitalization database	CCDM	0.1 million	2-4 weeks	positive and negative PCR tests from COVID regisity	COVID-19 vaccine administration date and brand. Homologous vac. 88.606 aged 5-11 and 103.584 aged 12-14y	Death information (death date, cause of death) are always available when during hospitalization. If not, most likely to be reported by the FPs in the database.	20131 Participation in ACCESS & VAC4EU studies on COVID-19 disease and vaccines ORCHESTRA – EU funded project VERDI – EU funded project
ES	BIFAP	AEMPS	General practice	Primary care medical records linked with Hospital and ICU admissions and discharge	CCDM	4,237,801 aged <14 years 717,602 aged 14-17 years 12,727,542 aged 18-65 years 3,148,910 aged >65 years	3 months	COVID register, negative (in some regions)/positive tests COVID-related hospitalisations and ICU admissions	COVID-19 vaccine administration date, brand and dose available. 3,919,506 homologous and 8,933 heterologous vaccinated adults up-to October 2021. That needs to be updated for the current study.	Administrative death date	21501 Participation in ACCESS, ECVM, CVM studies on COVID-19 disease and vaccines EUPAS41134: Ongoing COVID-19 vaccines effectiveness study
ES	SIDIAP	IDIAP	General Practice	Primary care medical records Hospital discharge for part of population	CCDM	5.8 million	3 months	COVID tests negative/positive PCR, antigen tests	COVID-19 vaccine administration date, brand and dose available. 72,836 ,aged 5-11 and 570,795, aged 12-19. And 3,565,625 boosters among adults.	Administrative death date	4646 Participation in ACCESS, CVM studies on COVID-19 disease and vaccines

Country	Data source	DAP & relation	HC Setting	Type of data	CDM	No. Active subjects (children, adolescents and adults)	Lag time	COVID-19 data	COVID-19 vaccine data (approximate N of vaccinations provided by DAPs on the 27/02/2022)	Death information (death date, cause of death)	Reference ENCePP data source Register (no. and link) and experience in COVID-19 studies
UK	CPRD-Aurum	UU	Primary care, sample of national	Primary care medical records	CCDM	16 million	2-4 weeks	COVID tests negative/positive PCR, antigen tests	COVID-19 vaccine administration date, brand and dose available. Homologous vac. 49,671 children-adolescent and 4,935,754 adults. Heterologous: 136children/adolescents and 1,163,932 adults. Booster: 3,225,991 adults.	Death date recorded in GP record	https://cprd.com/primary-care Participation in ACCESS, ECVM, CVM studies on COVID-19 disease and vaccines
NL	PHARMO	PHARMO Institute	Primary care, sample of national	Primary care medical records	CCDM	2 million	1-6 months	COVID-19 positive PCR test	COVID-19 vaccine administration date, brand and dose available, with missingness in brand for GP data. Estimated 500,000 adults with booster.	Death date recorded in GP record	45066 Participation in ACCESS, ECVM, CVM studies on COVID-19 disease and vaccines

9.5.1. Caserta LHU database (INSPIRE srl), Italy

The Caserta database is a claims database containing patient-level data from the city of Caserta, in the Campania region. The catchment population in Caserta consists of more than 1 million persons (15% of the Campania regional population). The Caserta linkage databases consists of several databases which are linked through a unique patient identifier: a demographic registry, pharmacy claims database with information on concerning all dispensed drugs reimbursed by the Italian NHS, a as well as hospital discharge diagnose databases, emergency department admissions database, claims for diagnostic and laboratory tests ordered, and a registry of patients exempt from reasons for healthcare service co-payment exemptions (e.g. diabetes mellitus, dementia, and other chronic diseases), emergency department visit diagnoses and diagnostic tests. Patient level data from these claims databases, including other drugs reimbursed by the NHS and dispensed by community pharmacies, can be linked together, using a unique patient identifier. The healthcare information in the databases is coded using international coding systems, such as International Classification of Diseases, 9th Edition (ICD 9 CM) for diagnoses and Anatomic Therapeutic and Chemical (ATC) classification for drugs. A COVID-19 registry including all positive cases with clinical follow up is also available. COVID-19 outcomes and algorithms: the Covid-19 registry is the most reliable source to identify SARS-CoV-2 infection and severity of COVID-19 disease.

9.5.2. Pedianet pediatric data source, Italy

Pedianet, a pediatric general practice research database, was set up in 2000. It contains reason for accessing health care, health status (according to the Guidelines of Health Supervision of the American Academy of Pediatrics), demographic data, diagnosis and clinical details (free text or coded using the ICD-9-CM [International Classification of Diseases, Ninth Revision, Clinical Modification]), prescriptions (pharmaceutical prescriptions identified by the ATC code), specialist appointments, diagnostic procedures, hospital admissions, growth parameters, and outcome data of the children habitually seen by approximately 140 family pediatricians distributed throughout Italy. Pedianet can link to other databases using unique patient identifiers. In the first database, information on routine childhood vaccination is captured, including vaccine brand and dose. In the second database, information on patient hospitalization date, reason for hospitalization, days of hospitalizations, and discharge diagnosis (up to six diagnoses) is captured. The family pediatricians' participation in the database is voluntary, and individuals and their parents provide consent for use of their data for research purposes. In Italy, each child is assigned to a family pediatrician, who is the referral for any health visit or any drug prescription; thus, the database contains a detailed personal medical history. The data, generated during routine practice care using common software (JuniorBit®), are anonymized and sent monthly to a centralized database in Padua, Italy, for validation. The Pedianet database can be linked to regional vaccination data, which was successfully tested in several large European projects (e.g., ADVANCE) where it was characterized and deemed fit for purpose to evaluate prescriptions including pediatric routine vaccines. Timeframe for data availability: data are extracted every trimester. COVID-19 outcomes and algorithms: COVID-19 positive test from the COVID-19 registry (we have data also for negatives and undetermined). Free text algorithms on COVID-19 signs and symptoms are being developed for specific outcomes. Hospitalization clinical data are available as free text, including ICU admission start and end date.

9.5.3. SIDIAP, Spain

The Information System for Research in Primary Care (Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària [SIDIAP]) in Catalonia, Spain, is a primary care database set up by the Institute of Research in Primary Care (Institut Universitari D'Investigació en Atenció Primària Jordi Gol [IDIAP Jordi Gol]) and Catalan Institute of Health (Institut Català de la Salut). The database collects information from 278 primary health care centres and includes more than 5.8 million patients covered by the Catalan Institute of Health (approximately 78% of the Catalan population) and is highly representative of the Catalan population. SIDIAP data comprise the clinical and referral events registered by primary care health professionals (i.e., GPs, paediatricians, and nurses) and administrative staff in electronic medical records, comprehensive demographic information, community pharmacy invoicing data, specialist referrals, and primary care laboratory test results. SIDIAP can also be linked to other data sources, such as the hospital discharge database, on a project-by-project basis. Health professionals gather this information using International Classification of Diseases, 10th Revision (ICD-10) codes, ATC codes, and structured forms designed

for the collection of variables relevant to primary care clinical management, such as country of origin, sex, age, height, weight, body mass index, tobacco and alcohol use, blood pressure measurements, and blood urine test results. In relation to vaccines, information on all routine childhood and adult immunisations is included in addition to the antigen and the number of administered doses. Currently, because of the Covid-19 pandemic, having shorter-term updates to monitor the evolution of the pandemic is a possibility. Recent reports have shown SIDIAP data to be useful for epidemiological research. SIDIAP is listed under the ENCePP resources database (http://www.encepp.eu/encepp/resourcesDatabase.jsp). SIDIAP was characterised in the IMI-ADVANCE project and considered fit for purpose for vaccine coverage, benefits, and risk assessment (http://www.encepp.eu/encepp/viewResource.htm?id=4646). After EMA approval, the protocol must be evaluated by the SIDIAP Scientific Committe and by the IDIAPJGol Ethics Committee, the approval can take 4-6 weeks. The timeframe for data availability after the approval by the two local Committees is one month. **COVID-19 outcomes and algorithms.** We have two ways to get a COVID+ in SIDIAP data:

- Through diagnosis codes ICD10CM B34.2, B97.21 and U07.1;
- and using tests (PCR, antigens and antibodies).

COVID-19 admitted to hospital or ICU will be able to be identified. We also have registered the negatives and the undetermined tests.

9.5.4. BIFAP, Spain

BIFAP (Base de Datos para la Investigación Farmacoepidemiológica en Atención Primaria), a computerized database of medical records of primary care (www.bifap.aemps.es) is a non-profit research project funded by the Spanish Agency for Medicines and Medical Devices (AEMPS). Information collected by PCPs includes administrative, socio-demographic, lifestyle, and other general data, clinical diagnosis and health problems, results of diagnostic procedures, interventions, and prescriptions/dispensations. Diagnoses are classified according to the International Classification of Primary Care (ICPC)-2, ICD-9 and SNOMEDCT system, and a variable proportion of clinical information is registered in "medical notes" in free text fields in the EMR. Additionally, information on hospital discharge diagnoses coded in ICD-10 terminology is linked to patients included in BIFAP for a subset of periods and regions participating in the database. All information on prescriptions of medicines by the PCP is incorporated and linked by the PCP to a health problem (episode of care), and information on the dispensation of medicines at pharmacies is extracted from the e-prescription system that is widely implemented in Spain. The project started in 2001 and the current complete version of the database with information until December 2020 includes clinical information of 14,810 primary care practices (PCPs) and pediatricians. Nine participant autonomous regions send their data to BIFAP every year. BIFAP database currently includes anonymized clinical and prescription/dispensing data from around 20 million (17 active population) patients representing 92% of all patients of those regions participating in the database, and 32% of the Spanish population. Mean duration of follow-up in the database is 9 years. Information up to the end of 2021 and Covid-19 is also available for several regions from registries linked to the database. The BIFAP database was characterized in the ADVANCE project and considered fit for purpose for vaccine coverage, benefits, and risk assessment (http://www.encepp.eu/encepp/viewResource.htm?id=21501) and participated in the EMA-funded Early Covid-19 Vaccine Monitor study for COVID-19 vaccine safety monitoring. We estimate that the timeframe for data availability since the EMA approval is less than a month, since the approval from BIFAP Scientific Committee is required, which has a meeting every month, and from an Ethics Committee, which meets twice a month. **COVID-19 outcomes and algorithms.** COVID-19 infection was identified by:

- using positive tests (PCR, antigens and antibodies), which was the most reliable source to identify SARS-CoV-2 infection;
- COVID-19 diagnosis recorded through ICPC-2 and ICD-9 codes. Validations parameters using positive test as gold-standard have been estimated in a previous study;

The cases of COVID-19 related hospital and ICU admissions was identified as follows:

- COVID-19 was recorded as the cause of hospital or ICU admissions (information provided by some regions participating in the database);

- or a hospital or ICU admission with COVID-19 diagnosis recorded in the 30 days after a COVID-19 positive test (if the previous was not provided by the region).

9.5.5. CPRD, United Kingdom

The CPRD collates the computerised medical records of a network of general practitioners (GPs) in the UK who act as the gatekeepers of health care and maintain patients' life-long electronic health records. The data are sourced from over 2,000 primary care practices and include 62 million patients, of whom 16.5 million are currently registered and active (16). General practitioners act as the first point of contact for any non-emergency health-related issue, which may then be managed within primary care and/or referred to secondary care, as necessary. Secondary care teams also provide feedback information to GPs about their patients, including key diagnoses. The data in the CPRD are updated monthly and include demographic information, prescription details, clinical events, preventive care, specialist referrals, hospital admissions, and major outcomes, including death (17,18). Most of the data are coded using Read or SNOMED codes. Data validation with original records (specialist letters) is available. Depending on the type of electronic medical software used by the general practice, data are collected into either the CPRD GOLD (General Practitioner Online Database) or the CPRD Aurum database. The dataset is generalisable to the UK population based on age, sex, socioeconomic class, and national geographic coverage when the CPRD GOLD and the CPRD Aurum versions are used. Data include demographics, all GP/health care professional consultations, diagnoses and symptoms, results from laboratory tests, information about treatments (including prescriptions), data on referrals to other care providers, hospital discharge summaries (date and Read/SNOMED codes), hospital clinic summaries, preventive treatment and immunisations, and death (date and cause). Lag time for the CPRD GOLD and CPRD Aurum is 1 month. Information about vaccinations from mass vaccination campaigns during the pandemic is expected to transfer to GPs and into the patient's medical records (via National Health Service [NHS] systems rather than patients informing the GP); however, the lag time for this transfer varies. The present study included only active CPRD practices (Aurum). The CPRD is listed under the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) resources database, and access was provided by the University Utrecht (UU). CPRD approval would be required in order to approve the study protocol and it required approximately 4-6 weeks. For the current study, only primary care data in CPRD were used. This has information on positive tests, as well as diagnoses of COVID-19 (a CPRD provided lists for diagnoses of COVID-19 were used). Thus, CPRD cannot provide data for severe (hospitalised/ICU) COVID-19 analyses.

9.5.6. PHARMO, the Netherlands

The PHARMO Database Network, which is maintained by the PHARMO Institute for Drug Outcomes Research, is a population-based network of electronic health record databases that combines anonymous data from different primary and secondary health care settings in the Netherlands. These different data banks—including data from the general practices, inpatient and outpatient pharmacies, clinical laboratories, hospitals, the cancer register, the pathology register, and the perinatal register—are linked at the patient level through validated algorithms. To ensure data privacy in the PHARMO Database Network, the collection, processing, linkage, and anonymisation of the data are performed by STIZON, which is an independent, ISO/IEC 27001–certified foundation that acts as a trusted third party between the data sources and the PHARMO Institute. The PHARMO Institute is always seeking new opportunities to link with additional databanks and is currently exploring linkage with the COVID-19 immunisation register, which is collected by the Dutch National Institute for Public Health and the Environment (RIVM). Currently, the PHARMO Database Network covers over 6 million active persons of 17 million inhabitants of the Netherlands. Data collection period, catchment area, and overlap between data sources differ. Therefore, the final cohort size for any study depended on the data sources included. All electronic patient records in the PHARMO Database Network include information on age, sex, socioeconomic status, and mortality. Other available information depends on the data source. The lag time of all databases is 1 year, except for the General Practitioner Database, which is updated every 3 months or less. For this study we used the general practitioner database which comprises data from electronic patient records registered by GPs. The records include information on diagnoses and symptoms, laboratory test results, referrals to specialists, and health care product/drug prescriptions. Primary care data are available for a portion of the population of

approximately 3.2 million inhabitants (approximately 20% of the Dutch population). Information on lifestyle variables (e.g., BMI, smoking, alcohol consumption) is available in the General Practitioner Database if recorded by GPs in the electronic medical records. The PHARMO Institute uses de-identified data from existing databases without any direct enrollment of subjects. Ethical approval or informed consent is not necessary according to the Dutch law regarding human medical scientific research (Wet Medisch-wetenschappelijk Onderzoek met mensen (WMO)), which is enforced by the Central Committee of Research involving Human Subjects (Centrale Commissie Mensgebonden Onderzoek (CCMO)). Studies performed on the PHARMO Database Network are reviewed afterwards by the PHARMO Compliance Committee to assess whether the WMO requirements are met. **COVID-19 outcomes and algorithms.** COVID-19 outcomes are captured from data from the primary care level, using an algorithm with identifies: 1) COVID-19 episodes recorded by GP (either ICPC R83.03 or free text) and 2) COVID-19 tests, including test results. Date of infection is determined using multiple sources (date GP journal, mail, or COVID-19 test). Death date recorded in GP records.

9.6. Protection of human subjects

This is a non-interventional study using secondary data collection and does not pose any risks for individuals. Each data source research partner applied for an independent ethics committee review according to local regulations. Data protection and privacy regulations were observed in collecting, forwarding, processing, and storing data from study participants.

9.6.1. Patient information

This study involves data that exist in anonymized structured format and contain no patient personal information. All parties comply with all applicable laws, including laws regarding the implementation of organisational and technical measures to ensure protection of patient personal data. Such measures include omitting patient names or other directly identifiable data in any reports, publications, or other disclosures, except where required by applicable laws. Patient personal data are stored at DAPs in encrypted electronic form and password protected to ensure that only authorised study staff have access. DAPs implemented appropriate technical and organisational measures to ensure that personal data can be recovered in the event of disaster. In the event of a potential personal data breach, DAPs shall be responsible for determining whether a personal data breach has in fact occurred and, if so, providing breach notifications as required by law.

9.6.2. Patient consent

As this study does not involve data subject to privacy laws according to applicable legal requirements, obtaining informed consent from individuals is not required.

9.6.3. Ethical conduct of the study

This study adheres to the Guidelines for Good Pharmacoepidemiology Practices (GPP) and has been performed in line with the ENCePP Guide on Methodological Standards in Pharmacoepidemiology. The ENCePP Checklist for Study Protocols has been completed. The study is a post-authorisation study of vaccine effectiveness (PAES) and complies with the definition of the non-interventional (observational) study referred to in the International Conference on Harmonisation tripartite guideline Pharmacovigilance Planning E2E and provided in the EMA Guideline on Good Pharmacovigilance Practices (GVP) Module VIII: Post Authorisation Safety Studies and with the 2012 EU pharmacovigilance legislation, adopted June 19, 2012. The study has been registered in the EU PAS Register (EUPAS47725) before data collection started. The research team and study sponsor adhere to the general principles of transparency and independence in the ENCePP Code of Conduct and the ADVANCE Code of Conduct. There is no sponsor in the current study. The research team apply for the ENCePP Study Seal. The study is conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value, and rigour, and followed accepted research practices described in the Guidelines for Good Pharmacoepidemiology Practices (GPP) issued by the International Society for Pharmacoepidemiology (ISPE), and Good Epidemiological Practice guidelines issued by the International Epidemiological Association.

9.6.4. Institutional review board (IRB)/Independent ethics committee (IEC)

Each DAP followed the local country and data custodian requirements to apply for access to the data. All correspondence with the institutional review board or independent ethics committee and applicable documentation are retained as part of the study materials.

9.7. Bias

Efforts to assess and address potential sources of bias at the design stage are described in sections 9.4-Study Design and Limitations.

9.8. Study size

All patients meeting the eligibility criteria in the sources of data have been included in the source population, comprising, at the end, to more than 20 million patients (see Table 4 and Result section). 20 million full vaccinated patients aged ≥12 years and 287,000 vaccinated children aged 5-14 years (at the study start date) have been identified in the study. In Table 5, we provide the estimation performed prior vaccine effectiveness analyses of the probability that the upper limit of the 95% CI of the risk ratio (RR) is below 1.00 (a correlate of the lower bound of the vaccine effectiveness estimate being above 0.00 demonstrating a protective effect of vaccination) according to:

1. the population in the data sources participating in each analysis: around 19 millions of fully vaccinated patients for COVID-19 infections analysis according to the ROC20 interim reports (https://www.encepp.eu/encepp/viewResource.htm?id=42637).

2. the incidences in two different periods:
   - at the beginning of booster recommendations, i.e. November 2021 with low incidence of infection (21/100,000 hab.) but high hospitalization proportion (4.3%; incidence: 0.9 hospitalised COVID-19 per 100,000 habitants), and
   - at the beginning of 2022 with higher incidences of infections (716/100,000 hab.) but lower hospitalisation proportion (1.4%; 10 hospitalised COVID-19 per 100,000 habitants).

Under different VE scenarios (I.e. 25%, 51.3%, 75% or 95%), selecting 1 vaccine non-boosted to each boosted patient (ratio of 1:1) and a low incidence of 21 infections per 100,000 habitants in non-boosted patients, the probability could be lower than 1 as estimated and displayed in the following Table 5 (overall and by brand).

Based on European/national policies and previous studies of the same data sources (ROC19 EMA-2018-28-PE_ROC19_ECVM_Cohort Monitoring of Adverse Events), the proportion of patients receiving a heterologous second dose among those fully vaccinated (initial course) ranged by the source of data from 0.0-3.8% for AZ, 0.0-0.47% for PF, and 0.0-0.16% for MD at the beginning of the campaign.

As expected, the numbers of heterologous primary vaccinations and booster in all databases are much lower (1-5%) than that of the homologous ones. That is limiting the probabilities to match and estimating comparative effectiveness between homo- and heterologous primary vaccinations, and heterologous boosters versus non-booster. However, we expected enough power based on the large population covered by data sources, and the high percentage of full primary regimens (see Table 1) in the participating countries. Therefore, we are matching by several factors and finding good representation of clinical subgroups that will inform the meta-analysis.

The low number of vaccinated children identified, due to the lack of updated instances beyond December 2021, limited many of the analysis planned.

As we expected, the sample is around 20 million fully vaccinated individuals. If we hypothesize a very low theoretical 25% effectiveness value of booster vs non-booster against COVID-19 infection and considering a 21.2/100,000 habitants risk value: only in the case of the study population would suffer a strong reduction equal to 1/10 of the total expected (reduced till 1 million in each compared matched cohort as displayed in the Table 5), the probability will be lower than 0.97 (i.e. 0.77). However, even in this unlikely condition, the sample size will always be reached.

Table 5. Study Size Precision Estimates assuming an effectiveness of 25% of the booster against covid-19 infection in comparison with non-boosted patients, under an incidence of 21.2 Covid-19 infections per 100,000 non-boosted patients, allocation ratio of 1, and maximum sample size of 19,287,049 adults over all data sources.

Size	Number in boosted	Number in non-boosted	SE (ln(RR))	Span	Probability of the Upper Limit of the 95% CI to Be Below 1.00
1928704.9	964352.5	964352.45	0.106823	-0.20937	**0.768257**
3857409.8	1928705	1928704.9	0.075535	-0.14805	**0.967744**
5786114.7	2893057	2893057.35	0.061674	-0.12088	**0.996581**
7714819.6	3857410	3857409.8	0.053411	-0.10468	**0.999694**
9643524.5	4821762	4821762.25	0.047773	-0.09363	**0.999976**
11572229.4	5786115	5786114.7	0.04361	-0.08547	**0.999998**
13500934.3	6750467	6750467.15	0.040375	-0.07913	**1.000000**
15429639.2	7714820	7714819.6	0.037767	-0.07402	**1.000000**
17358344.1	8679172	8679172.05	0.035608	-0.06979	**1.000000**
19287049	9643525	9643524.5	0.03378	-0.06621	**1.000000**

Using simulated calculations under difficult conditions we are facing, the severe COVID-19 analysis could undergo a reduction in the probability to 0.87 in case the VE was 25% under an incidence of hospitalised Covid-19 of 0.9/100,000. The sample size could be reduced to limit the probability to 0.28. But, this need to be calculated with the finally observed matched data.

9.9. Data transformation

This study has been conducted in a distributed manner using a common protocol, common data model (CDM), and common analytics programs based on existing health data. The following steps have been accomplished:

- Extraction, transformation, and loading (ETL) of data to a CDM. To harmonise the structure of the data sets stored and maintained by each data partner, a shared syntactic foundation has been used. The CDM is the Conception CDM v2.2 for EHR. In this CDM, data are represented in a common structure, but the content of the data remain in their original format.
- The ETL design is shared in a searchable FAIR catalogue. The VAC4EU FAIR Molgenis data catalogue is a meta-data management tool designed to contain searchable meta-data describing organisations that can provide access to specific data sources.
- To reconcile differences across terminologies, a shared semantic foundation is built for the definition of events under study by collecting relevant concepts in a structured fashion using a standardised event definition template. The Codemapper tool has been used to create diagnosis code lists based on completed event definition templates for each outcome and comorbid risk condition. Based on the relevant diagnostic medical codes and keywords, as well as other relevant concepts (e.g., medications), one or more algorithms are constructed (typically one sensitive, or broad, algorithm and one specific, or narrow, algorithm) to operationalise the identification and measurement of each event. These algorithms may differ by database, as the components involved in the study variables may differ. Specifications for both ETL and semantic harmonisation were shared in the catalogue.
- Third, following conversion to harmonised study variable sets, R programs for the calculation of incidence and risk have been distributed to data access providers for local deployment.
- The aggregated results produced by these scripts have been then be uploaded to the Digital Research Environment (DRE) for pooled analysis and visualisation (see Figure 3). The DRE is made available through UMCU (University Medical Center Utrecht) (https://www.andrea-consortium.org/). The DRE is a cloud-based, globally available research environment where data are stored and organised securely and where researchers can collaborate (https://www.andrea-consortium.org/azure-dre/).

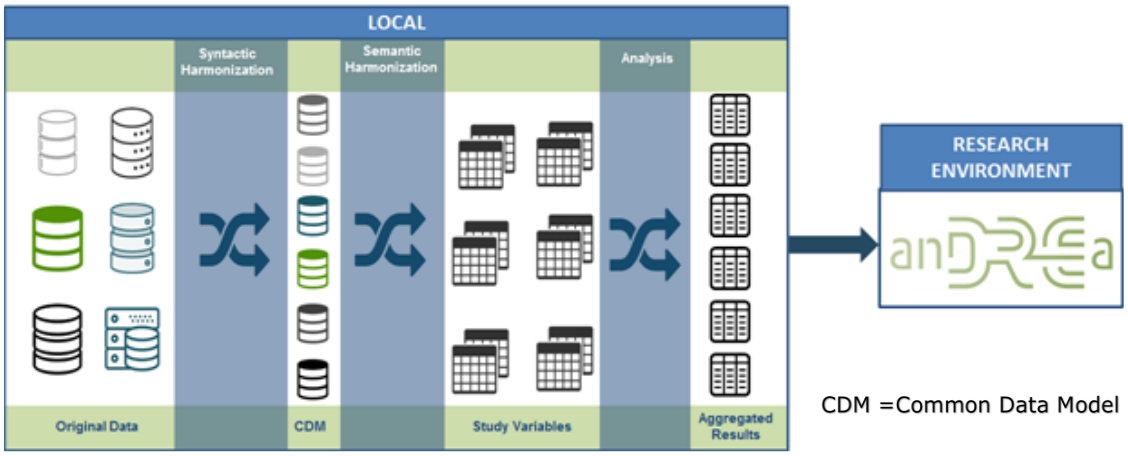

Figure 2. Data management flow.

9.9.1. Quality management and control

Data transformation into the CDM has been conducted by each subcontracted research partner in its associated database, using the processes described in the following sections (see below), each of these steps is fully transparent and was signed of/reviewed. Standard operating procedures or internal process guidance at each research centre have been used to guide the conduct of the study. These procedures include rules for secure and confidential data storage, backup, and recovery; methods to maintain and archive project documents; Quality control procedures for programming; standards for writing analysis plans; and requirements for scientific review by senior staff.

9.9.2. Quality checks of DAPs data

Data partners have been asked to perform data quality checks (Figure 3) for level 1 (Completeness), level 1b (Data presence), and 2 (consistency) (or any EMA quality framework checks when available); level 3 is checking for study variables and assess whether data are fit for purpose.

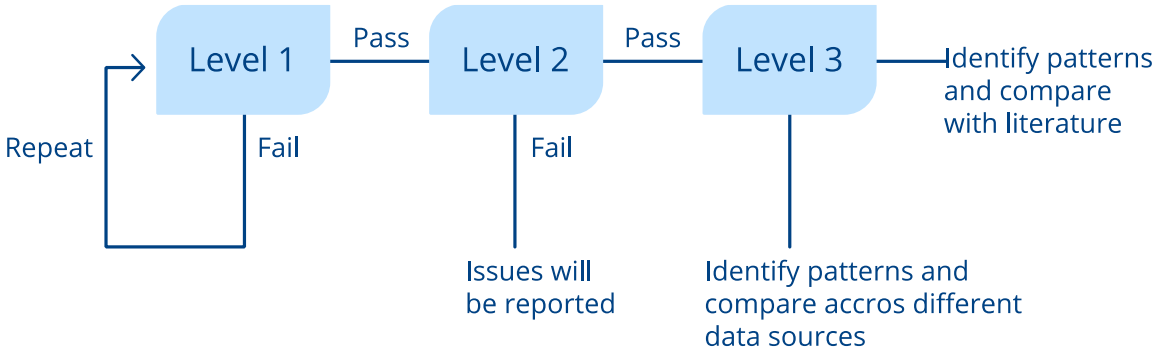

Figure 3. Data quality check cycle as developed in ConcePTION Quality Framework.

Generic open-source data quality check scripts are available from the IMI-ConcePTION quality framework that are publicly available on GitHub.

9.9.2.1. Level 1 - Data completeness

The purpose of the level 1 check is to verify the completeness of the ETL process and the data in the variables. Examples of tests are:

- Presence of variables in each of the CDM tables in D2
- Checks for misspelling and letter case in variable names in each of the CDM tables
- Verification of vocabularies
- Check date formats
- Check conventions of values
- Missing data analysis
- Frequency tables for categorical variables

- https://github.com/IMI-ConcePTION/Level-1-checks

9.9.2.2. Level 1b – Data presence

The aim of this script is to explore which data is stored in the CDM. This can be compared to the study specific specifications of the study variables. https://github.com/UMC-Utrecht-RWE/ConcePTION-Level1b#readme

9.9.2.3. Level 2 - Data logic/consistency

Real data are not random but follow certain logical constraints that reflect rules governing real-world situations. Examples of indicators generated by level 2 checks are:

- Event dates before date of birth
- Event dates after date of death
- Event dates out of observation periods
- Subjects having an observation but not present in the PERSONS table
- Observations associated with a visit id and occurred before/after the visit start/end date
- Subjects younger than 12 years old reported as parents
- Age at the observation period older than 115 y old
- https://github.com/IMI-ConcePTION/Level-2-checks

9.9.2.4. Level 3 – Fit for purpose

Level 3 checks review patterns of study variables over time, age within and between datasources. There are 8 modules, which may be used depending on the study variables:

- Source and study population.
- Medicines
- Vaccines
- Diagnoses
- Pregnancy
- Populations of interest.
- Health-seeking behaviour and lifestyle factors.
- EUROCAT indicators.
- https://github.com/IMI-ConcePTION/Level-3-checks

9.9.3. Quality checks of R-coding

Data Management and Statistical Analysis followed standard operating procedures for UMCU. All Statistical Analysis programs have been double coded or reviewed by UMCU and ARS. UMCU created clear documentation (graphical and in Excel spreadsheet) of the data management steps, beginning with describing the required variables from the CDM and each of the subsequent transformation steps and intermittent data tables. ARS double coded and/or conducted code review of the datasets built in R by UMCU using R and from instructions provided by UMCU. Discrepancies have been resolved.

9.9.3.1. Coding conventions (process quality)

We used GitHub (and the underlying git version control system) to collaborate with multiple parties on several projects involving writing scripts and functions. At its core, GitHub tracks all changes and shows which, when, who and why changes were made. In the chain of events, any previous state can be recovered easily. Regarding proposed changes or potential bugs, GitHub provides a platform to discuss details. Using GitHub Actions, standard workflows have been defined and executed after a submitted change. An example has been executing unit tests to ensure that scripts are correct. The main coordinator of the GitHub is the UMCU who creates a repository for the study and provides the 'main' functions to be used in each study.

- A readme file is initialized with relevant information about the scripts.
- For each study, a 'branch' is created in which scripts are tailored to the respective study.
- After each version update, the coordinator requests from all teams to incorporate the changes using 'merge'. One responsible from each team is appointed and allowed access to the repository. In case the main scripts contain an error, the 'issues' functionality is used to report the bug. If possible, a bug fix can be proposed by creating a 'pull request'. The `issues` platform also provide a means to ask for further clarification regarding new versions.

We used one set of standard conventions for all parties to facilitate collaboration and minimize bugs in scripts. Coding conventions are categorized into three parts:

- Notation (e.g., name scripts, functions, and objects).
- Syntax (e.g., spacing, braces, indentation)
- Documentation (e.g., writing comments, dividing code into sections)

Script names have been informative, where words are separated with an underscore or a hyphen. For scripts that are executed sequentially, the names are prefixed with numbers that indicate the order. For naming functions and objects, we suggest adopting the "snake style", where words are separated with an underscore. For syntax rules, we have implemented the tidyverse style guide found at https://style.tidyverse.org/. To facilitate implementing these rules, we used the 'formatR' R package. This package automatically restyles R code to adhere to these rules. For documentation, comments have been provided that explain each part of the code. Each script file starts with a title, author, date, and version number. Comments are placed to describe functions and objects.

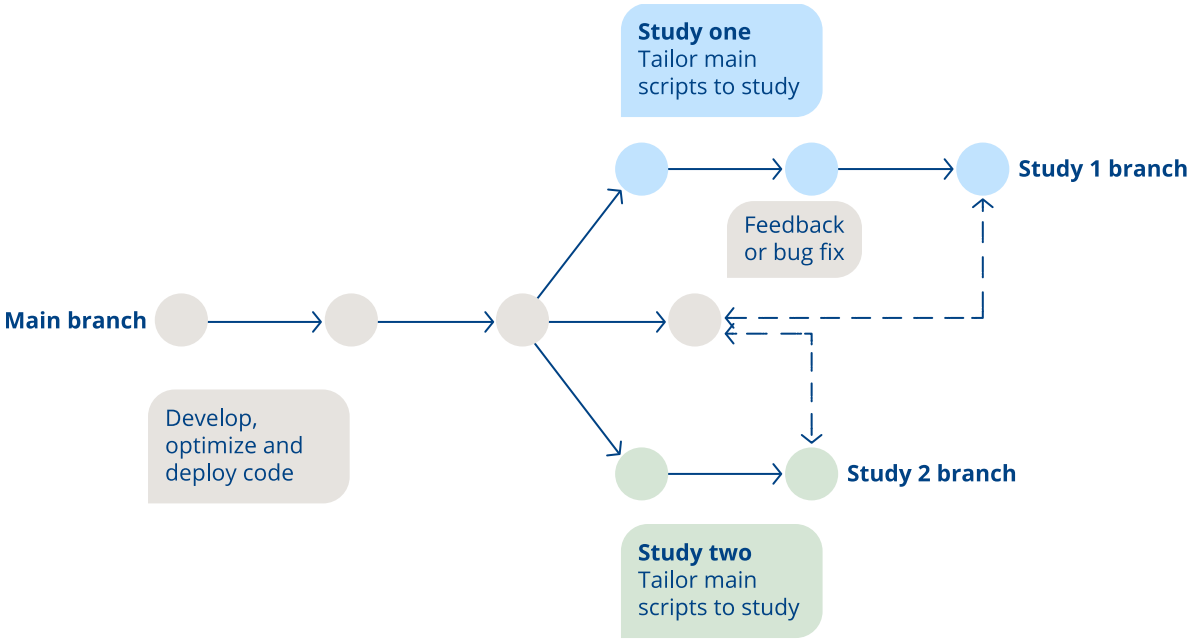

Figure 4. A workflow for collaborating on code with several studies in GitHub.

9.9.3.2. Function creation and release

Functions are modules of code that accomplish a specific limited task. The development of functions consists of a series of 5 steps:

1. Specifications of the function by the initiator
2. Approval of the function specification by the owner.
3. Programming of the function by a qualified programmer
4. Testing of the function by another qualified programmer (the tester), who completed a test specification form. For unit testing, all the test scenarios are added to a test script that needs to be performed after every update of the function.
5. The function owner checks if all steps are complete, and deployment is approved.

9.9.3.3. Standard/bespoke analyses script creation, testing and release

Study scripts connect and package functions using a structured design and follow the statistical analysis plan. Study scripts have been created in 4 steps:

1. Defining a map of the script, which includes specification of the folder structure, data model, graphical representation of the steps, use of functions, allocation of responsibilities and timelines, plus review schedules.
2. Programming of the code by a qualified programmer plus statistician. Test with code profiler to monitor bottlenecks in the code.
3. Testing the scripts by second programmer on:
   a) Small, simulated dataset in the specific CDM

b) Test on big data: run the script on the 1 million test dataset and test performance (see QAC1 for creation)
c) Test script on one real data partner before making it available for all the DAPS.

4. Take the script into deployment

For each update, steps 1-4 are repeated.

9.9.4. Data transformation

Based on the diagram in (Figure 5), the transformation steps (T) listed below in this section have been implemented by the data partners.

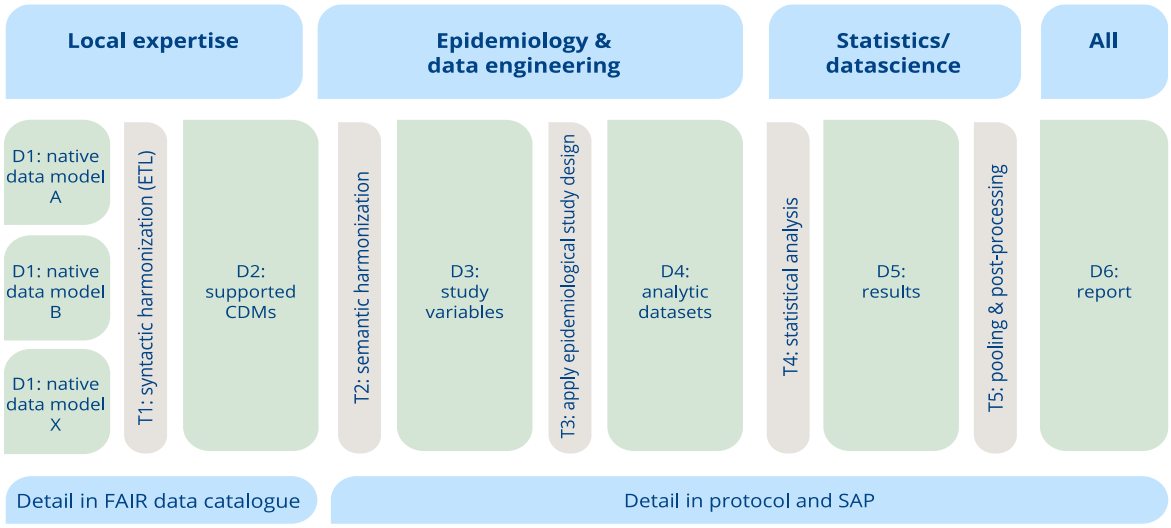

Figure 5. Data Management Process.

CDMs, common data models; Dn, data type; FAIR, findable, accessible, interoperable, and reusable; Tn, transformational step, for explanation, see below.

9.9.4.1. T1: Syntactic harmonisation (ETL)

Syntactic harmonisation through an extraction, transformation, and loading (ETL) process of native data into the ConcePTION CDM. In this CDM, data are represented using a common structure, but the content of the data remain in their original format. The CDM version that has been used is v2.2, which is available as an open-source CDM. The CDM was developed as part of the IMI-ConcePTION project (project number IMI-821520). The ETL process has various structured steps as described in Thurin et al. 2021 (19,20):

- DAPs have been asked to share the data dictionaries of their data banks (tables and variable names/structure) with the principal Investigator.
- Based on the data dictionaries of the data banks, an interview has been conducted by the principal Investigator that explores what action(s) prompts the creation of a record, what is missing, and the context of each of the data banks.
- Metadata (descriptive data about the data sources and databanks), data dictionaries, and interview answer sheets are uploaded in the VAC4EU FAIR (findable, accessible, interoperable, and reusable) data catalogue, according to the metadata structure for electronic health data that was defined in IMI-ConcePTION and the European Medicines Agency (EMA)–funded MINERVA project.
- An overview is created of all the required study variables and definitions in order to create the code lists to identify the outcomes and covariates (see T2).
- Instructions for the ETL design are provided by the WP2, these instructions comprise:
  - The required CDM tables
  - Mandatory variables
  - The calendar period over which data needs to be extracted
  - Code lists for the data to be extracted
  - The ETL design for each study can be conducted on paper or directly in the FAIR catalogue. The VAC4EU FAIR data catalogue (software and server provided through MOLGENIS:

https://www.molgenis.org) is a metadata management tool designed to contain searchable metadata describing organisations that can provide access to specific data sources (VAC4EU). The ETL design section details how the native data are mapped to the different CDM tables. Review of the ETL design is conducted by the study team.

- Once the ETL design has been approved, it has been executed by the DAP using its programming language. The output are CSV (comma-separated values) files.
- Once the ETL has been conducted, Level 1 and 2 data quality checks are conducted to measure the integrity of the ETL, as well as internal consistency within the context of the CDM. University Medical Center Utrecht (UMCU) is responsible for the Level 1 and 2 scripts.
- Level 1 data checks are performed to assess the completeness and content for each variable in all CDM tables to ensure that the mandatory variables contain data and conform to the formats specified by the CDM specifications, e.g., data types, variable lengths, formats, acceptable values. Each DAP was responsible for running the Level 1 R-scripts to complete the Level 1 checks (see https://github.com/IMI-ConcePTION/Level-1-checks/tree/master). A standard R Markdown report describing results of the checks for each table of the CDM was produced by the script. Level 1 checks have been reviewed by the UMCU team with each of the DAPs, to assess completeness and consistency. After addressing any issues identified in Level 1 checks, DAPs may rerun the script and inspect the results. This control and correct process could have been repeated until the ETL has been judged to be sufficiently complete and correct by the DAP.

- Level 2 data checks are performed to assess the logical relationship and integrity of data values within a variable, or between 2 or more variables within and between tables. Records occurring outside of recorded person-time, i.e., before birth, after death, or outside of recorded observation periods, are assessed. Each DAP was responsible for running the Level 2 R-scripts to complete the 2 checks https://github.com/IMI-ConcePTION/Level-2-checks). An R Markdown report describing results of the Level 2 checks for each CDM table has been produced. After addressing any issues identified in Level 2 checks, DAPs may rerun the script and inspect the results together with the UMCU team. This control and correct process could have been repeated until the ETL is judged to be sufficiently complete and correct by the DAP.
- Each of the corrections, changes, and edits between sequential Level 1 and 2 checks and adaptations is documented in specific database instance reports and signed off by the DAP using specific forms.

9.9.4.2. T2: Semantic harmonisation

To reconcile differences between terminologies and native data availability, a shared semantic foundation has been built for the creation of relevant study variables. This has been a multistep process:

- **Definition of study variables**, which is done using an event definition form that systematically captures the following items and is a living document that was closed upon study ending.
  1. Purpose of the event: covariate or outcome
  2. Version
  3. Document history
  4. Clinical definition
  5. Synonyms/lay terms (for text mining purposes)
  6. Laboratory tests specific for diagnosing event
  7. Diagnostic tests specific for diagnosing event
  8. Drugs that are used to treat event
  9. Procedures used to treat event
  10. Setting where condition is diagnosed (hospital, outpatient, GP)
  11. Diagnosis codes or algorithms used in other papers (health outcomes of interest)
  12. Codes used for study
  13. Algorithm proposal
  14. References
  15. Examples of event definition forms can be found in the VAC4EU Zenodo repository (https://zenodo.org/communities/vac4eu/?page=1&size=20)
- **Initial code lists** are created using the VAC4EU CodeMapper tool (https://vac4eu.org/codemapper/) to assist in the creation of code sets from case definitions for several coding systems simultaneously while keeping a record of the complete mapping process. Study variables are named in a standard hierarchical fashion based on body system.

- **Review of the codes by DAPs:** The output of the CodeMapper is a Microsoft Excel list, which has been inspected by the DAPs and commented on at the VAC4EU SharePoint.
- **Consolidation:** Comments from DAPs are consolidated by the study team. The code lists are read automatically through R/SAS code and:
    - Check for ranges in output
    - Check for strange codes
    - Insert the codes in the creation of concept sets, which are used to extract the data from the various CDM tables
- Based on the relevant diagnostic medical codes and keywords, as well as other relevant components (e.g., medications), one or more algorithms have been constructed to operationalise the measurement of each study variable. These algorithms may differ by database, as the components relevant for the study variables may differ.
- During the T2 step, transformations occur for a series of steps related to completion of missing features in the data, e.g., dose of vaccines, sorting on record level, combination of concepts for algorithm, and rule-based creation of study variables on a personal level for the study population, specific if needed per DAP.
- Once the study variables are created, Level 3 checks have been deployed, targeting to assess the patterns of study variables between data sources and against external benchmarks. A public example is available from the IMI-ConcePTION GitHub (https://github.com/IMI-ConcePTION/Level-3-checks).
    - The Level 3 checks are divided in 8 major modules, which can be tailored to the specific study variables:
    - Source and study population
    - Medicines
    - Vaccines
    - Diagnoses
    - Pregnancy
    - Populations of interest
    - Health-seeking behaviour and lifestyle factors
    - EUROCAT indicators

An R Markdown report describing results of the Level 3 checks has been produced by the script. Level 3 checks have been reviewed by the study team with each of the DAPs, to assess whether study variables were fit for purpose. After addressing any issues identified in Level 3 checks, DAPs may rerun the script and inspect the results. This control and correct process could have been repeated until the Level 3 checks are judged to be sufficiently complete and correct by the DAP.

9.9.4.3. T3: Application of epidemiological study design

Based on the creation of the study variables on a person level or a medicines level, epidemiological designs has been applied, such as sampling, matching (on specific variables and/or propensity scores), and selection based on inclusion and exclusion criteria. For some data access providers (DAPs), before extensive medical data can be extracted, preliminary matching of vaccinated and comparators on key demographic (e.g., age, sex, calendar time, and region) were needed since not all the medical data could be extracted for the D1 data instance. These designs have been defined in the study protocol and may differ per study objective. Subsequent matching on medical conditions and other variables have been conducted as part of the T3 step, once the required medical information was extracted. The designs were implemented for the various study objectives using R-scripts, using the existing functions (R-cran) or macros.

9.9.4.4. T4: Statistical analysis

This step in the data transformation pipeline produced statistical estimates for descriptives (counts, percentages), distributions (mean, percentiles), rates (prevalence, incidence), regression coefficients, and other relevant estimates. This has been conducted using R or Stata scripts.

9.9.4.4.1. Scripting and deployment

The analytical R scripts that produce the T2-T4 steps was produced on VAC4EU GitHub for version control, links to the latest script have been distributed to DAPs for local deployment. Any issues was notified on the GitHub, and the data engineers who are responsible for the R code worked with the local DAP to resolve issues when occurred. Scripts were developed independently, based on a data

engineering program and codebook in R (UMCU/ARS), and the outputs were compared against each other for validation.

9.9.4.5. T5: Results and pooling post-processing

The aggregated results produced through T4 have been uploaded to the Digital Research Environment (DRE) for pooled analyses and visualization. The DRE is made available through VAC4EU and UMCU (The anDREa consortium 2021). The DRE is a Microsoft Azure cloud-based, research environment with double authentication where researchers can collaborate using data that are stored and organized securely. UMC Utrecht is responsible for data processing and data security.

All researchers who need access to the DRE have access to study-specific secure workspaces by VAC4EU/UMCU. Access to the workspaces is possible only after double authentication using an identification code and password together with the user's mobile phone for authentication.

Uploading of files is possible for all researchers with access to the workspace within the DRE. Downloading of files is possible only after requesting and receiving permission from a workspace member with an "owner" role, who is a UMCU team member.

9.10. Statistical Methods

Analyses have been conducted using R version R4.0.3 or higher (Foundation for Statistical Computing, Vienna, Austria; https://www.R-project.org). Additional tools as RevMan for Meta-analysis 5.4 or Stata® version 16 or higher were also be used when they were needed.

9.10.1. Descriptive study

Distributions of baseline and COVID-19 vaccination characteristics at time 0 have been assessed and reported in the all-COVId-19 vaccinated population and in the matched populations to describe and illustrate differences between the compared groups by scheme (homologous versus heterologous; or non-vaccinated in children; booster and non-booster) (see ANNEX 1). The number and proportion of different brands used in homologous and heterologous primary vaccinations and boosters, over the total patients vaccinated by month have been reported in the Results section. For continuous variables, means, standard deviations, medians, and quartiles have been estimated and reported by compared groups. For categorical variables, counts and proportions have been reported by compared groups. The missingness of variables is also described.

Furthermore:

- Incident rate (IR) of each COVID-19 outcome (i.e., severe covid, death with covid and all covid) and 95% confidence intervals (CIs) by **primary vaccination** matched cohorts are estimated overall, by age groups (adults, adolescents and children), by brand of the 1st dose and time after full primary vaccination (time 0; only the pairs in which both individuals are still at risk at the beginning of each studied period are included).

- IR (and 95% CI) of each COVID-19 outcome (i.e. severe covid, death with covid and all covid) CIs are estimated in the **booster and non-booster** matched cohorts overall adults and by type of primary vaccination scheme (heterologous or homologous), type of booster (heterologous or homologous) and time since booster-time 0 (only the pairs in which both individuals are still at risk at the beginning of each studied period were included).

- Generate IPW-weighted Kaplan-Meier curves to depict the cumulative incidence of the outcomes by matched cohorts over time after time 0 (for adults, adolescents and children separately, and overall and by brand i.e. PF, MD, AZ) and booster-time 0 (for adults by booster and no-booster matched cohorts overall), by primary scheme (heterologous and homologous separately) have been created following the steps:

1. Compute the propensity score (PS) model and use it to predict the conditional probability of receiving treatment for each individual based on the potential confounders (including matching variables and the listed below:
   Influenza vaccination (yes/no in the previous 4 years)
   Other vaccinations (number in the previous 4 years)

COVID-19 tests (total number ever before time0 when available)
Comorbidity (up-to 4y before unless otherwise specified):

- Diabetes mellitus (types 1 and 2)
- Hypertension
- Coronary artery disease
- Cerebrovascular disease
- Chronic respiratory disease
- Chronic kidney disease
- Chronic liver disease
- Autoimmune disorders
- Parkinson disease
- Dementia
- Sepsis
- Heart failure
- Bladder incontinence
- OsteoArthritis
- Coagulation deficiencies
- Coronary artery disease
- Chronic liver disease

Comedication (up-to 4y before unless otherwise specified):

- Antibiotics 30 days
- Antiviral 30 days
- Corticosteroids
- Non-steroidal anti-inflammatory drugs
- Analgesic
- Psychotropics
- Statins
- >5 drugs (as a proxy of high level of morbidity)

2. Create the weights: for those vaccinated the IPW was 1/PS and for those unvaccinated the IPW was 1/(1-PS)
3. Use these weights to weigh the KM curve.

9.10.2. Effectiveness study (primary objectives 1,2,3, 4 and secondary objective)

For each matched compared cohort, the following comparative analyses have been performed:

- Study the risk of COVID-19 related outcomes, applying Cox regression to fit a proportional hazards model and derive average hazard ratio (HR) and 95% CIs. Additional confounding, was used as IPW to control for in the hazard models. Only the pairs in which both individuals were still at risk at the beginning of each studied period were included in the analysis of that period.
- The adjusted vaccine effectiveness (VE) has been estimated as 1 minus the adjusted HR multiplied by 100 (and its confidence intervals were estimated as 1 minus the confidence intervals of HR multiplied by 100). VE for COVID-19 outcomes are presented for adults, adolescents and children separately, overall matched cohorts and by brand of the 1st dose (i.e. PF, MD, AZ).

Severe COVID-19 and COVID-19 with death analysis have been stratified by previous COVID-19 infection (yes/no) and by periods of country's predominant circulating COVID-19 genetic variants at time 0 and booster time 0 as a proxy of variant-specific vaccine effectiveness (i.e., pre-alpha; alpha; delta; omicron) (21). VE has been provided by data source.

Since many point estimates and 95% CIs have been produced, reporting all them in text would make it difficult to read and interpret. In order to easily read and simplify interpretation, we do not report in text the 95% confidence intervals (that can be found in tables), neither the non-significant VE (we

only mention in text that 'non-significant VE was found'). Non-significant VE (and its 95% CI) values can be found in tables. Also, in scenarios with many estimated VEs, we are reporting VE point estimates in ranges, meaning that we report the lowest and highest VE point estimates found across data sources.

9.10.2.1. Meta-analysis

For analysis restricted to subgroups of people (reaching reduced sample size) or rare outcomes, meta-analysis have been performed, i.g.:

- VE of booster against severe COVID-19 among adults in clinical subgroups i.e. patients with Immunodeficiency or Immunosuppressant, Cancer or malignant tumor, Transplant recipient, Severed renal disease or Down Syndrome
- VE of homologous vaccination versus non-vaccinated against COVID-19-related death among children severe COVID-19 and/or death with COVID-19.

Some meta-analysis could not be performed due to lack of or insufficient number of cases (see ANNEX_6) (i.e., severe COVID-19 among children, etc.).

Using the main estimates from each data source, appropriate random-effects meta-analytic methods (inverse variance method or others, as needed) have been obtained and combined effect estimated. The analysis of outcomes is based on individuals, not on number of events. We used Stata and RevMan Version 5.4 for analyses. Both adjusted hazard ratios and incidence rate differences have been used to summarize findings and a forest plot was produced with the data sources' estimates and the pooled estimate. We used the Chi^2 and I^2 statistics to test for heterogeneity of treatment effect between trials. We considered a Chi^2 value $p < 0.05$ or I^2 value $> 50\%$ as indicative of heterogeneity. When data exhibits substantial heterogeneity ($I^2 > 60\%$), we investigated possible causes. As explained above, a sensitivity analysis for VE in children and for adults with comorbidities with booster vaccination (in which compared groups will not receive a vaccination at time0 and booster-time0) have been performed stratifying by whether the data source included information of primary care consultations or not. This analysis provided an estimation of the confusion of healthy vaccinee effect present in data sources not including primary care consultations/data.

9.10.3. Sensitivity analyses

Sensitivity analyses were conducted restricting to pairs with a negative COVID-19 test recorded prior to time0 or booster-time0 in data sources with negative COVID-19 tests information (BIFAP-ES, SIDIAP-ES and PEDIANTE-IT). A comparison of the VEs resulted between that analysis and main analysis in those data sources provided an estimation of the effect of the differential access (if any) to COVID-19 testing among compared groups in the effectiveness.

9.11. Quality control

Data transformation into the CDM is conducted by each subcontracted research partner in its associated database, with processes as described in this report and the corresponding study protocol. Standard operating procedures or internal process guidance at each research center are used to guide the conduct of the study. These procedures include rules for secure and confidential data storage, backup, and recovery, methods to maintain and archive project documents, quality check procedures for programming, standards for writing analysis plans, and requirements for scientific review by senior staff.

10. Results

For the main analyses (matching for patients without prior COVID-19), the vaccine effectiveness (VE), crude and adjusted HR, Kaplan-Meier curves, and data sources' details about clinical subgroups and outcomes are available in the ANNEXES, specified in each subsection. The main results are described using the estimated VE in percentage (only for statistically significant) comparing different cohorts of interest, otherwise '95%CI' is stated.

10.1. Primary objective 1 (adults and adolescents) and 4 (waning of immunity) results: effectiveness of primary vaccination, homologous versus heterologous (= reference group)

Vaccine effectiveness and Kaplan-Meier curves calculated in the main analysis (i.e., among pairs without prior COVID-19) are displayed in ANNEX_2 (Adults VE), ANNEX_3 (Adolescents VE) and ANNEX 5 (KM Curves). Further detail among DAPs, subgroups, and outcomes are available in ANNEXES 2, 3, and 5.

Among **adults,** 3,807 cases of non-severe COVID-19, 8 of severe COVID-19 and none (or <5) of death with COVID-19 among individuals with homologous primary vaccination were captured in 5 data sources (all except IT-PEDIANET), versus 2,944, 17 and none (or <5) respectively among heterologous primary vaccination. Overall, homologous primary vaccinations showed a slightly decreased VE compared to heterologous primary regimens (–36%, -33%, –27% in ES-BIFAP, ES-SIDIAP and IT-INSPIRE, respectively) against **non-severe COVID-19** infection, while for UK-CPRD and NL-PHARMO the lower VE observed (-10% and –41%, respectively) was not statistically significant. This evidence varied with time since vaccination, with no clear pattern of increase or decrease in effectiveness in any of the data sources. Similarly, no pattern was seen by age categories. In ES-BIFAP, the decreased VE of homologous versus heterologous primary regimens was more marked during the Delta than the Omicron predominance period (-39% and –24%, respectively). The slightly higher different protection offered by heterologous vaccine regimens compared with homologous ones seemed to be driven by those who received a first dose of AZ, the most represented vaccine brand in this analysis. This was observed in all data sources, except for NL-PHARMO, where the VE estimate was not statistically significant.

No clear differences were observed from mRNA vaccines as the first administered doses. VE for **severe COVID-19** outcomes (hospitalization and death with COVID-19) could be only estimated in ES-BIFAP and ES-SIDIAP data sources. No differences between homologous and heterologous primary vaccine regimens were found. <5 cases of **death with COVID-19** were found in any data source, thus, no VE could be estimated for this outcome.

Among **adolescents**, 36 cases of non-severe COVID-19 among individuals with homologous versus 32 among heterologous primary vaccination were captured in three data sources (ES-BIFAP, ES-SIDIAP, and UK-CPRD). None (or <5) cases of severe COVID-19 and COVID-19 with death among adolescent pairs were found. This is due to the scarce number of heterologous schemes among this age population, for which few pairs (in ES-BIFAP, ES-SIDIAP and UK-CPRD) or even none (in Italian data sources or the Netherlands) could be identified to answer the research questions. VE estimates of homologous versus heterologous primary vaccine schemes against **non-severe COVID-19** were very unstable among databases and subgroups, with very wide 95% Confidence Intervals. This is due to the small number of individuals and cases compared. Overall, no differences were found. No VE could be calculated for **severe COVID-19 disease** and **death with COVID-19**.

10.1.1. Primary objective 1 Stratified by clinical subgroups: effectiveness of primary vaccination (adults and adolescents)

ANNEX_2 (Adults VE) and ANNEX_3 (Adolescents VE) show the analyses stratified by clinical subgroups and for patients with COVID-19 prior vaccination. We report in the text the results of subgroups or cohorts with enough size to estimate VE. For **non-severe COVID-19 disease**, the slightly higher protection offered by heterologous primary schemes was seen in the subgroup of **adult** patients with **immunodeficiency conditions** but only in ES-SIDIAP (-29%; 95% CI: -57% to –6%). When it comes to patients with **prior COVID-19 infection**, heterologous regimens appeared to be even more effective than homologous ones (-129%; 95% CI: -122% to –63%), but, again, this effect was only seen in one database (NL-PHARMO). The low number of **severe COVID-19** and **death with COVID-19** cases did not allow us to estimate VE by clinical subgroups. Among **adolescents** diagnosed with immunodeficiency, the low number of cases (<5) in the comparator groups prevented us to make any conclusion.

10.2. Primary objective 3 (children) and 4 (waning of immunity) results: effectiveness of primary vaccination in children

Among the youngest children (5-11yo), 1,569 cases of non-severe COVID-19, no (or <5) severe COVID-19 cases, and no (or <5) death with COVID-19 were captured among vaccinated, whereas 1,651, no (or <5), and no (or <5) among unvaccinated controls, respectively. Among those aged 12-14 years, 13,474 cases of non-severe COVID-19, 21 severe COVID-19 cases, and no (or <5) death with COVID-19 among vaccinated, whereas 12,908, 47, and no (or <5) among unvaccinated controls, respectively, were captured.

Among **children aged 5-14 years old**, the VE of primary vaccination against **non-severe COVID-19** was different between **Delta** and **Omicron** variant dominance periods as observed for the two DAPs whose data covered both periods (ES-BIFAP and IT-INSPIRE). Therefore, we describe the VE for these two periods separately. During the Delta predominance period, the VE against non-severe COVID-19 disease varied across databases (29% in ES-BIFAP, 49% ES-SIDIAP, 60% IT-INSPIRE, 70% in IT-Pedianet and 77% in NL-PHARMO). During the Omicron variant predominance, VE decreased to 42% in IT-INSPIRE (95% CI: 35-48%) and reverted to an increased risk among vaccinated children in ES-BIFAP (-44%; 95% CI: -49% to –39%). Regardless of the variant dominant period, in the **sensitivity analysis** restricted to **patients with negative tests before** time 0, the VE slightly increased in ES-SIDIAP (from 49% to 57%, overall) and decreased in IT-PEDIANET (from 70% to 59%, overall). On the other hand, restricting to children with **prior COVID-19 infection**, analyses showed that two doses of mRNA vaccines did not provide protection against non-severe disease across all databases, with no differences between vaccinated and controls in ES-SIDIAP and NL-PHARMO, and with even a higher risk among vaccinated children in ES-BIFAP. In particular, for children **aged 5-11** years old, VE against non-severe COVID-19 was: 70% (95% CI: 58-78%) in IT-INSPIRE, very low (8%; 95% CI: 1-14%) in ES-BIFAP, and not statistically significant (73%; 95% CI: -25% to 94%) in ES-SIDIAP.

Regarding **waning of immunity**, in those DAPs whose data covered only the Delta variant period, when most 12–13-year-olds completed their primary vaccination schemes (i.e., August- September 2021), the VE estimates against non-severe COVID-19 infection were statistically significant up to the 4-5th month after vaccination (ES-SIDIAP, IT-PEDIANET, NL-PHARMO and IT-INSPIRE). In ES-BIFAP, which also covered the Omicron period, an effectiveness of 57% was observed during the week after full vaccination, but decreased to 20% during the 3rd month, being finally lost and even reversed to a higher risk among vaccinees (EV: -64% to –10%) from the 4th to 7th month. This was in line with the reduction in effectiveness observed in this data source: from 29% during the Delta variant period to -44% during the Omicron variant period for non-severe outcomes. We should bear in mind that most infections occurred during the Omicron period, thus, the overall VE estimation did not show any protection.

VE of primary vaccination against **severe COVID-19** was 53% in ES-BIFAP (95% CI: 18-74%), slightly higher (61%), albeit non-statistically significant, during the Delta variant period, and 50% (95% CI: 5-74%) during the Omicron period. VE of 94% was found in ES-SIDIAP (95% CI: 52-99%), where only the Delta period was covered. PF vaccine was the most used brand. The pre-adolescent age group (12-14yo) was the most represented in this analysis (<5 cases occurred among children aged 5-11yo). Therefore, the overall VE is approximated in these two subgroups. The small number of severe cases did not allow the estimation of the waning of immunity after vaccination in the 5-11 years old population. In the **sensitivity analysis** restricting to patients with prior COVID-19 negative tests, only ES-SIDIAP confirmed a VE of 90% [broad confidence interval (95% CI: 23-99%)], while, in ES-BIFAP, a non-statistically significant 52% VE was estimated. Effectiveness estimates could not be calculated due to the very low number of cases in IT-INSPIRE or among pairs with COVID-19 infection prior to vaccination in all DAPs.

10.2.1. Primary objective 3 (children) stratified by clinical subgroups: effectiveness of primary vaccination in children 5-14 years old

Vaccine effectiveness and Kaplan-Meier curves calculated in the main analysis (i.e., among pairs without prior COVID-19) are displayed in ANNEX 4 (Children VE), and Meta Analysis-Results section. Further details among DAPs, subgroups, and outcomes are available in ANNEX 4. We only report in the text the results of subgroups or cohorts with enough size to estimate VE. The reported VE refers to statistically significant estimators unless the opposite is stated.

Among children (aged 5-14 years) diagnosed with **immunodeficiency,** the VE of full primary vaccination:

- Ranged, across databases, from 8% (95% CI; 1-14%) in ES-BIFAP to 73% (95% CI; 53-85%) in NL-PHARMO against **non-severe COVID-19**. The highest protection (80%), albeit not statistically significant, was observed in IT-PEDIANET.
- Was 85% (95% CI; 27-97%) in ES-BIFAP and non-statistically significant 74% (95% CI; -133% to 97%) in ES-SIDIAP against **severe COVID-19.**

Among children (aged 5-14 years) diagnosed with **cancer**, the VE of full primary vaccination was:

- Non-statistically significant 19% (95% CI; -70% to 61%) in ES-BIFAP, and 95% (95% CI; 62-99%) in ES-SIDIAP against **non-severe COVID-19 infection.**

VE estimated in other clinical subgroups (e.g., severe renal disease), or death with COVID-19 could not be calculated due to the lack of enough COVID-19 episodes for the analysis.

10.3. Primary objective 2 (boosting in adults and adolescents) and exploratory objective 5 (waning of immunity after booster) results

10.3.1. Primary objective 2 (boosting), and exploratory objective 5 (waning of immunity after booster) results against non-severe COVID-19

Vaccine effectiveness and Kaplan-Meier curves calculated in the main analysis (i.e., among pairs without prior COVID-19) are displayed in ANNEX_2 (Adults VE), ANNEX_3 (Adolescents VE) and ANNEX 5 (KM Curves). Further detail among DAPs, subgroups, and outcomes are available in ANNEXES 2, 3 and 5.

A total of 79,076 cases of **non-severe COVID-19** among **adults** who received any booster versus 138,638 among unboosted comparators were captured in 5 data sources (IT-PEDIANET does not have adult data). The VE of booster could be estimated for all three types of primary schemes and brands. To avoid repetitions, the VE estimations mentioned below are all statistically significant unless the opposite is stated. The VE of a booster broadly varied by data source as follows for three similar vaccinations, only similar two first doses and different two first doses, respectively: 28%-69%, 31%-62%, 34%-71% for PF booster; 41%-74%, 44%-71%, 47%-69% for MD booster). The highest VE was estimated for PF (in Italy) and for MD (in Italy and The Netherlands). Data from Spain gave the lower VE (≤58%). The VE of AZ booster was only observed in the UK (90%) after two doses of AZ. NL-PHARMO did not show effectiveness when AZ booster was administrated following any homologous scheme (low analysis power). Other data sources had <5 cases in at least one of the compared groups, so no VE was produced. In Italy, a slightly higher significant VE was suggested during the **Omicron** (70% and 73%) in comparison with the **Delta** (66% and 65%) period for homologous primary vaccinations receiving a (homo- or hetero-) booster. The opposite is observed for heterologous primary vaccinations (79% in Omicron vs 74% in Delta). In Spain, the VE was 58-60% during Delta and decreases to 19-36% during Omicron, especially among those receiving three similar doses (19%). No information during the Omicron period was reached for the UK or The Netherlands.

Most of the **boosted** patients **with a heterologous primary vaccination** initiated with AZ benefited from: 71% VE in IT-INSPIRE (similar to PF or MD booster) and 41-58% in Spanish sources (63% and 34% PF, and 47% and 57% MD, in ES-SIDIAP-ES and BIFAP, respectively). Other data sources (NL-PHARMO and UK-CPRD) had low incidences and did not show statistically significant VE. Any **booster among homologous primary schemes** showed a VE of 42-58% in data from Spain and the Netherlands. In Spain, VE decreased to 36% during **Omicron**. In IT-INSPIRE, VE was estimated during both periods (65% Delta; 73% Omicron). **Homologous booster vaccination** with AZ showed VE (90%) only in the UK, VE with PF was significant in four DAPs (not in NL-PHARMO) ranging from 28% (ES-BIFAP) to 69% (IT-INSPIRE), and VE with MD is statistically significant with enough power in all DAPs except UK-CPRD, ranging from 41% (ES-BIFAP) to 74% (NL-PHARMO). VE in patients with **homologous primary vaccinations and any heterologous booster** ranges from 42% to 58% in Spain and the Netherlands, respectively, and was 70% in Italy. No UK data are available for this cohort. The **VE lasted** until the 4th month (IT-INSPIRE; ES-BIFAP) and the 3rd month (ES-SIDIAP) after booster administration. In ES-BIFAP, VE decreased during the 3rd month, with an increment of infection risk among the vaccinated. In Italy and the UK, VE was similar in all **age groups** but increased to 78% (UK-CPRD) and decreased to 61% (IT-INSPIRE) among those

aged 80+. In ES-BIFAP, VE is low in all age groups (27% for 80+ years old). VE in ES-SIDIAP fluctuated among age groups, achieving 45% for 80+ years old. In **sensitivity analysis** (ES-SIDIAP and BIFAP) restricted to individuals tested at least once before the matching date, the VE of any booster did not change, remaining between 30-55%. Here, VE disappeared after 4 months from vaccination (fluctuating in PHARMO). **Booster doses** in patients **with heterologous primary vaccinations** starting with AZ (i.e., AZ(1)-mRNA(2)-mRNA(3) platforms), showed significant VE (39% ES-BIFAP; 51% ES-SIDIAP; 69% IT-INSPIRE) in comparison with no-booster. This effect was already evident during the first week after booster vaccination and lasted until 1st month after the booster dose without observed VE differences. Not statistically significant VE was observed with a booster when heterologous primary vaccinations started with PF (46%) or MD (54%). The small cohort size did not allow the VE calculation in UK-CPRD and NL-PHARMO.

408 cases of **non-severe COVID-19** among **adolescents** who received any booster versus 936 among unboosted comparators were captured in 5 data sources (ES-BIFAP, ES-SIDIAP, IT-INSPIRE, UK-CPRD and NL-PHARMO). Among adolescents with a **homologous primary vaccination**, the VE of a **homologous booster** dose (mainly PF) against non-severe infection varied from 67% (IT-INSPIRE; using only PF) to 38% (ES-BIFAP; 64% for MD and 35% for PF). The VE of a **heterologous booster** (mainly MD) was estimated from four data sources. Enough cases for VE estimation were present in ES-BIFAP, with an overall VE of 48% (95% CI: 22-65%). Considering the **vaccine brand**, adolescents vaccinated with PF (doses 1 and 2), who then received an MD booster, were the highest in number, showing a 48% VE. This estimation was similar for adolescents vaccinated with MD (doses 1 and 2) and boosted with PF (51%) but did not reach statistical significance. During **Delta** predominance, the VE of the booster was only observed in IT-INSPIRE (69%; 95% CI: 28-87%). During **Omicron** variant predominance, the VE was 67% (IT-INSPIRE) and 44% (ES-BIFAP) for the homologous booster, and 51% (ES-BIFAP) for the heterologous booster (with a homologous primary scheme). The evaluation of **VE by time** after the booster dose administration could only be observed for homologous booster schemes, ranging from the first week after booster vaccination (71% in Italy and 63% in Spain) until 1 month in Italy (75%; later on <5 cases occurred) or 2 months in Spain (45%). After, IRs were not different between boosted and non-boosted groups in BIFAP-ES. The **sensitivity analyses** restricted to individuals with a prior test in ES-BIFAP did not show statistically significant VE. Boosted adolescents with a **heterologous primary vaccination** were only identified in UK-CPRD and could not be analysed (i.e., 22 person-days).

10.3.2. Primary objective 2 (boosting), and exploratory objective 5 (waning of immunity after booster) results against severe COVID-19

Vaccine effectiveness and Kaplan-Meier curves calculated in the main analysis (i.e. among pairs without prior Covid-19) are displayed in ANNEX_2 (Adults VE), ANNEX_3 (Adolescents VE) and ANNEX 5 (KM Curves). Further detail among DAPs, subgroups, and outcomes are available in ANNEXES 2, 3, and 5.

1,015 cases of **severe Covid-19** among **adults** who received **any booster** versus 3,362 among comparators without 3rd dose were captured in 3 data sources with hospitalization information. Cases were mostly identified in ES-BIFAP and SIDIAP and only a few in IT-INSPIRE. **All cases happened in the homologous primary vaccination cohort**. Overall, the adjusted VE of receiving any booster dose was ≥61% (statistically significant). It was 80% (for homologous booster) in Italy, whereas 61-67% (for homologous booster) and 75-79% (with heterologous booster) in Spanish databases (ES-BIFAP and SIDIAP). In Italy, only the overall VE could be estimated due to insufficient statistical power for stratification analyses. In Spain, by **vaccine brand**, we found that:

- adults with two doses of PF showed: 64-67% VE with a 3rd PF dose and 74-78% VE with an MD booster.
- adults with two doses of MD showed: 42-65% VE with a 3rd MD dose and 73-78% VE with PF booster;
- adults with two doses of AZ had: 76-81% VE from receiving any mRNA platform (PF or MD) as booster dose.

AZ booster doses were not sufficient to be analysed. The VE of any booster was estimated for both the **Delta** and **Omicron periods** in ES-BIFAP (68% and 67% for a homologous booster, respectively; and 77% and 74% for a heterologous booster, respectively). In ES-SIDIAP, the VE (61% for a homologous and 79% for a heterologous booster) referred only to Delta as no follow-up during

Omicron was covered. **By age groups**, the VE of homologous boosters was mainly observed for ≥50 years old pairs in ES-BIFAP and ≥70 years old pairs in ES-SIDIAP. Most of the severe cases of COVID-19 were registered in ES-SIDAP. VE of heterologous boosters showed a trend to decrease from 50-59 years old (90% VE in ES-SIDIAP and 73% in ES-BIFAP) to >80 years old (67% VE in ES-SIDIAP and 66% VE in ES-BIFAP). **Over time**, the beneficial effect appeared from the first week (which could reveal some undetected bias) until the 1st (in ES-SIDIAP) or 3rd (in ES-BIFAP) month after vaccination with any heterologous booster (PF or MD). The VE lasted until the 2nd (ES-SIDIAP) or 5th (ES-BIFAP) month after vaccination with any homologous booster dose**.** The significant effectiveness was confirmed by **sensitivity analysis** after restricting to individuals with negative tests before comparison (VE between 55-59% for homologous booster and 70-81% for heterologous booster, with the Spanish data).

Among **adolescents**, we observed 88,892 person-days of follow-up in ES-SIDIAP, ES-BIFAP and IT-INSPIRE. No (or <5) cases of severe COVID-19 were identified in either group (boosted or non-boosted), so no VE could be estimated.

10.3.3. Primary objective 2 (boosting), and exploratory objective 5 (waning of immunity after booster) results against death with COVID-19

Vaccine effectiveness and Kaplan-Meier curves calculated in the main analysis (i.e. among pairs without prior COVID-19) are displayed in ANNEX_2 (Adults VE), ANNEX_3 (Adolescents VE) and ANNEX 5 (KM Curves). Further details among DAPs, subgroups, and outcomes are available in ANNEXES 2, 3 and 5.

313 cases of death with COVID-19 among adults who received any booster versus 1,367 among comparators without that 3rd dose were captured in the 2 Spanish data sources, NL-PHARMO and UK-CPRD. Cases were mostly identified in ES-BIFAP and ES-SIDIAP and a few in UK-CPRD. NL-PHARMO had <5 cases. Numbers from the heterologous primary vaccination cohort were not enough to produce VE (only 422,000 person-day of follow-up for each compared group). **All cases were observed in the homologous primary vaccination** cohort, and most of them received a homologous booster. Independently from the booster dose type, the VE against COVID-19 with death among adults was ≥74% (statistically significant), ranging (by data source) from 74% to 80% for **homologous booster** and from 82% to 86% for **heterologous booster**. VE for complete homologous schemes ranged from 76% to 80% during **Delta** (Spanish data sources and UK-CPRD) and was 72% during **Omicron** only for ES-BIFAP (no other data source covered the Omicron period here). VE for heterologous booster was 80% and 86% during **Delta** (in the two Spanish data sources) and 83% during Omicron (only in ES-BIFAP). Considering the vaccine brands, the VE for the booster homologous scheme ranged (by data source) from 72% to 79% with PF and was 88% (only ES-BIFAP data) with MD when compared with the no boosted from the homologous primary scheme. Those VE were observed only for people aged ≥60 years old. The confidence intervals did not show differences among the 60-69, 70-79, and ≥80 **years old groups**. Regardless of the **brand** of the primary schedule, a booster dose with MD, after PF or AZ as primary scheme doses, showed from 83% to 86% VE in Spanish data sources, whereas a booster dose with PF, after MD or AZ as primary doses, showed 77% VE only in ES-BIFAP. People with AZ as the booster dose had <5 cases, insufficient for VE estimation. Among adults with AZ as primary doses, a booster dose with any mRNA vaccine (PF-MD) showed a VE of 80% and 86% (non-statistically significant; based on 8 cases among unboosted and <5 cases among boosted pairs). The effectiveness was lost at the 6th month and 5th month after homologous and heterologous booster vaccinations, respectively, in ES-BIFAP. Other data sources contributed to shorter **follow-up in terms of time.** Statistically significant VE could only be observed until the 2nd month after homologous and the 2nd week after heterologous booster vaccinations in ES-SIDIAP, and, until the 1st week after homologous booster in UK-CPRD (i.e. in time-windows with <5 cases in at least one compared group). The statistically significant VE was confirmed in **sensitivity analysis** in ES-BIFAP and -SIDIAP, after restricting to individuals with negative COVID-19 tests before the matching date: 67% and 79% for homologous booster and 81% and 77% for heterologous booster, respectively. **Over time**, the beneficial effect appeared from the 1st week after vaccination, which could reveal some unmanaged bias.

Among **adolescents**, neither in this case, we could produce any VE for any booster schedules in the participating data sources (no [or <5] cases of death with COVID-19).

10.3.4. Primary objective 2 (boosting) by clinical subgroups

Vaccine effectiveness and Kaplan-Meier curves calculated in the main analysis (i.e., among pairs without prior Covid-19) are displayed in ANNEX_2 (Adults VE), and Meta Analysis-Results section. Further details among DAPs, subgroups, and outcomes are available in ANNEXE 2. We only report in the text the results of subgroups or cohorts with enough size to estimate VE, otherwise the opposite is stated.

10.3.4.1. Primary objective 2 (boosting) by clinical subgroups, adults

In clinical subgroups, only homologous primary vaccination contributed to the booster analyses.
Among **adults** diagnosed with **immunodeficiency,** the VE of a 3rd dose (homologous or heterologous) varied by data source, outcome and type of booster:

- 68-72% VE against **non-severe COVID-19** in IT-INSPIRE, 69% in UK-CPRD, 31-55% in the Spanish DAPs, and 41% in NL-PHARMO for heterologous 3rd doses. VE was slightly lower (Italy and Spain) or not effective (The Netherlands) when a homologous 3rd dose was administrated.
- 60-63% (homologous booster) and 72-73% (heterologous booster) VE against **severe COVID-19** in Spanish data sources, and 78% borderline in IT-INSPIRE (homologous booster; with not enough sample size for heterologous booster evaluation).
- 71-81% against **death with COVID-19** in Spanish DAPs, for both, homo- and hetero-logous boosters.

Among adults diagnosed with **cancer,** VE was:

- against **non-severe COVID-19:** 65-70% in Italy, 54% in the UK, and 27-56% in Spain, independently from the 3rd dose. Non-significant in the Netherlands.
- against **severe COVID-19:** 61-77% (ES-SIDIAP) and 49-54% (ES-BIFAP) for homologous and heterologous 3rd doses, respectively. In IT-INSPIRE, the number of outcomes was scarce for VE estimation.
- against **death with COVID-19**: 75-83% in Spanish data sources (slightly higher for heterologous booster). No (or <5) cases in UK-CPRD or NL-PHARMO.

Among patients diagnosed with **severe renal disease**, VE was:

- against **non-severe COVID-19**: 62% in ES-BIFAP; only for homologous 3rd doses. No data was observed in the other four DAPs.
- against **severe COVID-19**: 75% in ES-BIFAP for homologous 3rd doses, and 29% (non-statistically significant) in ES-SIDIAP or in IT-INSPIRE (scarce population was for VE estimation).
- against **death with COVID-19**: 90% in ES-BIFAP and 42% in ES-SIDIAP (non-significant) for homologous 3rd doses.

Patients with **transplants** were identified only in Spanish data sources, where VE could not be estimated due to the low number of cases for any outcome or did not show statistical significance (i.e., for homologous booster against severe COVID-19).

Patients with **Down Syndrome** were identified in the Spanish or Italian data sources. VE against non-severe COVID-19 was 78% in ES-SIDIAP but non-statistically significant, 2% in ES-BIFAP and 55% in IT-INSPIRE. Not enough severe COVID-19 cases or death with COVID-19 occurred for any VE estimation.

Matched pairs with **heterologous primary vaccinations** were scarce: no severe COVID-19 or death with COVID-19 was identified to estimate VE in clinical subgroups. The VE against non-severe COVID-19 was 59% in ES-SIDIAP, 72% in IT-INSPIRE, and non-statistically significant 26% in BIFAP-ES among patients with **immunodeficiency**, or non-statistically significant 42% in IT-INSPIRE for patients with a **cancer** diagnosis.

Among patients with **COVID-19 infection prior to vaccination:**

- against **non-severe COVID-19 infection:** heterologous booster VE was 33% (ES-SIDIAP) and 38% (NL-PHARMO) but was not observed in ES-BIFAP (where increased risk was observed among 3rd doses-patients; HR: 1.27, 95% CI: 1.20-1.34). In the five participating data sources, we were able to estimate incidences, but none showed VE of the homologous 3rd doses.

- against **severe COVID-19**: the heterologous booster showed higher VE (69%-71%) than homologous 3rd doses (43% in ES-SIDIAP and 32% -non statistically significant- in ESBIFAP) in comparison with their unboosted pairs.
- Against **death with COVID-19:** the VE was 92% for heterologous (ES-BIFAP) and 68% for homologous 3rd doses (ES-SIDIAP; and 34% -non statistically significant- in ES-BIFAP). In UK-CPRD or NL-PHARMO, none (or<5) cases were observed.

10.3.4.2. Primary objective 2 (boosting) by clinical subgroups, adolescents

Among **adolescents**, only cases of **non-severe COVID-19** occurred. Among adolescents diagnosed with **immunodeficiency**, the VE of a 3rd dose against **non-severe COVID-19** was only observed for a homologous schedule in INSPIRE-IT (VE: 65%). In the two Spanish DAPs, only incidence rates could be calculated (especially among homo- and heterologous 3rd doses).

No effectiveness against **non-severe COVID-19** was observed for pairs with **COVID-19 prior to vaccination**. Caution for interpretations is required for subgroups with a small sample size. Other subgroups were too small for VE estimations.

10.4. Outcome data (included in effectiveness sections)

The overall number of the COVID-19 outcomes (non-severe infection, severe COVID-19 disease, and death with COVID-19) and all-cause death episodes are included in each corresponding 'Main results' for effectiveness section, and detailed by data source (including their incidence rates; IR) and matched cohorts in ANNEX 2 (Adult VE), ANNEX 3 (Adolescents VE), ANNEX 4 (Children VE), and ANNEX 5 (KM Curves).

10.5. Meta-analysis

Meta-analyses are reported below with detailed pooled adjusted HR and VE with sufficient severe COVID-19 and death with COVID-19 episodes to be analyzed. Meta-analyses refer to patients without previous COVID-19 infection.

Meta-analysis was not planned for Primary objectives 1 (adults and adolescents) and 4 (waning of immunity).

Meta-analysis for Primary objective 2 (boosting) was planned per protocol for rare outcomes (i.e., severe COVID-19 or Death with COVID-19) among patients in the following clinical subgroups with reduced sample sizes:

Among **adults with homologous primary vaccination**, the pooled VE of **homologous booster against severe COVID-19** was 62% (95% CI: 57 to 67%; I^2=0%) among patients with **immunodeficiency** or immunosuppressant, 54% (95% CI; 41 to 64%; I^2=18%) among patients with **cancer**, 24% (95% CI; -54 to 63%; I^2=0%) among patients receiving **transplants**, and 57% (95% CI; -20 to 84%; I^2=65%) among patients with severe renal disease.

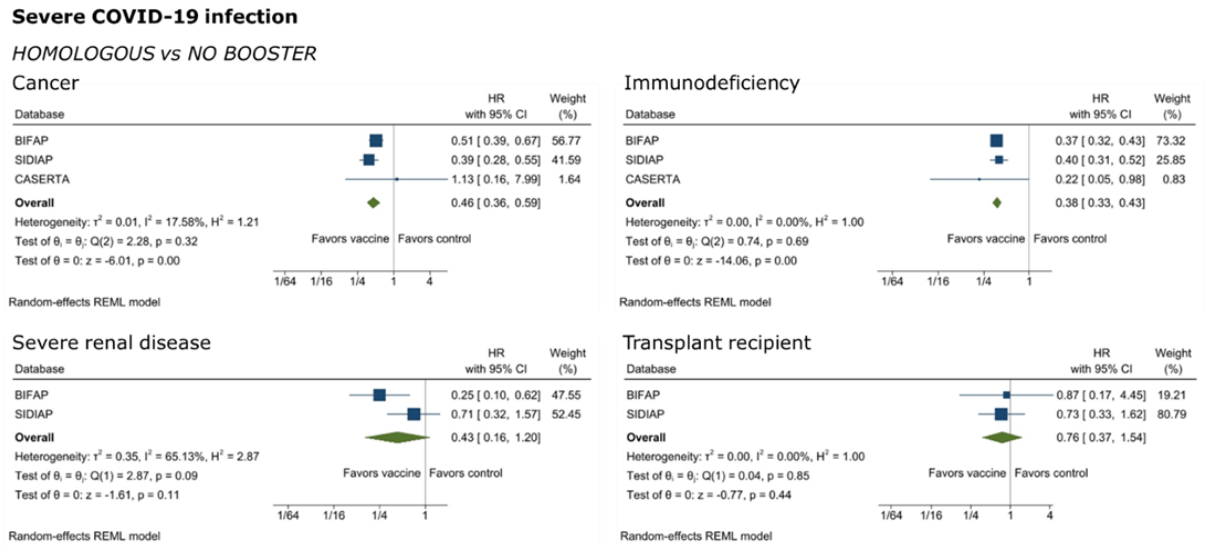

Figure 7. Forest plot for HR of severe COVID-19 among booster versus no booster among adults by clinical subgroups.

Table 7. Adjusted HR and VE (1 minus the HR)

Comparison and population	Adjusted HR (95% CI)	Adjusted VE (95% CI)
Homo-primary + Homo-Booster vs Homo-primary + No Booster among adults		
Patients with immunodeficiency or immunosuppressant	0.38 (0.33-0.43)	62% (57 to 67%)
Patients with cancer	0.46 (0.36-0.59)	54% (41 to 64%)
Patients receiving transplants	0.76 (0.37-1.54)	24% (-54 to 63%)
Patients with severe renal disease	0.43 (0.16-1.2)	57% (-20 to 84%)
Patients with Down Syndrome	no data	

In addition, the pooled VE of **homologous booster against death with COVID-19** was 73% (95% CI; 63 to 80%; I^2=15%) among patients with immunodeficiency or immunosuppressant, 75% (95% CI; 65 to 82%; I^2=0%) among patients with cancer, and 75% (95% CI; -38 to 96%; I^2=63%) among patients with severe renal disease. ES-BIFAP and -SIDIAP databases contributed to all estimators, while IT-INSPIRE and UK-CPRD contributed to two meta-analyses each on severe COVID-19 and death with COVID-19, respectively.

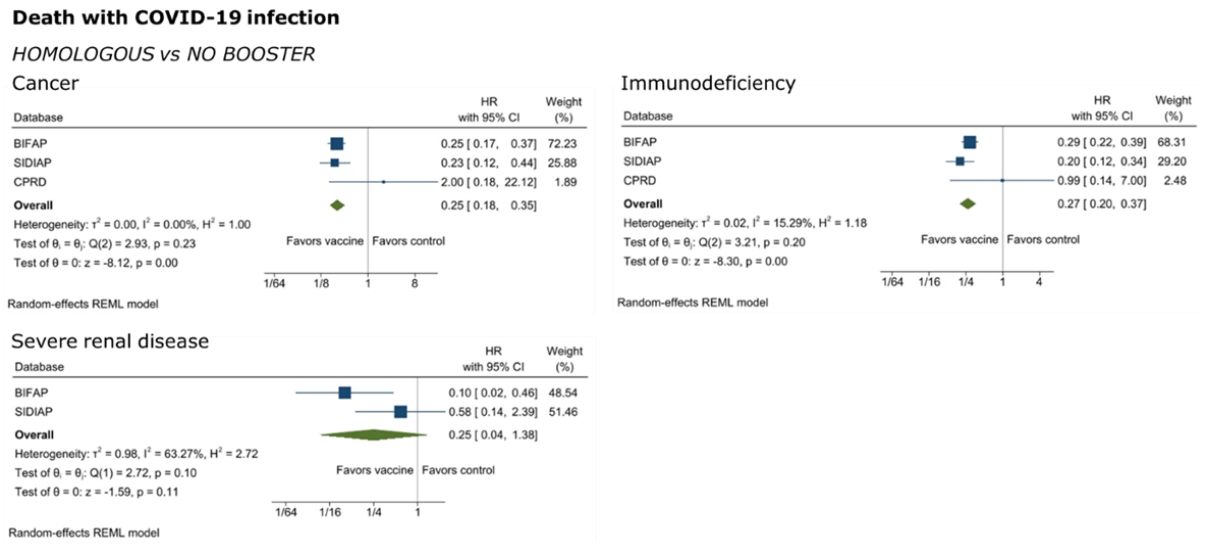

Figure 8. Forest plot for HR of death with COVID-19 among booster versus no booster among adults by clinical subgroups.

Table 8. adjusted HR and VE (1 minus the HR).

Comparison and population	Adjusted HR (95% CI)	Adjusted VE (95% CI)
Homo-primary + Homo-Booster vs Homo-primary + No Booster among adults		
Patients with immunodeficiency or immunosuppressant	0.27 (0.20-0.37)	73% (63 to 80%)
Patients with cancer	0.25 (0.18-0.35)	75% (65 to 82%)
Patients receiving transplants	no data	
Patients with severe renal disease	0.25 (0.04-1.38)	75% (-38 to 96%)
Patients with Down Syndrome	no data	

Among **adults with homologous primary vaccination**, the pooled VE of **heterologous booster against severe COVID-19** was 72% (95% CI; 66 to 77%; I^2=0%) among patients with **immunodeficiency** or immunosuppressant, and 68% (95% CI; 36 to 84%; I^2=77%) among patients with **cancer.** Among patients with **severe renal disease,** VE was only available for ES-BIFAP so meta-analysis is not applicable.

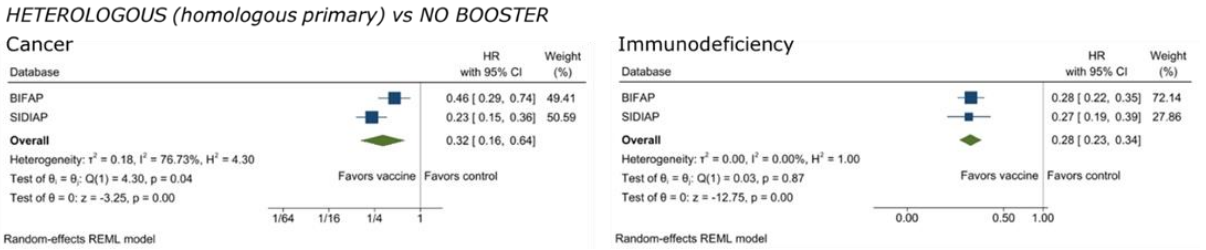

Figure 9. Forest plot for HR of severe COVID-19 among heterologous booster versus no booster among adults with homologous primary vaccination by clinical subgroups.

Table 9. Adjusted HR and VE (1 minus the HR).

Comparison and population	Adjusted HR (95% CI)	Adjusted VE (95% CI)
Homo-primary + Hetero-Booster vs Homo-primary + No Booster among adults		
Patients with immunodeficiency or immunosuppressant	0.28 (0.23-0.34)	72% (66 to 77%)
Patients with cancer	0.32 (0.16-0.64)	68% (36 to 84%)
Patients receiving transplants	no data	
Patients with severe renal disease (only BIFAP-ES)	0.67 (0.16-2.87)	33% (-187 to 84%)
Patients with Down Syndrome	no data	

In addition, the pooled VE of **heterologous booster against death with COVID-19** was 80% (95% CI; 70 to 86%; I^2=0%) among patients with **immunodeficiency** or immunosuppressant, and 81% (95% CI; 70 to 89%; I^2=0%) among patients with **cancer**. ES-BIFAP contributed to all estimators, while ES-SIDIAP contributed to all meta-analyses except for severe COVID-19 in patients with severe renal disease.

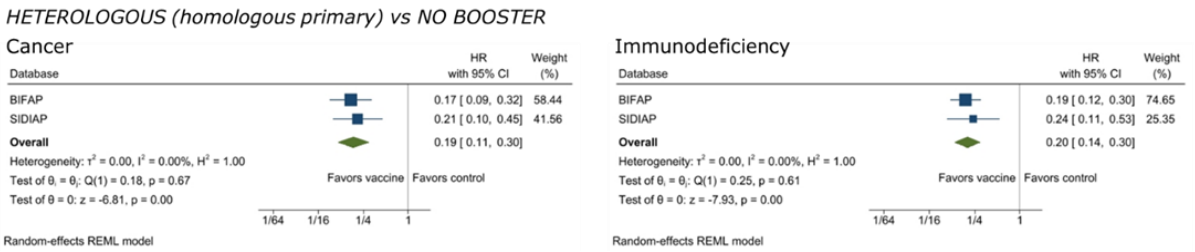

Figure 10. Forest plot for HR of death with COVID-19 among heterologous booster versus no booster among adults with homologous primary vaccination by clinical subgroups.

Table 10. Adjusted HR and VE (1 minus the HR).

Comparison and population	Adjusted HR (95% CI)	Adjusted VE (95% CI)
Homo-primary + Hetero-Booster vs Homo-primary + No Booster among adults		
Patients with immunodeficiency or immunosuppressant	0.20 (0.14-0.30)	80% (70 to 86%)
Patients with cancer	0.19 (0.11-0.30)	81% (70 to 89%)
Patients receiving transplants	no data	
Patients with severe renal disease	no data	
Patients with Down Syndrome	no data	

Among **adults with heterologous primary vaccination**, no data was retrieved to be meta-analyzed. Among **adolescents and patients with Down syndrome**, an insufficient number of COVID-19 outcomes were reached to be meta-analyzed.

Meta-analysis for Primary objective 3 (children) was planned per protocol for rare outcomes (i.e., severe COVID-19 or death with COVID-19):

Among **children and pre-adolescents** aged 5-14 years old, the pooled adjusted VE of full vaccination against **severe COVID-19** was 79% (95% CI; -50 to 97%) with substantial heterogeneity (I^2=75%). When results were restricted to participants during the Delta variant period, the pooled adjusted VE of full vaccination against severe COVID-19 was 82% (95% CI; -10 to 97%), also with substantial heterogeneity (I^2=62%). ES-BIFAP and -SIDIAP contributed to both meta-analyses. Regarding the Omicron variant period, only ES-BIFAP data was available and meta-analysis is not applicable. No data on death with COVID-19 for children was available.

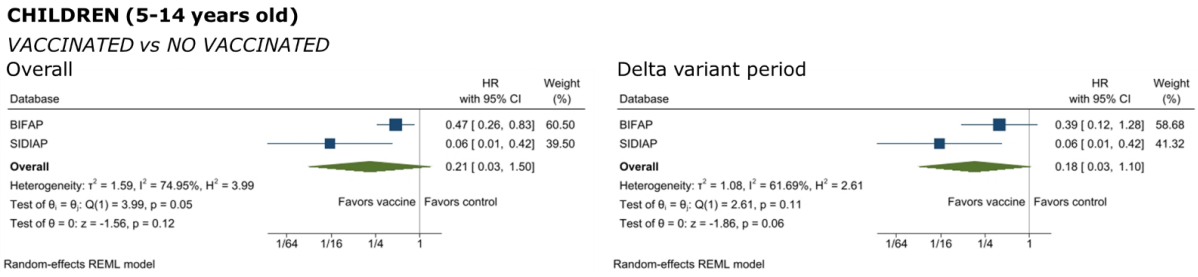

Figure 6. Forest plot for HR of severe COVID-19 among homologous primary vaccination versus non-vaccination in children (5-14 years old) overall Delta and omicron periods (left side) and in Delta period only (right side).

Table 6. Adjusted HR and VE (1 minus the HR).

Comparison and population	Adjusted HR (95% CI)	Adjusted VE (95% CI)
Homo-primary vs Non-vaccinated among children		
Overall	0.21 (0.03-1.50)	79% (-50 to 97%)
Delta	0.18 (0.03-1.10)	82% (-10 to 97%)
Omicron (only BIFAP-ES)	NA	NA

10.6. Secondary analyses (any cause of death)

The VE of any booster dose (whether homologous or heterologous) after a homologous primary vaccination ranged between 72% and 96% among 10-by-10 age categories and data sources (ES-BIFAP and -SIDIAP, UK-CPRD and NL-PHARMO) among patients aged >60 years old (ANNEX 2, Adults VE). VE was higher during the Delta (75% for homologous and 83% for heterologous booster) than the Omicron variant period (67% for both booster types). Moreover, a few heterologous primary vaccinations were used among patients aged: a) 60-69 years old in ES-BIFAP, b) 70-79 years old in UK-CPRD, and c) 80+ years old in UK-CPRD. Only this last group showed VE (74%).

10.7. Representativeness of the matched participants

In this section, we describe the full primary vaccination population identified in the data sources from which the matched cohort groups were selected. We aim to present an overall interpretation of the population represented in the effectiveness estimations (before matching).

10.7.1. Adults and Adolescents with complete primary vaccination before matching

The full primary COVID-19 vaccination study population aged **≥12 years old**, contained more than 20 million individuals from 6 healthcare databases of 4 European countries: Spain, Italy, Netherlands, and the United Kingdom:

- 7.7 million Homologous & 160,513 Heterologous in ES-BIFAP
- 3.7 million Homologous & 41,710 Heterologous in ES-SIDIAP
- 616,954 Homologous & 34,585 Heterologous in IT-INSPIRE
- 1,737 Homologous & 5 Heterologous in IT-PEDIANET
- 842,582 Homologous & 52,543 Heterologous in NL-PHARMO
- 6.9 million Homologous & 118,603 Heterologous in UK-CPRD

Table A1 (ANNEX 1) shows the number of included individuals ≥12 years old across countries and databases, details about the total numbers of included primary homologous and heterologous vaccinations, vaccine brands, sex, age at time 0, and databases-available COVID-19 severity risk factors before matching. Across all the countries, we encountered more than **400,000 heterologous (2%) and 19.8 million homologous (98%) primary scheme**-vaccinated individuals before matching. Among them, around 4% were aged 12-17 years old. Data from Spain (BIFAP and SIDIAP) has the largest number of both heterologous and homologous vaccinated individuals, followed by the United Kingdom (CPRD), the Netherlands (PHARMO), and Italy (INSPIRE and PEDIANET). The most prevalent heterologous scheme was with the AZ vaccine as the first administered dose in ES-SIDIAP (72%), Italy (75%), and the United Kingdom (86%). Differently, in the Netherlands and ES-BIFAP, 78% and 50% of the heterologous schemes have the mRNA PF vaccine as the first administered dose, respectively. The use of the MD vaccine as the first administered dose resulted to be the lowest among all countries in both homologous and heterologous schemes (first and second dose). In all the countries, except for the UK, the most prevalent homologous scheme has PF (73-77%). The UK has AZ (53%) as the most frequently administered homologous vaccine, with PF administered in 44% of individuals. The gender distribution was almost 50% among all countries and all the different homologous and heterologous schemes herein presented, with a slight unbalance, not overpassing 56%, in favor of females. **Across the total population, 94% did not have encountered SARS-CoV-19 infection prior to vaccination**. The prevalence of COVID-19 infection before vaccination ranged 1.8-6.9% from homologous (main population) and 0-62% from heterologous primary vaccination cohorts across the data sources.

10.7.2. Children with complete primary vaccination before matching

Among **children** (**5-11 years old** when entering the study population and **12-14 years old** at the vaccination date), the full primary COVID-19 vaccination study population contains approximately 308,000 individuals with around 3 months of follow-up:

- 192,421 Homologous and 1,061 Heterologous in ES-BIFAP
- 79,029 Homologous and 113 Heterologous in ES-SIDIAP
- 14,570 Homologous and <5 Heterologous in IT-INSPIRE
- 1,599 Homologous and 6 Heterologous in IT-PEDIANET
- 17,945 Homologous and <5 Heterologous in NL-PHARMO
- 1,400 Homologous and 5 Heterologous in UK-CPRD

The latest available data lock dates for the participating data sources allowed to include low numbers of fully (two doses) vaccinated children. Tables A12 (ANNEX 1) show the total numbers of included homologous (majority) and heterologous (residual) vaccination schemes by brands (>90% of inclusions had PF as 1st administered dose, MD vaccine was residual ≈10%), and COVID-19 severity risk factors. The majority of children and pre-adolescents were fully vaccinated in January and February 2022 (>69%). UK-CPRD has the majority of vaccinated children in October-December 2021. A peak of 26% and 59% administered doses were also observed in September 2021 for ES-BIFAP and ES-SIDIAP, respectively, and of ≈50% in August 2021 for NL-PHARMO and IT-PEDIANET. These doses might represent vaccinated pre-adolescents. Between 11-32% of the vaccinated children before matching, for Spain and Italy, respectively, present a diagnosis of immunodeficiency (overall, ≥ 44,500 children). **Most of the vaccinated children (>95%) were free of prior COVID-19**. Overall, the highest amount of vaccinated (>77%) registered as aged 12-14 years old that were 5-11y at the study start date.

11. Discussion

In this study, using real-world data from 6 different databases in Europe, various schemes of COVID-19 vaccination (primary vaccination and booster doses) have been identified among children, adolescents, and adults in Spain, Italy, the Netherlands, and the UK from the beginning of the vaccination campaign to February 2022.

Most of the **adult population** participating in our study have been fully vaccinated with the primary scheme from mid of 2021. Adults also received booster doses in the last months of 2021. These results are in line with the availability of vaccine brands in Europe over the past year, public health recommendations, and publicly available information on vaccination coverage (24). The majority (>75%) of primary COVID-19 vaccinations and booster doses, among adults, consisted of homogeneous schemes, in all countries, with a good sample size of PF but also AZ, mainly due to the inclusion of the UK and the Netherland, and the matching designed to condition on the brand of the 1st dose. As expected, we obtained a **booster vaccines coverage in the order of millions**, which allowed us to perform analyses of their immediate effectiveness with enough power, comparing boosted individuals with those who had not received any booster dose at that time.

Our data registered that most of **the adolescent population** completed the vaccination during the Autumn of 2021. During this period, mRNA vaccines were the only available/approved for such populations. Thus, most of them belong to the homogeneous vaccination scheme cohorts, whereas heterologous vaccinees were marginally present in the participating data sources. Some adolescents also received booster doses during the last month of 2021-beginning of 2022.

For **children and pre-adolescents**, data from Spain, Italy and The Netherlands showed that they could contribute to the follow-up for the COVID-19 outcomes for approximately 2-3.5 months. The increase in the distribution of administered vaccinations in August-September 2021 may refer to pre-adolescent individuals, being the youngest children vaccinated from December 2021 to February 2022. Again, our data refers to mRNA vaccines for such populations.

In the current study, a relevant sample size of patients diagnosed with **immunosuppression** or under immunosuppressant drugs or with prior **cancer** diagnosis was obtained. Those are less represented in randomized control trials (RCT) and small observational studies, especially among young patients. We could estimate the effectiveness of primary vaccination and 3rd doses (in some data sources).

In this study, the designed matching criteria (see section 9.3.1) have been maximized to reach the best possible comparable situation among individuals compared with different vaccine schemes or not vaccinated. This design permits gaining an excellent balance in terms of different comedications and comorbidities captured among the matched cohorts as potential confounders. This is a solid starting point for the statistical analyses and final interpretations.

11.1. Key results

In **primary objectives 1 (adults and adolescents), and 4 (waning of immunity)**, **no differences between primary homologous and heterologous schemes were observed against severe COVID-19 among adults**, although that could only be assessed in Spain due to the reduced sample size in the Italian database INSPIRE. At the start of the vaccinations, little evidence was available on the comparative effectiveness of the heterologous and homologous primary vaccination schedules against severe COVID-19 and death with COVID-19. Herein, we provide results for severe COVID-19, although no robust VE estimations could be produced due to the low number of cases in each comparison group. The differential effectiveness against death with COVID-19 could not be tested due to zero cases observed. In general, small cohorts could be matched due to the low use of heterologous primary vaccinations. **Against non-severe COVID-19, we observed slightly higher effectiveness of heterologous versus homologous primary vaccinations** among patients who initiated with an AZ dose and received a 2nd mRNA dose in comparison with receiving two AZ doses. The VE against non-severe COVID-19 was not different

among people who received two different mRNA doses. This was observed **in all the countries studied** and was similar to small clinical trials (22,23), whose endpoint was immunogenicity, and other previous observational studies (24–28).

Under the scenario of a reported very high level of protection with the COVID-19 vaccination against infection, symptomatic disease, and severe disease during Delta predominance in the EU/EEA countries (29), in the current project, **no differences were observed in adolescents**-matched cohorts **against non-severe COVID-19**, between those who used the same mRNA vaccine or combination of mRNA vaccines for the primary schedule. In agreement with the very low crude risk of hospitalization, ICU admission and death for 12-17 year-olds in the EU/EEA countries (29), we did not observe severe cases among adolescents that precluded the VE assessment in the current project.

Among both, adults and adolescents, no clear diverse patterns of waning of immunity for homologous versus heterologous schemes, or changes in the comparative effectiveness among age groups or variant predominance periods were observed. This is also due to the small sample size contributed to this primary analysis (29).

For **primary objective 2 (boosting)**, a **high VE of PF and MD booster** doses **against severe COVID-19 and COVID-19 with death** were observed **in adults** regardless of whether booster doses were the same or different from the brand of the two doses administrated as primary vaccination. Those results were confirmed with more precision in Spanish cohorts (where most cases were captured), than in Italian cohorts (being able to provide only overall estimates for severe COVID-19 and not for death), and in the UK (being able to provide only estimates for death and complete homologous three doses). The use of AZ as boosters was not sufficient to estimate its effectiveness against those severe outcomes. However, mRNA-boosted individuals showed high effectiveness against severe COVID-19 also among those who used AZ as primary vaccination.

The majority of patients who died with COVID-19 belonged to the oldest groups who were hardly vaccinated with AZ. Consequently, the sample size was not sufficient to find enough fatal cases for quantification of the benefit of a booster among those initiating the immunization with AZ (recommended initially for younger populations). **The mRNA booster doses were observed effective during the Omicron** predominance period against both severe outcomes (according to one data base ES-BIFAP) **as well as during the previous Delta** period (according to three sources from Spain and the UK). As expected, people with a third homologous dose were more represented in the study cohorts than those with heterologous ones.

For **exploratory objective 5 (waning of immunity after booster)**, **VE against severe COVID-19** was observed from the 1st week after vaccination and **remained until the 2nd and 5th months since homologous booster** (in ES-SIDIAP and ES-BIFAP, respectively) **and until the 1st or 3rd month, respectively for heterologous booster in comparison with their unboosted pairs**. **The reduction of the effectiveness could be triggered by both waning immunity and lower effectiveness against the more aggressive Omicron variant**, which became dominant in Europe at that time (January 2022, approximately, as most adults received their booster dose around November-December 2021). In particular, the heterologous booster started to be administrated approximately one month later than the homologous one (September 2021 in Spain, and 71-86% of them in December 2021), with the starting of Omicron predominance that could affect the efficacy and its duration. Although boosters were also effective during Omicron, the heterologous showed lower duration of the VE than the homologous. This could also be influenced by the one-month shorter follow-up time for heterologous schedules. Insufficient follow-up might have prevented the identification of more precise estimates of vaccination effectiveness in ES-SIDIAP. Of note, individuals vaccinated with heterologous booster doses could have different characteristics and risk of infection/prognosis than those receiving homologous booster, so the direct comparison may also be confounded by those. Effectiveness is here provided in comparison with non-boosted pairs carefully selected as described in the methods section for pairs comparability. The **effectiveness against death with COVID-19 remained until enough number of cases were detected**, **with a maximum duration of 5-6 months after administration of the booster** in a Spanish data source. The short follow-up time limited that evaluation in the other two data sources.

Finally, **individuals who received a 3rd dose of the studied three vaccine brands, regardless of the primary vaccination scheme, were more protected against non-severe COVID-19**

than their non-boosted pairs. However, **important variations** in the effectiveness estimation (28-74%) were observed **by brand and/or country**. Effectiveness was high for AZ boosters in the UK, medium for PF or MD in Italy, the UK and The Netherlands, and very low for all the brands in Spain (that did not improve after restricting to individuals with negative tests before vaccination in sensitivity analyses). These results on the effectiveness of booster doses concur with previous studies (3,29–31), specifically with those of the test-negative case-control study by Andrews et al., who evaluated the effectiveness of the booster-dose vaccine against symptomatic illness caused by the Omicron and Delta variants in England. Overall, they found that the administration of a booster dose of PF or MD, after either a primary homologous course of AZ or PF vaccines, substantially increased protection. The protection was seen from the 2nd to the 4th week after vaccination and waned over thereafter (1,8,31). Again, the above-presented booster effectiveness data correspond to the subgroup of adults with no prior documented SARS-CoV-2 infection. Also, the **benefits of administering a heterologous booster** dose (any brand) in **preventing severe COVID-19 outcomes among individuals with prior infection was medium-high** in Spain (mainly against COVID-19 with death). It was not replicated in the UK or Netherlands, with a reduced number of cases (<5), and no data were available for those events in Italy. The VE of the homologous booster was low against those severe outcomes and also against non-severe SARS-CoV-2 infection (in one data source in Spain and in the Netherlands) or even an increased risk of SARS-CoV-2 infection was identified in one database (ES-BIFAP). This last estimation could be confounded by the period with the start of the Omicron variant and other differences (such as more use of self-tests). The vaccination scheme for individuals suffering from a prior SARS-CoV-2 infection, both the completion of the primary regimen and the administration of a booster dose, was slightly different at the beginning of the pandemic situation (i.e., doses delayed in time, compared with those free of prior infection). Thus, a lower number of adults was available for this subgroup for being compared with a sufficient follow-up time to produce robust VE estimates. Finally, unfortunately, the planned matching by the moment of the prior infection (within the same month) could not be implemented in the current study. Thus, selection bias cannot be discarded (32) in the estimations herein presented and interpretation of effectiveness in these subgroup with prior SARS-COV2 infection should be cautious.

Previous studies among immunocompromised patients (33–35) showed significant improvement in the immunogenicity of the vaccines when administrating a third dose. We could estimate the VE of a third dose among **vulnerable patients who received homologous primary schemes** (heterologous were scarcely used). Among adults who are immunocompromised, we found that **the VE of a 3rd dose (homologous or heterologous) was medium-high in up to five data sources against severe COVID-19 events except for non-severe COVID-19 (where VE was low-medium in Spain or the Netherland)**. In comparison with no booster, our estimations suggest a higher VE with heterologous than homologous boosters, while similar protection against COVID-19 with death was observed. It is important to consider that this direct comparison of hetero and homo booster may be confounded by the vaccination moment (omicron or delta) as well as other baseline characteristics of the patients that triggered the receipt of either the hetero- or homologous 3rd dose.

Among **patients previously diagnosed with cancer**, **medium-high effectiveness of a 3rd homologous or heterologous dose against severe COVID-19 or death with COVID-19 was observed (only Spanish data) or against non-severe infection (in three data sources)**. Also, **high VE** was observed against severe outcomes (only homologous could be tested) in adult **patients with severe renal disease** in one (ES-BIFAP) out of three data sources tested.

Since we expected a reduced sample size of patients at high risk of COVID-19 severe infection or related comorbidity, several **meta-analyses** were planned and performed in certain clinical sub-groups to test effectiveness against the less frequent (severe) outcomes. We estimated pooled VE of adults with **homologous primary vaccination followed by a homologous or heterologous booster dose against severe COVID-19 and death with COVID-19**. According to the different clinical subgroups assessed, those **patients with immunodeficiency or immunosuppressant drugs and cancer would seem to benefit substantially from a booster in terms of both outcomes (VE ranged from 54% to 75%, low or no heterogeneity).** This provides important real-world insights about the benefits for these patients less represented in the studies.

In this study, **adolescents with boosters after homologous primary vaccination** were also identified and analyzed, although no severe COVID-19 cases occurred for such effectiveness

estimation. Among adolescents, the level of antibodies observed after vaccination was like that of adults in some previous studies and other reported effectiveness with real world data are in agreement. For instance, Fowlkes et al. estimated an adjusted VE at 14-149 days after the second dose of 87% against symptomatic and asymptomatic Delta infection that decreased to 59% against Omicron infection in individuals aged 12-15 years (36). Then, an additional dose provided a 3.7 (95% CI: 2.7-5.2%) fold increase in immediate (up to 60 days) protection against documented infection compared to the primary vaccination course only and, in a systematic review, the VE was estimated at 73% (95% CI: 67-79%) when compared to individuals who completed the primary course by January 2022 (32,37). In agreement with that, in the current study, we assessed the **VE of booster doses among adolescents** and found that the level of protection against SARS-CoV-2 infection was **medium for three doses of Pfizer in Italy (VE: 67%) both during Delta and Omicron but was low in Spain**. **Three homologous Moderna doses** could only be analysed in Spain, showing also **medium effectiveness (64%) in Spain** in comparison with similar two doses without booster. Effectiveness lasted 1-2 months. **In Spain**, the observed medium effectiveness was **only significant during the Omicron** period. No statistically significant effectiveness was observed overall after controlling by prior testing in BIFAP-SP. The Italian VE and others in Spain agree with those aforementioned, as well as a reduction of the effectiveness among the youngest was also reported (iscii report). Nevertheless, some of the low VE in Spain could be the consequence of several reasons.

A decrease in infection notification rates was observed for adolescents in the EU/EA countries in the early weeks of 2022 (29), possibly due to changes in testing approaches (more frequent due to the high incidence of the infection and/or use of self-test), the Omicron peak, and/or reporting delay. Those could have also affected the booster VE estimation in the current study, mainly administrated during the last month of 2021-beginning of 2022. Also, adolescents were not called to boost till February 2022, thus early boosted patients could be at high risk of infection or related comorbidities, underestimating the VE. Furthermore, more frequent patients with prior infection along the time and/or more underlined prior infections in the control groups (as observed in previous studies) (38) could be playing a role in the reduced estimates. A decrease in the effectiveness was also observed against severe events from winter 2021 (being 43.9% in February 2022 [95% CI: -20.2 to 71.3] among vaccinated adolescents with or without booster doses in a previous study in Spain (isciii report). Pooled estimates were not calculated for adolescents since no severe outcomes were identified.

In the literature, the odds of COVID-19 hospitalization and ICU admission among adolescents with underlying comorbidities were 9 and 25 times greater, respectively, than those without co-morbidities (39). Unfortunately, we could not assess the prevention of those severe COVID-19 outcomes due to the lack of cases. Regarding the protection **against non-severe COVID-19 infection**, the **VE of a 3rd homologous dose was medium among Italian immunocompromised adolescents**. Effectiveness was not proved in the other two databases (in Spain) analyzed, in which incidences could be estimated but the sample size could not be enough to show effectiveness against that mild infection. Similarly, the reduced sample size in young patients in vulnerable subgroups or with prior COVID-19 impeded the VE estimation.

Considering Primary objectives 3 (children) and 4 (waning of immunity), we showed that **two doses of an mRNA vaccine protected children against non-severe SARS-CoV-2 infection during the Delta variant predominance period**. However, this protection varied widely among data sources in this study (29% to 77%). A higher efficacy of the PF vaccine was calculated in a RCT though (90.7% effective at preventing symptomatic COVID-19, although the true rate was between 67.7% and 98.3%) (40). **During the Omicron period, VE became moderately lower** in one data source, whereas an even higher risk of infection among vaccinated children was found in another one (as consequence of differential surveillance between vaccinated and unvaccinated as observed in sensitivity analysis). These results are not far from a previous observational study that calculated a VE of 31% (95% CI = 9%-48%) against any Omicron infection 14-82 days after receiving a second dose of the PF vaccine (41). **From the 1st week after full immunization, children were protected against non-severe SARS-CoV-2 infection**. This protection was maintained but **decreased** in magnitude **over time until the 5th month** consistently across all data sources. This decrease in effectiveness could be the result of both, the expected waning of immunity and the effect of the dominance of the new Omicron variant around the 4th or 5th month after the most represented age group in this analysis (those 12-14 years old) have received the vaccination. There was an increase in the incidence in all age groups of COVID-19 from December 2021, which

was a consequence of the onset of the Omicron variant era and the relaxation in Public Health measures by national and local governments as it was considered a less severe variant. Moreover, IRs could be underestimated due to the increasingly high use of self-COVID-19-tests which results may not have been reported/surveilled and other factors aforementioned for adolescents. It should be noted that in the estimations of the two Spanish data sources and IT-PEDIANET, compared with unvaccinated children, we took into account a possible healthy vaccinee effect by controlling for the number of contacts with the health system in the week prior to time 0, but this was not the case in the other participating data sources due to the lack of availability of this variable. Those findings supports that the protection since the 1st week from vaccination found in ES-BIFAP and -SIDIAP was not likely to be affected by a healthy vaccinee effect. However, other factors aforementioned for adolescents do also apply here, including the reduced notification rate of the infections In the last studied months or a potentially reduced VE with the global higher exposure to the virus, among others.

The obtained VE estimates for **younger children (5-11 years) had low precision and were very heterogeneous across databases** due to the small number of individuals and COVID-19 cases, approximately 10 times lower than those for 12–14 years old. In this age group, vaccination started in December 2021. They were recommended to receive the second dose 8 weeks after the first one, so they were underrepresented in our study (most data sources were current as of December 2021), which led to the calculated VE estimates reflecting mostly the effect of vaccination in older children (preadolescents, i.e., 12-14 year old) studied together in the same group 5-14 years old.

The results of the **sensitivity analysis**, in which we restricted the population to **children who had been tested for COVID-19 at least once before the time 0**, **suggested important effect modification in the VE estimates in the ES-BIFAP data source (from -11% overall to 19%).** This could be the result of a higher rate of COVID-19 testing among vaccinated children, probably because they are thought to be more exposed or vulnerable to infection, by visiting health services more frequently or by parental preferences.

A complete primary vaccination regimen offered a higher level of protection for severe COVID-19 disease in children (12-14 years old, as no cases occurred in those 5-11), as observed in the two Spanish databases, the only ones with a sufficient number of cases to estimate this outcome. However, during the predominance of the **Delta** variant, IRs were nearly 3 times higher in ES-SIDIAP (0.27/100,000pd) than ES-BIFAP (0.08/100,000pd) for control children, while, for the vaccinated group, IRs were quite similar (0.02 and 0.03/100,000pd, respectively). This resulted in a quite different VE between these two databases because the **high effectiveness seen in ES-SIDIAP-ES (94%) was not observed in -BIFAP (non-statistically significant 61%). In the Omicron period, though, with data only from ES-BIFAP, vaccination has shown to provide moderate protection against severe disease**, while, for non-severe disease, the protection was very limited. Similar VE against hospitalized COVID-19 (49%; 95% CI: 26-64%) among people aged 12-17 years was reported in February 2022 in Spain after suffering a decreasing trend from November 2021 (previously, up to October 2022, VE was kept ≥90%) (isciii report). In a recent study, it was similarly observed that the effectiveness of two doses of PF vaccine against COVID-19–related hospitalization was higher than estimates of the effectiveness against infection, but uncertainties were greater due to a few of events (42). Furthermore, we believe that VE was underestimated in ES-BIFAP due to 1) confounding by surveillance (i.e. testing) as observed in sensitivity analysis; 2) differential COVID-19 severity misclassification, since in a manual review to confirm/validate the reason for hospitalization of patients aged 12-18 years old who were identified as cases of severe COVID-19, non-vaccinated were more often confirmed (63%) than vaccinated individuals (24%).

It appears that **children with any immunodeficiency condition have high protection (85%, 95%CI= 27%-97%) against severe COVID-19 disease after being fully vaccinated.** This was only observed in ES-BIFAP (although high, 74% not statistically significant estimate was also produced in ES-SIDIAP). Nonetheless, this result should be interpreted with caution because very few cases were found in each comparator group. What is important is that we were able to provide some evidence on VE in more vulnerable groups (immunocompromised children and cancer patients - in the latter case only for the non-severe infection outcome), those who could benefit most from vaccination. Further studies in this subgroup of children with special conditions are warranted.

Overall, in children, there were no deaths related to COVID-19, as in previous studies (41)(isciii report).

Data coming from the performed **meta-analyses suggested a high VE of full primary vaccination against severe COVID-19 in children although it did not reach statistical significance**. The lack of statistical significance was a consequence of the random effects performed and was coherent with the found heterogeneity found, since it allows differences in the treatment effect. Post-hoc, we replicated the meta-analysis using fixed effects (i.e., assuming all individual studies were estimating the same treatment effect) and so weighting more the data sources with lower standard error (ES-BIFAP: 93% vs 7%). Heterogeneity was observed among children (in the two Spanish data sources participating) and among adults' subgroups (with cancer, immunodeficiency, or those receiving transplants (that are less precise due to low sample sizes)). Heterogeneity could be the consequence of the differential effect of (un-)measured confounders among regions, different COVID-19/covariates recording habits or algorithms among data sources producing more/less predictive values or sensitivity for adjustment (inequitable control of confusion); differential contribution of virus variants (i.e., only Delta or Delta and Omicron periods included); different vaccination schedules; parents' vaccination preferences between the regions; different brand predominance in individuals contributing in each group per data source among other.

The main objective of the current study was to estimate and report the VE of individual data sources and we recommend interpreting as such. Also, to avoid wrong interpretations, we recommend interpreting with caution the pooled VE of meta-analysis with heterogeneity >60%.

Due to the lack of enough sample size and cases of severe COVID-19 and death, we were unable to provide VE estimates in children for all outcomes and objectives (comparing VE between primary regimens) like for adults and adolescents.

As for **all-cause mortality (secondary objective)**, data from four databases (Spain, the UK, and the Netherlands) showed lower risk after a booster dose with homologous (for aged ≥60 years), and one database (the UK, only for aged 80 years old) after heterologous primary vaccinations, when compared with unboosted pairs. **These results complement the effectiveness of booster doses found against death with COVID-19 in the primary objective.** This secondary analysis could be including potential fatal COVID-19 cases for which information on the infection status may be missing in the electronic health records before death (11). That is recognized to be the case for data instances during the periods of disrupted or overwhelmed health systems (as occurred during the pandemic period), not recorded self-diagnosis tests, or among patients not being attended in healthcare centers. However, we do not know the proportion of false cases (like death for other reasons) that we could be included in this study, limiting more precise interpretations.

The findings in ES-BIFAP about **vaccinated children or boosted adults with prior SARS-CoV-2** infection with **an increased risk of non-severe outcomes could be confounded by the Omicron variant period.** Vice versa, other differences among considered study periods (such as higher use of self-tests, triggering outcome misclassification) should be taken into account when assessing the effectiveness during both predominant periods. That could have also influenced the VE estimations in other data sources, for which prior tests were not available for sensitivity analysis. The vaccination scheme for individuals suffering from infection prior vaccination, both the completion of the primary regimen and the administration of a booster dose, was slightly different at the beginning of the pandemic situation (i.e., doses delayed in time, compared with those free of prior infection). Thus, a lower number of adults was available for this subgroup for being compared with a sufficient follow-up time to produce robust VE estimates. Of note, unfortunately, the planned matching by the moment of the prior COVID-19 infection (within the same month) could not be finally implemented in the current study. Thus, selection bias (32) cannot be discarded in the estimations presented and interpretation of effectiveness in this subgroup should be cautious.

The VE reported for each data source refer to the scenarios (people characteristics and calendar moments among other) described in the Tables of ANNEX I. As reported, a broad range could be observed sometimes for the VE among the data sources that can be the results of many factors including those baseline characteristics of the matched vaccinated populations and their pairs (that could vary among sources as observed in descriptive Tables of ANNEX I) for instance immunocompromised children represented between the 6% (NL-PHARMO) and the 21% (ES-SIDIAP)

of the total participants, the covered regions (with different virus prevalence, predominance and potential public health recommendation, people habits and believes, etc) or calendar moments.

11.2. Limitations

Electronic health records still face difficulties in research on almost real-time data. Further investigation of the multiple reasons for those difficulties is needed, together with further research of solutions. We encountered difficulties in extracting the most updated data in most data sources (ES-SIDIAP, UK-CPRD, NL-PHARMO, IT-PEDIANET) [precluding the study of VE among children aged 5-11y in all countries (mostly vaccinated from December 2021 ahead), long-term follow-up among boosted people (mostly administrated from Autumn 2021 ahead) during the period of the Omicron predominance (from January 2022)] and in having death information in one data source (IT-INSPIRE). Small cohorts did not allow the analysis of rare COVID-19 outcomes with precision. However, an important number of patients at high risk of severe disease were represented, especially patients diagnosed with immunosuppression, although not enough sample size was reached to answer all the research questions in all countries.

In the current study, we sacrificed analytical power to maximize comparability by matching multiple factors. As consequence, the sample size was insufficient to estimate the differential effectiveness between homologous and heterologous primary vaccinations mainly among adolescents and adults' subgroups. However, relaxing the other matching criteria (i.e., days-by-day of date of 1st dose, time, or booster-time0) would have lost the designed comparability that was also observed in the balance achieved in the numbers of subjects with comorbidity and comedications among all groups.

Similarly, booster doses were mainly homologous based on the brand used in the primary vaccination schemes. This affected the analysis of the heterologous booster schemes, which lacked of the power for COVID-19 cases in most of the data sources, especially the most severe and rare. We identified that the main vaccine brands used as booster doses were PF and MD, which later have become the two Omicron-adapted vaccine platforms currently administered. The booster doses were administrated during the final months of the study period for most of the data sources. However, we were able to estimate medium-term vaccine booster effectiveness in some of them.

According to our estimations, the exclusion of data sources without the link to hospital data (UK-CPRD and NL-PHARMO) or death information (IT-INSPIRE) did not affect the probability to demonstrate a protective effect of the booster against most severe outcomes, even though we encountered new difficulties to finally estimate it in all remaining data sources.

Studies of COVID-19 vaccines may be subject to limitations common to non-randomized studies based on existing healthcare data.

The patient's test-seeking behavior may be associated with both severities of infection symptoms and personal health-seeking behavior, which may introduce selection bias or confounding. To avoid that effect, adjustments by the number of previous visits among compared groups (healthy vaccinee), as well as a sensitivity analysis restricting to patients with negative test results (controlling potential differential probabilities to infection, vulnerability to severe episodes, or access to COVID-19 testing) were planned. However, information on those factors was not available or could not be produced on time in three data sources (IT-INSPIRE, UK-CPRD, NL-PHARMO). That effect could overestimate the VE in those data sources since boosted patients and vaccinated children were probably 'healthy' (i.e., without symptomatology) at the vaccination moment, while we cannot affirm the sae for their unvaccinated controls.

Confounding of the relationship between booster receipt and COVID-19 outcomes may be likely. The use of eligibility criteria to define a comparable exposure group (i.e., matched on the date of the booster, date of each vaccination, type of primary vaccination, region, sex, and age) and further adjustment by factors (around 21 to 42 covariates in all data sources), taken into account in the statistical models, allowed investigators to adequately minimize confounding. Still, some known unmeasured confounding (not finally controlled for in all data sources (e.g., region or no number of healthcare consultations to control by health vaccinee effect or number of prior tests in some data sources) could play a role as aforementioned along the Discussion section.

Even though efforts were made to properly control bias and multivariate adjustment, minimizing confusion, information about the job and type of residence was not available in the data sources (i.e., prioritised groups to be boosted or certain scheme receipts, such as people living in nursing homes or health care workers). Consequently, confusion may still be present due to the higher probability of infection among them (and more frequently/earlier vaccinated) versus other social groups. That aspect would direct towards a reduction in the effectiveness estimations. However, in the current study, we believed that matching by date of the 1st and 2nd dose prevented the differential effect of prioritized groups in all databases as it determined the prioritized group (preference for full vaccination and boosters). In other words, people with a homologous scheme were compared to people with a heterologous vaccination if vaccinated with 1st and 2nd doses at the same moment (also in the boosted versus unboosted comparison, that strategy was used to control by prioritization).

According to preliminary results, most data sources did not have data about determinants, such as personal (or parents') health-seeking behavior, recorded, so the selection of patients included in the compared groups and/or their dates to start the comparison could be biased, i.e.:

- Patients prone to infections or to develop a severe/hospitalized infection or those with active respiratory infection/disease (prevalent active pneumonia, COPD, etc.) decided to attend to receive the vaccination earlier/more than those less prone to it, we could observe patients more affected by covid among the vaccinated group. That aspect would direct towards a biased lower effectiveness estimation due to selection bias).
- The opposite could also be true, i.e., patients more adhered to general prevention measures (personal health-seeking behavior such as applying social distancing, wearing masks, volunteer confinement, etc.) were more vaccinated, thus, we could be observing more infections in the unvaccinated group. That aspect would direct towards a biased higher effectiveness estimation due to selection bias.

Selective recruitment into the study of subjects for compared groups recorded in the database with quality criteria (up-to-standard information) that are not representative of the general respective groups' subjects respectively in the source population could produce selection bias. For instance:

- If we were losing people having died from COVID-19 even though they were vaccinated (i.e., non-effective vaccinations), selection bias would be present. Similarly, if more vaccinated participants were survivors of previous infections than vaccinated non-participants (i.e., they died by COVID-19) because those who attend the primary care physician (PCP) (having information and minimum anamnesis in the database required to participate), while vaccinated non-participants do not.
- Also, if vaccinated participants were recorded in the database because they are more closely surveyed by the PCP/nurse due to their predisposition to complicated COVID-19 infection (i.e., baseline health conditions), while vaccinated non-participants were not recorded in the database because they do not seek healthcare, we would be including in the study patients with more probability to severe infection than the real overall vaccinated individuals.

The selection bias could direct towards any direction the effectiveness estimates.

If an individual's personal health beliefs and behaviors increase the likelihood of both booster vaccines receipt and seeking health care for the milder disease, then the analysis of this outcome may be particularly subject to confounding by healthcare–seeking behavior.

Misclassification of vaccine exposure, outcome status, or covariates is possible in existing healthcare data not collected for research purposes. Information about COVID-19 infection (through test results or codes) may not be captured reliably in databases. Additionally, COVID-19 testing was not systematic throughout the study period, and laboratory confirmation of case status may not be available always. Bias may be minimised since COVID-19 testing becomes near universal and repeated with individuals, with similar testing capacity and practice across regions, though. However, the marketing of patients' auto-diagnosis became popular over the months which could imply missing positive cases recorded in tests performed in clinical settings. That could be influencing the evaluation during the Omicron predominance and last of Delta periods.

Since we did not have any better gold standard for case status than the information about confirmed infection through lab results (used for case definition in most databases in the current study), its

precision could not be evaluated but with the expected external incidences published by the health authorities. For that, before effectiveness estimation, a comparison (benchmarking) between the incidences of COVID-19 events in the participating databases and those reported by the corresponding country was performed. Benchmarking allowed us to confirm the events that were the expected in the data sources. Also, we consider that, if COVID-19 misclassification occurred, it would be similar in both heterologous and homologous primary vaccinations, booster and no-booster, and children with and without full vaccinations. However, surveillance or detection bias (different likelihood of screening or testing for COVID-19 between the groups) cannot be discarded. If boosted individuals had less likelihood of screening or testing than non-boosted, we would have artificially observed more cases among non-boosted, directing toward a biased higher booster effectiveness estimation. This may happen also for vaccination effectiveness among children but be less frequent when comparing heterologous and homologous primary vaccinations.

Data from multiple regions and sources were included in the current study. Thus, there may be variations in the capture and recording of various clinical elements. Additionally, distinct types of data sources were used (e.g., records from general practice, hospital, and lab results from other registers) as well as different coding systems. Thus, the variables defined in different data sources may not exactly represent the same exact concept across data sources. The heterogeneity of effectiveness across data sources may be due to the underlying heterogeneity of confounding control, misclassification, or other data source factors rather than true differences in the effectiveness.

Patterns of routine healthcare delivery and utilization may be disrupted during the pandemic as patients and providers forgo or delay routine preventive, elective, or non-emergency care. These disruptions in health care may result in under-ascertainment of important patient comorbidities in existing healthcare databases during periods of disruption. However, we controlled this effect (in all databases by design) by matching by the date, preventing the differential effect of disrupted periods between compared groups (i.e., the moment of vaccination with homologous versus heterologous second vaccination, with the booster and second dose vaccination in children).

With the increase in coverage, there is the potential for rapidly changing herd immunity in the population. This study is not designed to assess the overall and indirect effects of vaccination with COVID-19 vaccines.

Comparative analyses may not be possible in every setting and brand combination if, for instance, confounding is deemed to be insurmountable. However, descriptive information about vaccine recipients and crude incidence rates may still be informative and meaningful, even without calculations of vaccine effectiveness measures. Since matching was maximized in the current study, and according to current preliminary results, compared cohorts are also comparable in terms of potential confounders (i.e., comorbidity and comedication) captured.

The capture of over-the-counter medications, potentially indicative of short-term disease status (e.g., painkillers, cough medicines, and fever reducers) may not be captured reliably.

The use of secondary data could lead to a misclassification of both the exposure and the outcomes in terms of the result and the date. Nonetheless, both outcomes and exposures have been validated in previous studies. Asymptomatic COVID-19 infection could be underestimated as it is difficult that people without symptoms to do the test. Underestimation of positive at-home tests could be higher over time, especially during the last months of 2021 and during the Omicron period. At-home COVID-19 tests may not follow surveillance or not be compulsory declared in all countries. Some data sources do not contain the cause of death, but a proxy has been established (as death occurring in COVID-19 people in a plausible temporary window).

This study aimed to compare different brand options, available in participant countries, for the 2nd dose and subsequent booster. Thus, one-dose Janssen COVID-19 vaccine or one dose among patients with prior COVID-19 infection (i.e., homogeneous by defaults) were not considered in the current study. The comparability between the time after the one-dose scheme and two-doses schemes may be compromised in observational studies by confusion immediately after the start of the vaccination. Thus, people vaccinated with Janssen's vaccine (around 7 million in the covered countries; Table 1) were not represented in the current study.

11.3. Interpretation

Interpretation of results about the study objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence has been addressed in the Discussion section.

11.4. Generalisability

The methodology used in this study aimed to maximize comparability between matched pairs, albeit reducing or sacrificing generalisability (i.e., external validity), similarly to clinical trials, having left out of the comparisons many other individuals for whom no match was found. Thus, in our study, a higher proportion of individuals with a heterologous primary COVID-19 vaccine regimen were selected compared to those in the general population.

The above-presented results are mostly generalizable to a particular period when COVID-19 variants were different from the variants and subvariants that are prevalent these days or will be soon, so generalizability in terms of circulating variants may have certain limitations. Only two databases include follow-up up to February 2022 and thus captured sometime of Omicron predominance as it entered circulation on the continent at the end of 2021. Another reason limiting the follow-up time in this study is the fact that many individuals censored their follow up very rapidly as their corresponding partners are vaccinated soon after being selected as a control.

However, our study does represent individuals from the participating data sources. People with a booster or third dose are well represented in terms of vaccination coverage, age and sex. On the other hand, specific populations such as immunocompromised patients are very well represented as they are comprehensively captured in the participating databases and thus among our study individuals.

12. Other Information

No additional information is available.

13. Conclusions

Focusing on the Delta and the early Omicron predominance era, herein, we assessed vaccine effectiveness of the EU-approved PF, MD, and AZ vaccines against SARS-COV-2 infections including hospitalization with COVID-19 in Southern Europe (Spain and Italy), death with COVID-19 in Spain and the UK, and non-severe COVID-19 in southern countries (Spain and Italy), the UK and the Netherlands.

From our results, the protection against severe COVID-19 outcomes was not different between adults with homologous or heterologous primary vaccinations, in line with the proven efficacy of homologous schemes in randomized clinical trials (in comparison to non-vaccinated or other vaccines). That endorses the decision of some countries' health institutions to recommend switching brands for the second dose (especially for those first vaccinated with AZ). Likewise, against non-severe infection, adults and adolescents benefitted similarly when receiving a different brand for the 2nd dose, and adults were even more protected if vaccinated with AZ first.

mRNA 3rd doses offered an important level of protection against severe COVID-19 outcomes, high for heterologous and moderate for homologous schemes, regardless of the brand or the variant predominance periods, i.e., during Delta (Spain and UK) or Omicron (Spain). This was also true for more vulnerable subjects (i.e., the elderly and those at high risk of COVID-19 complications). Our results among patients at a higher risk of complications add real-world evidence to the immunogenicity of the vaccines already observed in clinical studies. Thus, a third dose seems to be a good strategy to prevent the most severe complications of SARS-COV2 infection in adults. In line with previous studies, we observed a wane in effectiveness over time in the early months that warrants further research to assess the proper time frame on when a booster should be administered.

Our results also suggest that a 3rd dose of Covid-19 vaccines, both mRNA platforms and AZ, conferred a certain level of protection at reducing SARS-CoV-2 infection in both variant predominance periods for adults and adolescents, overall, and for vulnerable patients, albeit it varies widely across vaccine platforms, countries and analyses subgroups.

Caution must be taken when interpreting these results for public health decision-makers, as they might be affected, among other reasons, by a well-known reduction in the notification of positive cases occurring since the last months of 2021, in line with reports from ECDC. This last limitation impacted substantially effectiveness estimates calculated for adolescents.

Our data infer that full primary vaccination protected 12-14-year-old children against hospitalized Covid-19 disease during the Delta and Omicron variant periods in Spain, including immunocompromised children. We cannot conclude the same for the youngest age groups of 5-11 years old due to the absence of cases. This highlights the need for bigger and longer studies to state if primary and booster doses are effective in disease prevention in children. We also observed protection against non-severe Delta and Omicron infections. These findings are consistent with the clinical trial–reported efficacy of PF and MD vaccines against SARS-COV-2 infection and observational studies.

The conclusions herein cannot fully address the current situation where the updated PF and MD vaccines with Omicron sub-variants are currently being used in the EU countries in a different SARS-COV2 virus epidemiological scenario with some natural immunity already conferred by prior SARS-COV2 infections in the population. Therefore, updated broader effectiveness estimations in all age groups on reinfections and disease severity are required. Still, this study provides evidence to inform the benefit-risk assessment of vaccines.

14. References

1. Andrews N, Tessier E, Stowe J, Gower C, Kirsebom F, Simmons R, et al. Duration of Protection against Mild and Severe Disease by Covid-19 Vaccines. New England Journal of Medicine [Internet]. 2022 Jan 27 [cited 2022 May 30];386(4):340–50. Available from: https://www.nejm.org/doi/10.1056/NEJMoa2115481
2. Pouwels KB, Pritchard E, Matthews PC, Stoesser N, Eyre DW, Vihta KD, et al. Impact of Delta on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK. medRxiv [Internet]. 2021 Aug 24 [cited 2022 May 30];2021.08.18.21262237. Available from: https://www.medrxiv.org/content/10.1101/2021.08.18.21262237v1
3. Bernal JL, Andrews N, Gower C, Robertson C, Stowe J, Tessier E, 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 [Internet]. 2021 May 13 [cited 2022 May 30];373. Available from: https://www.bmj.com/content/373/bmj.n1088
4. Hyams C, Marlow R, Maseko Z, King J, Ward L, Fox K, et al. Effectiveness of BNT162b2 and ChAdOx1 nCoV-19 COVID-19 vaccination at preventing hospitalisations in people aged at least 80 years: a test-negative, case-control study. Lancet Infect Dis [Internet]. 2021 Nov 1 [cited 2022 May 30];21(11):1539–48. Available from: http://www.thelancet.com/article/S1473309921003303/fulltext
5. Pritchard E, Matthews PC, Stoesser N, Eyre DW, Gethings O, Vihta KD, et al. Impact of vaccination on SARS-CoV-2 cases in the community: a population-based study using the UK's COVID-19 Infection Survey. medRxiv [Internet]. 2021 Apr 23 [cited 2022 May 30];2021.04.22.21255913. Available from: https://www.medrxiv.org/content/10.1101/2021.04.22.21255913v1
6. 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. The Lancet [Internet]. 2021 May 1 [cited 2022 May 30];397(10285):1646–57. Available from: http://www.thelancet.com/article/S0140673621006772/fulltext
7. Ismail SA, Garcia Vilaplana T, Elgohari S, Stowe J, Tessier E, Andrews N, et al. Effectiveness of BNT162b2 mRNA and ChAdOx1 adenovirus vector COVID-19 vaccines on risk of hospitalisation among older adults in England: an observational study using surveillance data. 2021;
8. Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. New England Journal of Medicine [Internet]. 2021 Aug 12 [cited 2022 May 30];385(7):585–94. Available from: https://www.nejm.org/doi/10.1056/NEJMoa2108891
9. Andrews N, Stowe J, Kirsebom F, Toffa S, Sachdeva R, Gower C, et al. Effectiveness of BNT162b2 COVID-19 booster vaccine against covid-19 related symptoms and hospitalization in England. Nat Med. 2022 Jan 14;
10. Questions and answers on COVID-19: Vaccines [Internet]. [cited 2022 May 30]. Available from: https://www.ecdc.europa.eu/en/covid-19/questions-answers/questions-and-answers-vaccines
11. Duarte-Salles T, Prieto-Alhambra D. Heterologous vaccine regimens against COVID-19. The Lancet [Internet]. 2021 Jul 10 [cited 2023 Apr 21];398(10295):94–5. Available from: http://www.thelancet.com/article/S0140673621014422/fulltext
12. Borobia AM, Carcas AJ, Pérez-Olmeda M, Castaño L, Bertran MJ, García-Pérez J, et al. Immunogenicity and reactogenicity of BNT162b2 booster in ChAdOx1-S-primed participants (CombiVacS): a multicentre, open-label, randomised, controlled, phase 2 trial. The Lancet [Internet]. 2021 Jul 10 [cited 2023 Jan 20];398(10295):121–30. Available from: http://www.thelancet.com/article/S0140673621014203/fulltext
13. EMA and ECDC recommendations on heterologous vaccination courses against COVID-19: 'mix-and-match' approach can be used for both initial courses and boosters | European Medicines Agency [Internet]. [cited 2022 May 30]. Available from: https://www.ema.europa.eu/en/news/ema-ecdc-recommendations-heterologous-vaccination-courses-against-covid-19-mix-match-approach-can-be
14. Monge S, Rojas-Benedicto A, Olmedo C, Mazagatos C, José Sierra M, Limia A, et al. Effectiveness of mRNA vaccine boosters against infection with the SARS-CoV-2 omicron (B.1.1.529) variant in Spain: a nationwide cohort study. Lancet Infect Dis [Internet]. 2022 Sep 1 [cited 2023 Jan 20];22(9):1313. Available from: /pmc/articles/PMC9162477/
15. Pottegård A, Kurz X, Moore N, Christiansen CF, Klungel O. Considerations for pharmacoepidemiological analyses in the SARS-CoV-2 pandemic. Pharmacoepidemiol Drug Saf [Internet]. 2020 Aug 1 [cited 2022 May 30];29(8):825–31. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/pds.5029
16. Data | CPRD [Internet]. [cited 2022 May 30]. Available from: https://cprd.com/data
17. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, Staa T van, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol [Internet]. 2015 Jun 1 [cited 2022 May 30];44(3):827–36. Available from: https://academic.oup.com/ije/article/44/3/827/632531
18. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. Int J Epidemiol [Internet]. 2019 Dec 1 [cited 2022 May 30];48(6):1740–1740g. Available from: https://academic.oup.com/ije/article/48/6/1740/5374844
19. Thurin NH, Bosco-Levy P, Blin P, Rouyer M, Jové J, Lamarque S, et al. Intra-database validation of case-identifying algorithms using reconstituted electronic health records from healthcare claims data. BMC Med Res Methodol [Internet]. 2021 Dec 1 [cited 2022 May 30];21(1):1–8. Available from: https://link.springer.com/articles/10.1186/s12874-021-01285-y
20. Thurin NH, Pajouheshnia R, Roberto G, Dodd C, Hyeraci G, Bartolini C, et al. From Inception to ConcePTION: Genesis of a Network to Support Better Monitoring and Communication of Medication Safety During Pregnancy and Breastfeeding. Clin Pharmacol Ther [Internet]. 2022 Jan 1 [cited 2022 May 31];111(1):321–31. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/cpt.2476
21. Data on SARS-CoV-2 variants in the EU/EEA [Internet]. [cited 2022 May 31]. Available from: https://www.ecdc.europa.eu/en/publications-data/data-virus-variants-covid-19-eueea

22. Jin P, Li J, Pan H, Wu Y, Zhu F. Immunological surrogate endpoints of COVID-2019 vaccines: the evidence we have versus the evidence we need. Signal Transduction and Targeted Therapy 2021 6:1 [Internet]. 2021 Feb 2 [cited 2023 Apr 21];6(1):1–6. Available from: https://www.nature.com/articles/s41392-021-00481-y
23. Feng S, Phillips DJ, White T, Sayal H, Aley PK, Bibi S, et al. Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection. Nature Medicine 2021 27:11 [Internet]. 2021 Sep 29 [cited 2023 Apr 21];27(11):2032–40. Available from: https://www.nature.com/articles/s41591-021-01540-1
24. Schmidt T, Klemis V, Schub D, Mihm J, Hielscher F, Marx S, et al. Immunogenicity and reactogenicity of heterologous ChAdOx1 nCoV-19/mRNA vaccination. Nature Medicine 2021 27:9 [Internet]. 2021 Jul 26 [cited 2023 Apr 21];27(9):1530–5. Available from: https://www.nature.com/articles/s41591-021-01464-w
25. Hillus D, Schwarz T, Tober-Lau P, Vanshylla K, Hastor H, Thibeault C, et al. Safety, reactogenicity, and immunogenicity of homologous and heterologous prime-boost immunisation with ChAdOx1 nCoV-19 and BNT162b2: a prospective cohort study. Lancet Respir Med [Internet]. 2021 Nov 1 [cited 2023 Apr 21];9(11):1255–65. Available from: http://www.thelancet.com/article/S221326002100357X/fulltext
26. Tenbusch M, Schumacher S, Vogel E, Priller A, Held J, Steininger P, et al. Heterologous prime–boost vaccination with ChAdOx1 nCoV-19 and BNT162b2. Lancet Infect Dis [Internet]. 2021 Sep 1 [cited 2023 Apr 21];21(9):1212–3. Available from: http://www.thelancet.com/article/S1473309921004205/fulltext
27. 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. The Lancet Regional Health - Europe. 2021 Dec 1;11:100249.
28. Hermosilla E, Coma E, Xie J, Feng S, Cabezas C, Méndez-Boo L, et al. Comparative effectiveness and safety of homologous two-dose ChAdOx1 versus heterologous vaccination with ChAdOx1 and BNT162b2. Nature Communications 2022 13:1 [Internet]. 2022 Mar 23 [cited 2023 Apr 21];13(1):1–7. Available from: https://www.nature.com/articles/s41467-022-29301-9
29. ECDC. COVID-19 vaccine effectiveness in adolescents aged 12–17 years and interim public health considerations for administration of a booster dose [Internet]. 2022 [cited 2023 Apr 21]. Available from: https://www.ecdc.europa.eu/en/publications-data/covid-19-vaccine-effectiveness-adolescents-and-interim-considerations-for-booster-dose
30. Andrews N, Stowe J, Kirsebom F, Toffa S, Sachdeva R, Gower C, et al. Effectiveness of BNT162b2 COVID-19 booster vaccine against covid-19 related symptoms and hospitalization in England. Nat Med. 2022 Jan 14;
31. Andrews N, Stowe J, Kirsebom F, Toffa S, Rickeard T, Gallagher E, et al. Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant. New England Journal of Medicine [Internet]. 2022 Apr 21 [cited 2023 Apr 21];386(16):1532–46. Available from: https://www.nejm.org/doi/full/10.1056/NEJMoa2119451
32. Monge S, Pastor-Barriuso R, Hernán MA. The imprinting effect of COVID-19 vaccines: an expected selection bias in observational studies. medRxiv [Internet]. 2023 Feb 13 [cited 2023 Apr 21];2022.11.30.22282923. Available from: https://www.medrxiv.org/content/10.1101/2022.11.30.22282923v2
33. Lin KY, Hsieh MJ, Chang SY, Ieong SM, Cheng CY, Sheng WH, et al. Serological response after COVID-19 mRNA-1273 booster dose in immunocompromised patients, Taiwan, July to August 2021. Journal of the Formosan Medical Association. 2022 Dec 1;121(12):2438–45.
34. Kamar N, Abravanel F, Marion O, Couat C, Izopet J, Del Bello A. Three Doses of an mRNA Covid-19 Vaccine in Solid-Organ Transplant Recipients. New England Journal of Medicine [Internet]. 2021 Aug 12 [cited 2023 Apr 21];385(7):661–2. Available from: https://www.nejm.org/doi/full/10.1056/NEJMc2108861
35. Manjappa S, Phi HQ, Lee LW, Onstad L, Gill DB, Connelly-Smith L, et al. Humoral and Cellular Immune Response to Covid-19 Vaccination in Patients with Chronic Graft-versus-Host Disease on Immunosuppression. Transplant Cell Ther. 2022 Nov 1;28(11):784.e1-784.e9.
36. Fowlkes AL, Yoon SK, Lutrick K, Gwynn L, Burns J, Grant L, et al. Effectiveness of 2-Dose BNT162b2 (Pfizer BioNTech) mRNA Vaccine in Preventing SARS-CoV-2 Infection Among Children Aged 5–11 Years and Adolescents Aged 12–15 Years — PROTECT Cohort, July 2021–February 2022. MMWR Morb Mortal Wkly Rep [Internet]. 2022 [cited 2023 Apr 21];71(11):422–8. Available from: https://www.cdc.gov/mmwr/volumes/71/wr/mm7111e1.htm
37. Effectiveness Studies | ViewHub [Internet]. [cited 2023 Apr 21]. Available from: https://view-hub.org/vaccine/covid/effectiveness-studies
38. Amir O, Goldberg Y, Mandel M, Bar-On YM, Bodenheimer O, Freedman L, et al. Initial protection against SARS-CoV-2 omicron lineage infection in children and adolescents by BNT162b2 in Israel: an observational study. Lancet Infect Dis [Internet]. 2023 Jan 1 [cited 2023 Apr 21];23(1):67–73. Available from: http://www.thelancet.com/article/S1473309922005278/fulltext
39. Bundle N, Dave N, Pharris A, Spiteri G, Deogan C, Suk JE, et al. COVID-19 trends and severity among symptomatic children aged 0-17 years in 10 European Union countries, 3 August 2020 to 3 October 2021. Eurosurveillance [Internet]. 2021 Dec 16 [cited 2023 Apr 21];26(50):2101098. Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2021.26.50.2101098
40. Comirnaty COVID-19 vaccine: EMA recommends approval for children aged 5 to 11 | European Medicines Agency [Internet]. [cited 2023 Apr 21]. Available from: https://www.ema.europa.eu/en/news/comirnaty-covid-19-vaccine-ema-recommends-approval-children-aged-5-11
41. Public Health Agency of Canada. Recommendations on the use of Moderna Spikevax COVID-19 vaccine in children 6 months to 5 years of age. :27.
42. Lin DY, Gu Y, Xu Y, Zeng D, Wheeler B, Young H, et al. Effects of Vaccination and Previous Infection on Omicron Infections in Children. New England Journal of Medicine [Internet]. 2022 Sep 22 [cited 2023 Apr 21];387(12):1141–3. Available from: https://www.nejm.org/doi/full/10.1056/NEJMc2209371