

# Comparing results from a traditional multivariable model and seven propensity score-based models for estimating COVID-19 vaccine effectiveness

Daniele Petrone<sup>1</sup>, Chiara Sacco<sup>1</sup>, Alberto Mateo-Urdiales<sup>1</sup>, Emmanouil Alexandros Fotakis<sup>1</sup>, Marco Alfò<sup>2</sup>, Patrizio Pezzotti<sup>1</sup> and Massimo Fabiani<sup>1</sup>

<sup>1</sup>Dipartimento di Malattie Infettive, Istituto Superiore di Sanità, Rome, Italy

<sup>2</sup>Dipartimento di Statistica, Sapienza Università di Roma, Rome, Italy

## Abstract

**Background.** Accurate estimation of vaccine effectiveness (VE) in real-world settings is essential for guiding immunization strategies, especially in older populations. However, observational studies are prone to bias due to confounding factors, and the choice of statistical method can significantly influence VE estimates.

**Materials and methods.** We compared the performance of a multivariable Cox proportional hazards model with seven propensity score (PS)-based models to estimate the relative vaccine effectiveness (rVE) of the bivalent Original/Omicron BA.4-5 mRNA vaccine as a second or third booster, compared to a first mRNA booster received  $\geq 120$  days earlier. Data from 11,879,461 individuals aged  $\geq 60$  in Italy (April-June 2023) were analyzed.

**Results.** All models produced consistent rVE estimates, with values ranging from 16.4% to 22.1%. Over time, booster effectiveness declined, with the reference model showing a drop in rVE from 45.6% (15-60 days) to 14.3% (181-265 days). PS-based methods improved covariate balance but did not outperform the Cox model in terms of precision or interpretability.

**Conclusions.** In large, relatively balanced datasets, traditional multivariable models remain a robust and reliable choice for estimating VE. While PS-based methods offer theoretical advantages, their practical benefit may be limited when confounding is well controlled.

## Key words

- COVID-19
- effectiveness
- propensity score
- matching

## BACKGROUND

Vaccination against COVID-19 was considered the key strategy in limiting the spread of the highly mutable SARS-CoV-2 virus. In this context, observational studies remain crucial to assess and monitor the effectiveness of approved vaccines. However, observational vaccine effectiveness (VE) studies, although crucial for understanding the real-world impact of vaccines, are prone to bias (e.g., due to confounding) [1, 2]. Compared to randomized trials, where all participants have an equal probability to receive the vaccine, in observational studies the choice to get vaccinated or not is not evenly distributed across the population. Indicatively, confounding factors like age, pre-existing health conditions, comorbidities, and socio-economic background

influence the individual decision to get vaccinated and the probability of a COVID-19 associated event, thus introducing potential bias in the VE estimates [3]. The risk of bias due to differences between vaccinated and unvaccinated populations could be amplified by the seasonal booster campaigns implemented in Italy and other European countries, targeting the elderly population and those with risk factors.

Several statistical methods have been proposed in the literature to consider the presence of potential confounding factors, aiming to balance the difference between those who received the intervention and those who did not. These methods include covariates adjustment, stratification, matching, regression models that use instrumental variables, and methods based on

propensity score (PS), representing the probability of receiving a specific treatment, given a set of observed covariates [4].

The performance of these methods in limiting bias could significantly impact the accuracy of VE estimates, particularly in the context of seasonal COVID-19 vaccine programmes. Nonetheless, there is a lack of studies comparing VE estimates considering potential confounding, especially in the context of large, real-world datasets.

Based on published data recorded in Italy [5], we aimed to compare estimates of relative VE (rVE) using different PS-based methods – i.e., inclusion of the PS as an adjustment variable, inverse probability weight (IPW), stabilized IPW, truncated stabilized inverse probability weights (STIPW) and matching – [6-12] with a reference approach based on a Cox proportional hazards regression analysis.

**MATERIALS AND METHODS**

**Study design**

We conducted a retrospective cohort study based on the individual data previously used by our study group to estimate the rVE of a booster dose of the bivalent (Original/Omicron BA.4-5) mRNA vaccine in preventing severe COVID-19 among people ≥60 years during a period with prevalent circulation of the Omicron XBB subvariant (3 April-4 June 2023) [5].

Specifically, we analyzed time to infection leading to severe COVID-19 consequences (i.e., hospitalization or death within 28 days of the swab date for a test that was SARS-CoV-2-positive) comparing individuals who had received only the first booster dose of an mRNA vaccine at least 120 days earlier and those who had received a second or third booster dose of the bivalent Original/Omicron BA.4-5 mRNA vaccine during the study period.

After splitting individual data into multiple records to account for the time-varying vaccination status, we estimated rVE both overall and by time since booster administration (i.e., 15-60 days, 61-120 days, 121-180 days and 181-265 days).

More details about the data sources, the selection of the study population, and the outcome classification have been published elsewhere [5].

**Statistical analysis**

**Reference model**

The reference model (M0) used in the original analysis was a multivariable Cox proportional hazards model M0 include the following demographic and clinical variables as covariates to adjust for: sex, age group (5-year age group from 60-64 to 90-94, then grouping individuals ≥95), country of birth (Italian-born or foreign-born), geographical area where the vaccination took place (NUTS 21 regions/autonomous provinces of Italy), and high-risk group (none or individuals residing in long-term care facilities and/or with health risk conditions; see *Supplementary Table S1 available online*). Additional covariates included the days since 1st booster dose (≤180, 181-365, >365), days since prior infection (≤180, 181-365, >365, no prior infection) and type of 1st booster vaccine (mRNA monovalent, mRNA bivalent Original/Omicron BA.1, mRNA bivalent Original/Omicron BA.4-5).

**Alternative models**

In order to take into account potential imbalance between the analyzed groups in the observed covariates we estimated adjusted VE against severe COVID-19 using seven alternative PS-based Cox regression models (M1-M7), as reported in *Table 1*. The alternative models are still Cox proportional hazards, but instead of using covariates for adjustment (as in the reference

**Table 1**  
Models description and covariates included in the study

Model name	Differences with respect to baseline Cox proportional hazard model	Covariates included
M0	-	Exposure level, sex, age group, country of birth, geographical area, high-risk group, days since 1st booster vaccine, days since prior infection, type of 1st booster vaccine
M1	Including (only) the PS	Exposure level, PS
M2	Including (also) the PS	Exposure level, sex, age group, country of birth, geographical area, high-risk group, days since 1st booster vaccine, days since prior infection, type of 1st booster vaccine, PS
M3	Weighting observations with IPW	Exposure level
M4	Weighting observations with SIPW	Exposure level
M5	Weighting observations with STIPW 1%	Exposure level
M6	Weighting observations with STIPW 5%	Exposure level
M7	Matching observations	Exposure level

Exposure level: time since administration of a second or third booster dose of the bivalent Original/Omicron BA.4-5 mRNA vaccine: 15-60 days, 61-120 days, 121-180 days and 181-265 days; sex: male, female; age group: 5-year age group from 60-64 to 90-94, then grouping individuals ≥90; country of birth: Italian-born or foreign-born; Geographical area: geographical area where the vaccination took place (NUTS 2, which are the 19 regions and two autonomous provinces of Italy); high-risk group: none or residents in long-term care facilities and individuals with health risk conditions (*Supplementary Table S1 available online*); days since 1st booster vaccine: ≤180 days, 181-365 days, >365 days; days since prior infection: ≤180 days, 181-365 days, >365 days, no prior infection; type of 1st booster vaccine: mRNA monovalent, mRNA bivalent Original/Omicron BA.1, mRNA bivalent Original/Omicron BA.4-5.  
PS: propensity score; IPW: inverse probability weight; SIPW: stabilized inverse probability weights; STIPW: stabilized truncated inverse probability weights.

model), covariates are used for the calculation of PS, weighting and matching.

To calculate the PS we used a logistic regression model based on individual data, estimating the probability of receiving a second or third booster dose of the bivalent Original/Omicron BA.4-5 mRNA vaccine as a function of the covariates used in model M0. The first alternative model (M1) was based on adjusting for the propensity score (PS), used as the only covariate in addition to the exposure variable (vaccination status). The second model (M2) was adjusted including the covariates considered in model M0 plus the PS. In models M3-M6, observations were weighted and only the exposure variable was included. Specifically, in model M3, observations were weighted using the inverse probability weighting approach (IPW) with individual weights defined as  $1/PS$  if the subject received the treatment and  $1/(1-PS)$  otherwise. To account for the different size of the treated/untreated groups and to reduce the variance of the estimated hazard ratios (HRs), we calculated the stabilized inverse probability weights (SIPW) by dividing the IPW by the total sum of the IPWs of the subjects in the same treatment group; and used these weights in model M4. To further reduce the variance due to extreme values of the PS and to study the sensitivity of the model to these values, we calculated the stabilized truncated inverse probability weights (STIPW) including only cases with PS between the 1st and the 99th percentile (M5) or between the 5th and 95th percentile (M6) when calculating SIPW. Finally, we matched 1:1 without replacement the observations according to the PS, before splitting the data, using the nearest neighbour matching with a calliper of 0.05, and then run a univariate Cox regression analysis on split data from the matched sample (M7).

All models based on IPW (M3-M6) were run using a robust variance estimator [6]. To verify the covariates balance after IPW, SIPW, STIPW and PS-matching (M3-M7) we used the Standardized Mean Difference (SMD), assuming a satisfactory balance where SMD values were less than 0.1.

For all models, the HRs and 95% confidence intervals (CI) were used to calculate the adjusted rVE against

severe COVID-19 using the formula  $((1-HR) \times 100)$ . In addition, we calculated the BIC and AIC for all models to compare their performance. All the analyses were carried out with RStudio 2023.12.1+402 under R 4.3.2 [13].

## RESULTS

We analysed data on 11,879,461 subjects included in the original study [2], of which 1,970,598 (16.6%) received at least the second booster dose. Among all subjects, there were 4,405 (0.04%) severe events, most of which occurred before receiving a second or third booster dose ( $n=3,565$ ; 0.036%). After receiving a second or third booster doses, the highest event rate was observed in the time interval 121-180 days post administration ( $n=467$ ; 0.025%) (Table 2).

All the coefficient estimates obtained with the logistic regression model used to calculate the PS were statistically significant, although there was not much heterogeneity between the two treatment groups (Figure 1). The estimated area under the curve (AUC) obtained with the model is 0.72.

After matching (3,941,196 subjects, 33.2% of the total, including all vaccinated individuals, were matched successfully), all the potential confounding variables appeared balanced according to the SMD values ( $SMD < 0.1$ ). The same was observed after weighting data through IPW, SIPW and the STIPW (truncation at level 1% removed 239,155 subjects, 2.01% of the total; truncation at level 5% removed 1,217,007 individuals, 10.24% of the total) (Figure 2 and Supplementary Figure S1 available online).

### Model comparisons

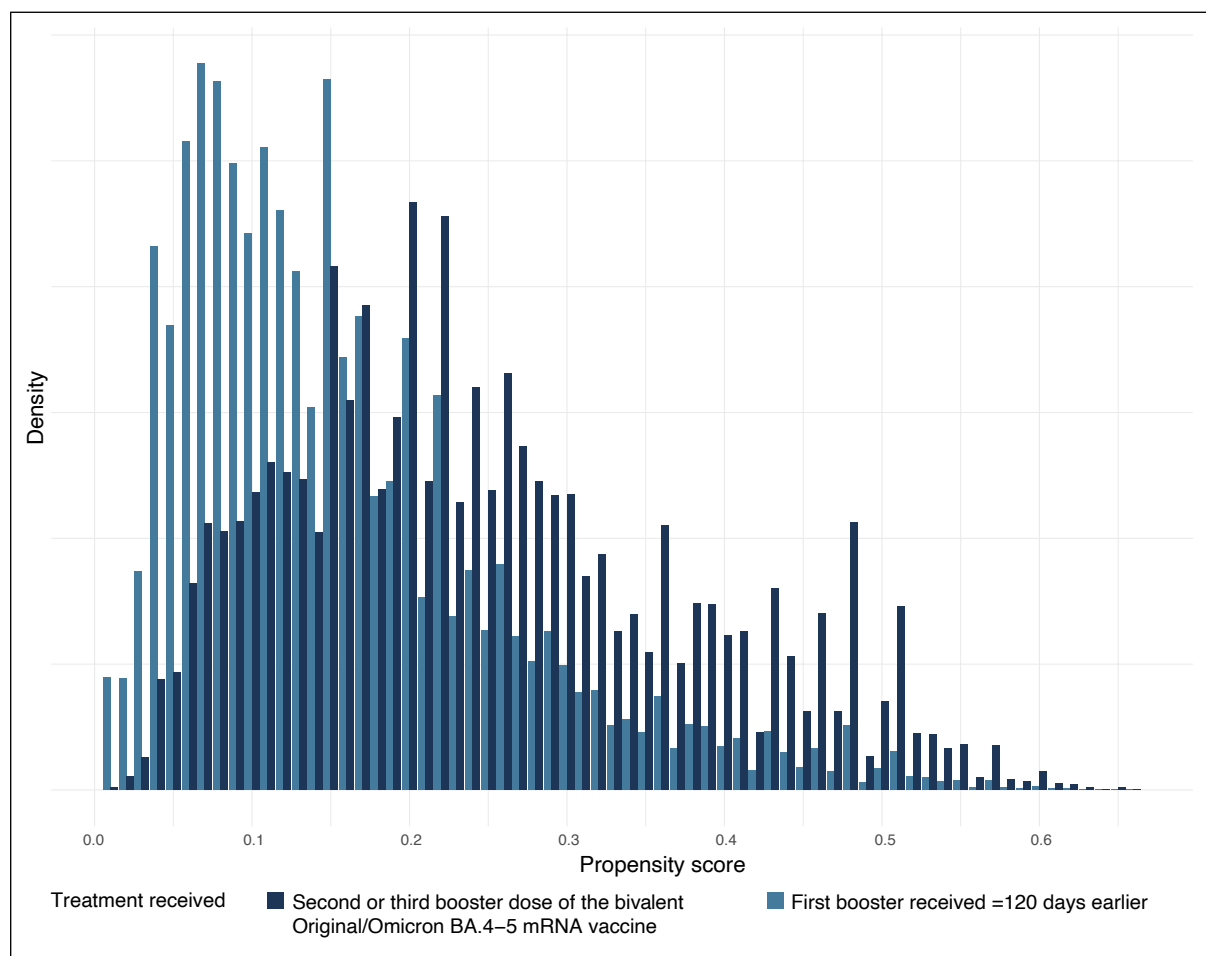
The overall analysis shows homogeneous results between all models, with the point estimate going from a minimum of 16.4% with model M7 to 22.1% with model M5. The 95% CI width also does not change much increasing slightly when the observations in the models M3-M6 are weighted (from 12.7 in the models M0-M2 to 15.4 in model M7). Both BIC and AIC, using the “lower is better” rule, indicate as better performing the models based on SIPW, with little difference between

**Table 2**  
Number and percentage of severe events by exposure level

Time since booster	Number of individuals <sup>a</sup>	Number of events	Percentage (%)
<b>First booster <math>\geq 120</math> days</b>	9,912,672	3,565	0.036
<b>Second or third booster</b>			
15-60 days	90,807	11	0.012
61-120 days	715,375	135	0.019
121-180 days	1,843,177	467	0.025
181-265 days	1,259,968	227	0.018
<b>Total</b>	<b>13,821,999</b>	<b>4,405</b>	<b>0.032</b>

<sup>a</sup>The sum of the number of individuals in each time interval is greater than the total number of individuals included in the analysis because of the possible change in vaccination status during the observation period leading an individual to be counted more than once.

First booster vaccine: Comirnaty monovalent Original strain mRNA vaccine (BNT162b2 mRNA, BioNTech-Pfizer); Spikevax monovalent Original strain mRNA vaccine (mRNA-1273, Moderna); Comirnaty bivalent Original/Omicron BA.1 mRNA vaccine (BNT162b2 mRNA, BioNTech-Pfizer); Spikevax bivalent Original/Omicron BA.1 mRNA vaccine (mRNA-1273.214, Moderna) or Comirnaty bivalent Original/Omicron BA.4-5 mRNA vaccine (BNT162b2 mRNA, BioNTech-Pfizer). Second or third booster vaccine: Comirnaty bivalent Original/Omicron BA.4-5 mRNA vaccine (BNT162b2 mRNA, BioNTech-Pfizer).



**Figure 1**

Density distribution of the propensity score for individuals aged  $\geq 60$  years who received either a second or third booster dose of the bivalent Original/Omicron BA.4-5 mRNA vaccine, Italy, 3 April-4 June 2023 ( $n=11,879,461$ ).

using truncation or not (Table 3 and Supplementary Figure S2 available online).

The results from the reference model (M0) show that, compared with a first booster of an mRNA vaccine received at least 120 days earlier, the rVE against severe COVID-19 for a second or a third booster dose of bivalent Original/Omicron BA.4-5 mRNA vaccine in the time interval 15-60 days post-administration was 45.6% (95% CI: 1.6-69.9). Relative VE progressively decrease to 14.3% (95% CI: 1.6-25.3) in the interval 181-265 days post-administration (Table 3 and Figure 3). The same decreasing trend is shown by all the models, except models M5 and M6 where we used the STIPW. In these cases, the rVE obtained in the first time-interval post-administration is lower than that obtained in the second one. The rVE estimates with 95% CIs obtained with models M1, M2 and M7 are similar to those obtained with model M0, although models M1 and M7 showed slightly lower estimates than M0 in the first time-interval post-administration (42.3% and 39.6%, respectively, both with a 95% lower bound of opposite sign than M0). In addition, model M1 shows slightly higher estimates in the other time-intervals. Model M2 shows estimates very close to those from model M0

(45.6% vs 45.8% in the 15-60 days interval, 24.7% vs 25.1% in the 61-120 days interval, 17.0% vs 17.3% in the 121-180 days interval, and 14.3% vs 14.5 in the 181-265 days interval). There are no differences between the rVE estimated with model M3 with respect to those estimated using model M4. The rVE estimates and 95% CIs obtained with the IPW-based models M3-M6 are generally similar to those derived from the reference model M0, except in the first period post-administration (15-60 days). This applies especially to model M6, which estimates a rVE more than 6 times lower and with a much higher variability than that derived from M0 (7.1%, 95% CI: -91.2-54.9). In the other time-intervals post-administration the differences appear much reduced, with slightly higher rVE estimates and slightly larger 95% CIs for models M3-M6 compared to M0. Looking at the AIC and BIC values, conclusions are similar to those made previously for the overall analysis (Table 3).

## DISCUSSION

Our study showed homogeneous VE estimates among all the models. In particular, two groups appear, whereby the rVE values estimated by models including



**Figure 2**  
Standardized mean differences for covariates used in the models before and after observation matching (A), and weighting with IPW (B) and weighting with SIPW (C).  
In the analysis we considered the Regions or Autonomous Provinces (NUTS 2) but for graphical reasons we represented the macroarea (NUTS 1). In both cases the covariates are balanced. IPW: inverse probability weight; SIPW: stabilized inverse probability weights.

**Table 3**

Point Estimation, 95% confidence interval and interval width of the estimated rVE against severe COVID-19 of a second or third booster of the bivalent Original/Omicron BA.4-5 mRNA vaccine relative to a first booster of an mRNA vaccine received at least 120 days earlier, Italy, 3 April-4 June 2023

Time since 2nd or 3rd booster		M0	M1	M2	M3	M4	M5	M6	M7
15-60 days	Point Es.	45.6	42.3	45.8	34.5	34.5	18.6	7.1	39.6
	95% CI	1.6-69.9	-4.5-68.1	2.0-70.1	-32.2-67.5	-32.2-67.5	-64.0-59.6	-91.2-54.9	-9.5-66.7
	In. Width	68.3	72.6	68.1	99.7	99.7	>100	>100	76.2
61-120 days	Point Es.	24.7	25.8	25.1	30.7	30.7	31.6	27.9	23.8
	95% CI	10.5-36.7	11.7-37.7	10.8-37.0	16.0-42.9	16.0-42.9	16.5-44.0	10.8-41.7	8.7-36.5
	In. Width	26.2	26.0	26.2	26.9	26.9	27.5	30.9	27.8
121-180 days	Point Es.	17.0	18.3	17.3	21.2	21.2	22.0	21.4	15.5
	95% CI	8.4-24.8	9.7-26.0	8.7-25.1	11.9-29.5	11.9-29.5	12.6-30.4	10.9-30.7	5.7-24.3
	In. Width	16.4	16.3	16.4	17.6	17.6	17.8	19.8	18.6
181-265 days	Point Es.	14.3	17.0	14.5	15.5	15.5	16.4	17.9	11.3
	95% CI	1.6-25.3	4.7-27.7	1.9-25.5	0.0-28.5	0.0-28.5	0.7-29.6	1.2-31.8	-2.8-23.4
	In. Width	23.7	23.0	23.6	28.5	28.5	28.9	30.6	26.2
Overall	Point Es.	18.2	19.7	18.5	21.6	21.6	22.1	21.3	16.4
	95% CI	11.6-23.3	13.1-25.8	11.9-24.6	14.5-28.2	14.5-28.2	14.8-28.7	13.3-28.6	8.3-23.7
	In. Width	12.7	12.7	12.7	13.7	13.7	13.9	15.3	15.4
AIC (overall)		138,461.7	142,356.6	138,461.5	278,371.3	2.0	2.0	2.0	47,330.3
BIC (overall)		138,704.6	142,369.3	138,710.7	278,377.7	8.4	8.4	8.2	47,335.9

Point Es.: Point Estimation; 95% CI: 95% confidence interval; In. Width: Interval Width; AIC: Akaike's information criterion; BIC: Bayesian information criterion; M0: multivariable model; M1: PS model; M2: PS and M0 covariates; M3: IPW; M4: SIPW; M5: STIPW 1%-99%; M6: STIPW 5%-95%; M7: matching model; rVE: relative vaccine effectiveness; PS: propensity score; IPW: inverse probability weight; SIPW: stabilized inverse probability weights; STIPW: stabilized truncated inverse probability weights.

PS as a covariate (M1 and M2) and by the model run after PS-matching (M7) are very close to the rVE values estimated by the reference multivariable model (M0). On the other hand, PS models based on IPW (M3), SIPW (M4) and STIPW (M5 and M6) yield estimates that are further away from the reference multivariable model (M0), especially in the early time-interval after vaccination for STIPW models (M5 and M6).

According to several studies [9-11], SIPW (M4) can increase the variance estimate compared to that obtained with IPW (M3). Our results show the same estimated values, probably due to the homogeneity (i.e., absence of excessive skewness) of the PS values between the treated and untreated groups. A factor that influenced the difference of the estimates obtained with models M5 and M6 compared to the other models is related to having removed records with extreme PS values from the sample (truncation at level 1% removed 239,155 individuals, 2.01% of the total; truncation at level 5% removed 1,217,007 individuals, 10.24% of the total). Truncation is used to improve estimation accuracy by truncating extreme values of the PS [12], many of them falling in the first exposure level (15-60 days after vaccine uptake), when the lowest percentage of severe events (study outcome) was observed.

Based on our results, we found no clear advantage in applying PS-based methods instead of multivariable

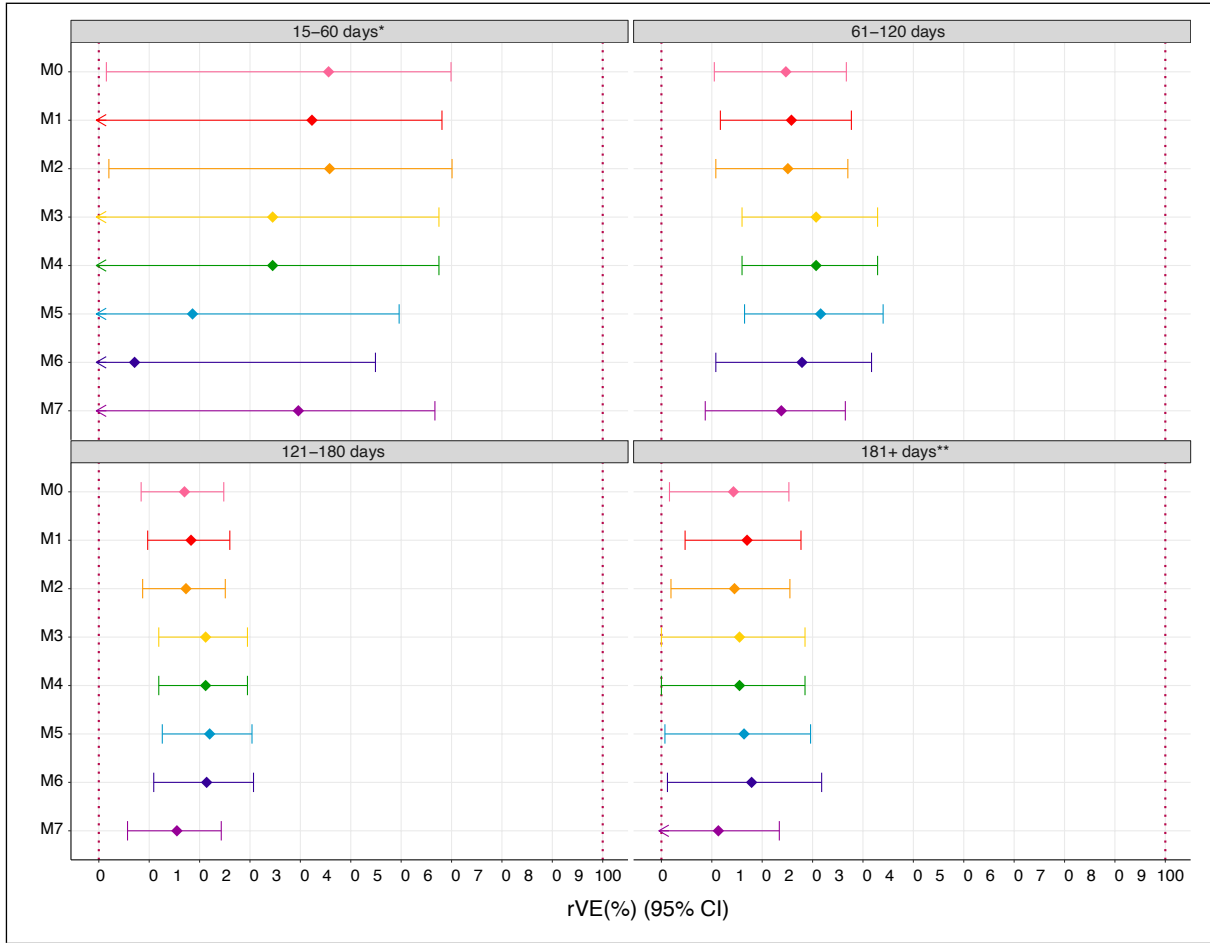
analysis. Previous studies have shown that these methodologies can produce similar results under certain conditions [14, 15], especially when the starting dataset is relatively balanced and the sample size is large [8, 16], as in our case. Moreover, it is important to note that, even using PS methods, a residual bias due to unmeasured confounders might remain, as with all analytic methods based on observational data [17, 18].

Although PS-based methods are sensitive to misspecifications of the model used to calculate it (e.g., omission of interaction effects, or misspecification of functional form of the included covariates [19]), they may still prove advantageous when many covariates are considered as potential confounders in the analysis, in which case reference models might result unstable. Furthermore in addition to the PS-based methods presented in this article, there are other PS-based methods in the literature, aiming to optimise survival data [20] and time-dependent exposure analyses [21-23]. The purpose of this study, however, was to present and compare the most widely used methods in the literature.

## CONCLUSIONS

Our study, based on a large study population and a relatively low number of potential confounders, does not show advantages in using PS-based models compared to using a traditional multivariable model.





**Figure 3** Point estimation and 95% confidence interval of rVE against severe COVID-19 of a second or third booster of the bivalent Original/Omicron BA.4-5 mRNA vaccine relative to a first booster of an mRNA vaccine received at least 120 days earlier with truncated lower bounds for visual clarity, Italy, 3 April-4 June 2023.

\*For graphical reasons, the negative lower bounds were set at -0.5. Below are the true values: M1: -4.5%; M3: -32.2%; M4: -32.3%; M5: -64.0%; M6: -91.2%; M7: -9.5%.

\*\*For graphical reasons, the negative lower bounds were set at -0.5. Below are the true values: M7: -2.8%.

### Acknowledgments

Preliminary results of this study were presented as an abstract at the PNRR meeting held in Pavia, Lombardy, Italy, 11-12 September 2024. Daniele Petrone has a temporary PNRR contract since 19 July 2023.

### Authors' contributions

PP, MF and DP designed the paper; DP, CS, AMU and MF ensured quality of COVID-19 surveillance; DP, supported by MF, AMU, CS and EAF, carried out the analysis; DP, MF, MA, AMU, EAF, CS and PP wrote the manuscript which was reviewed and approved by all Authors.

### Conflict of interest statement

The Authors declare that they have no competing interests.

### Funding source

This research was supported by EU funding within the NextGeneration EU-MUR PNRR Extended Partnership Initiative on Emerging Infectious Diseases (Project n. PE00000007 INF-ACT).

Received on 10 March 2025.

Accepted on 9 October 2025.

### REFERENCES

1. Kahn R, Schrag SJ, Verani JR, et al. Identifying and alleviating bias due to differential depletion of susceptible people in postmarketing evaluations of COVID-19 vaccines. *Am J Epidemiol.* 2022;191(5):800-11. doi: 10.1093/aje/kwac015
2. Barda N. The ups and downs of observational vaccine research. *Lancet Infect Dis.* 2023;23(7):767-8. doi: 10.1016/S1473-3099(23)00119-6
3. Primieri C, Bietta C, Giacchetta I, et al. Determinants of COVID-19 vaccination acceptance or hesitancy in

- Italy: an overview of the current evidence. *Ann Ist Super Sanità*. 2023;59(1):10-25. doi: 10.4415/ANN\_23\_01\_03
4. Austin PC. An Introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399-424. doi: 10.1080/00273171.2011.568786
  5. Fabiani M, Mateo-Urdiales A, Sacco C, et al. Relative effectiveness of bivalent original/omicron BA.4-5 mRNA vaccine in preventing severe COVID-19 in persons 60 years and above during SARS-CoV-2 Omicron XBB.1.5 and other XBB sublineages circulation, Italy, April to June 2023. *Eurosurveill*. 2023;28(32):2300397. doi: 10.2807/1560-7917.ES.2023.28.32.2300397
  6. Austin PC. Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. *Stat Med*. 2016;35(30):5642-55. doi: 10.1002/sim.7084
  7. World Health Organization. Evaluation of COVID-19 vaccine effectiveness: interim guidance, 17 March 2021. Geneva: WHO; 2021. Available from: <https://iris.who.int/bitstream/handle/10665/340301/WHO-2019-nCoV-vaccine-effectiveness-measurement-2021.1-eng.pdf?sequence=1>.
  8. Ioannidis JPA. Factors influencing estimated effectiveness of COVID-19 vaccines in non-randomised studies. *BMJ Evid Based Med*. 2022;27(6):324-9. doi: 10.1136/bmjebm-2021-111901
  9. Pezzi A, Cavo M, Biggeri A, et al. Inverse probability weighting to estimate causal effect of a singular phase in a multiphase randomized clinical trial for multiple myeloma. *BMC Med Res Methodol*. 2016;16(1):150. doi: 10.1186/s12874-016-0253-9
  10. Westreich D, Cole SR, Schisterman EF, et al. A simulation study of finite-sample properties of marginal structural Cox proportional hazards models. *Stat Med*. 2012;31(19):2098-109. doi: 10.1002/sim.5317
  11. Xu S, Ross C, Raebel MA, et al. Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value Health*. 2010;13(2):273-7. doi: 10.1111/j.1524-4733.2009.00671.x
  12. Crump RK, Hotz VJ, Imbens GW, et al. Dealing with limited overlap in estimation of average treatment effects. *Biometrika*. 2009;96(1):187-99. doi: 10.1093/biomet/asn055
  13. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2023. Available from: <https://www.r-project.org/>.
  14. Biondi-Zoccai G, Romagnoli E, Agostoni P, et al. Are propensity scores really superior to standard multivariable analysis? *Contemp Clin Trials*. 2011;32(5):731-40. doi: 10.1016/j.cct.2011.05.006
  15. Shah BR, Laupacis A, Hux JE, et al. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. *J Clin Epidemiol*. 2005;58(6):550-9. doi: 10.1016/j.jclinepi.2004.10.016
  16. Cepeda MS, Boston R, Farrar JT, et al. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol*. 2003;158(3):280-7. doi: 10.1093/aje/kwg115
  17. Xu S, Huang R, Sy LS, et al. A safety study evaluating non-COVID-19 mortality risk following COVID-19 vaccination. *Vaccine*. 2023;41(3):844-54. doi: 10.1016/j.vaccine.2022.12.036
  18. Amoah J, Stuart EA, Cosgrove SE, et al. Comparing propensity score methods versus traditional regression analysis for the evaluation of observational data: a case study evaluating the treatment of gram-negative bloodstream infections. *Clin Infect Dis*. 2020;71(9):e497-e505. doi: 10.1093/cid/ciaa169
  19. Chesnaye NC, Stel VS, Tripepi G, et al. An introduction to inverse probability of treatment weighting in observational research. *Clin Kidney J*. 2021;15(1):14-20. doi: 10.1093/ckj/sfab158
  20. Chen W-C, Lu N, Wang C, et al. Propensity score-integrated approach to survival analysis: leveraging external evidence in single-arm studies. *J Biopharm Stat*. 2022;32(3):400-13. doi: 10.1080/10543406.2022.2080701
  21. Thomas LE, Yang S, Wojdyla D, et al. Matching with time-dependent treatments: a review and look forward. *Stat Med*. 2020;39(17):2350-70. doi: 10.1002/sim.8533
  22. Zhang Z, Li X, Wu X, et al. Propensity score analysis for time-dependent exposure. *Ann Transl Med*. 2020;8(5):246. doi: 10.21037/atm.2020.01.33
  23. Wijn SRW, Rovers MM, Hannink G. Confounding adjustment methods in longitudinal observational data with a time-varying treatment: a mapping review. *BMJ Open*. 2022;12(3):e058977. doi: 10.1136/bmjopen-2021-058977