Journal of Biostatistics and Epidemiology

J Biostat Epidemiol. 2024;10(1): 124-131

Original Article

Estimating Causal Effect of Two-Dose COVID-19 Vaccination on Hospitalization: A Propensity Score Matching Approach

Mahboobeh Taherizadeh[1], Mohammad Taghi Shakeri[1*], Saeed Akhlaghi[1], Ehsan Mosa Farkhani[2,3]

[1]Department of Biostatistics, Faculty of Health, Mashhad University of Medical Sciences, Mashhad, Iran.
[2]Department of Epidemiology, School of Health, Mashhad University of Medical Sciences, Mashhad, Iran.
[3]Social Determinants of Health Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.

ARTICLE INFO

Received 23.01.2024
Revised 18.02.2024
Accepted 07.03.2024
Published 15.03.2024

Key words:
Propensity score matching;
Causal effect;
Observational study;
Logistic regression.

ABSTRACT

Introduction: To estimate the effectiveness of two-dose COVID-19 vaccination in reducing hospitalization, accounting for complex confounding factors in observational studies.

Methods: Researchers applied propensity score methods to adjust for confounding variables, comparing their performance to traditional covariate adjustment methods. Multiple Logistic Regression and Propensity Score Matching were employed to analyze the data, ensuring a balanced comparison between vaccinated and unvaccinated groups.

Results: Both analytical methods demonstrated a significant reduction in the likelihood of hospitalization among vaccinated individuals. The adjusted odds ratios (OR) were 0.29 (95% CI: 0.26, 0.31) via logistic regression and 0.32 (95% CI: 0.30, 0.34) using propensity score matching.

Conclusion: The study confirms the effectiveness of two-dose COVID-19 vaccination in decreasing hospitalization. It highlights the importance of using meticulous approaches like propensity score methods to assess real-world impacts in complex observational data settings.

Introduction

In the ongoing battle against COVID-19, vaccines have become vital tools to protect public health without promoting panic. Yet, assessing their true effectiveness in real-world scenarios presents formidable challenges, particularly within observational studies where the absence of randomization, a hallmark of clinical trials, complicates matters. In these studies, the lurking presence of confounding variables threatens to distort the genuine relationship between vaccination and outcomes, potentially skewing results if left unaddressed.

Amid this intricate landscape, propensity score analysis has risen as a stalwart methodological fortress, offering a robust defense against confounding in observational research.

*.Corresponding Author: shakerimt@mums.ac.ir

Propensity score analysis orchestrates balanced groups, similar to randomized allocation in clinical trials, by discerning the probability of treatment receipt, such as vaccination, based on observed baseline characteristics. This orchestration fosters a more precise evaluation of vaccination's causal impact on outcomes such as hospitalization, fortifying against the distorting effects of confounding factor.[1]

The imperative to comprehend the effectiveness of two-dose COVID-19 vaccination regimens in staving off hospitalization is paramount for guiding public health interventions. Thus, this paper embarks on a journey to explore the application of propensity score analysis in scrutinizing this very impact within the realm of observational studies. Through meticulous adjustments for potential confounders, our mission is to furnish robust estimates of vaccination's causal effect, providing invaluable insights into the real-world efficacy of vaccination strategies. Driven by rigorous methodological inquiry and exhaustive data scrutiny, this study endeavors to enrich the burgeoning corpus of evidence on COVID-19 vaccination efficacy. Ultimately, we aim to furnish policymakers with the knowledge necessary to craft strategies that effectively rein in the spread of the virus and curtail its associated burden of illness and mortality.

Methods

Data

The study population of this research comprised all individuals who underwent COVID-19 testing from the beginning of the year 1400 until the end of the same year within the community covered by Mashhad University of Medical Sciences, involving a total of 306630 individuals, aimed to assess the impact of a two dose COVID-19 vaccination on hospitalization. Data were collected from three databases: the Sina Health Information System, the Healthcare Services Monitoring System, and the Hospital Information System to ensure comprehensive coverage of the study population.

Study Design and Entry/Exit Criteria

Inclusivity criteria required individuals to have undergone at least one COVID-19 test, while exclusivity criteria excluded those residing outside the university's jurisdiction or receiving vaccination beyond the specified timeframe. Individuals with a vaccination-to-outcome interval of less than two weeks were also excluded. Participants were categorized as "exposed" if vaccinated and "unexposed" if not.

Statistical Methods

The propensity score (PS) for an individual represents the likelihood of being assigned to the "treatment" group based on all relevant covariates, denoted as

$$Pr(Z_k=1|X_k) \quad (1)$$

where $Z_k=1$ indicates the treatment assigned, and X_k is the vector of observed covariates.[2] Typically, a logistic regression is used to estimate propensity scores. The propensity score model includes both measured confounding variables and variables associated with the outcome. It's crucial to exclude variables solely associated with treatment decisions but not outcomes.[3] Additionally, predictive regression models should not include post-treatment decision variables. Consequently, age, body mass index, gender, marital status, education, occupation, place of residence, number of comorbidities, number of clinical symptoms, and history of COVID-19 infection were incorporated into the propensity score model. After calculating the PS, various methods can be used to estimate treatment effects. Propensity score matching (PSM), a widely used statistical technique in observational studies, pairs patients with similar PS to balance baseline characteristics between treated and control groups. In our study, PSM was employed to match vaccinated and unvaccinated patients using covariates associated with outcomes or true confounders. Following matching, balance checks were conducted using absolute standardized differences, while excluding patients outside the common support range. The standardized difference provides a means to compare the mean or prevalence of baseline covariates between treatment groups in the propensity score-matched sample. For continuous covariates, the standardized difference is defined as the difference in means divided by the pooled standard deviation. For dichotomous variables, the difference in prevalence or mean is compared in units of the pooled standard deviation. An ASMD exceeding 0.10 indicates a covariate imbalance.[1] The effect of vaccination was estimated using the average treatment effect (ATE) for the entire sample.[1, 3] Furthermore, vaccine effectiveness and adjusted baseline characteristics were determined through logistic regression modeling. Statistical analyses were performed using R 4.2.3 software, with a significance level set at ($p < 0.05$) and propensity score matching analysis was performed using the MatchIt package.

Outcome Model for propensity score-based method

Following propensity score matching, we meticulously assessed covariate balance to ensure equity between the groups. Leveraging logistic regression, our outcome model aimed to estimate the effect of vaccination on hospitalization. This analytical approach enabled us to examine the association between vaccination status and the probability of hospitalization while accounting for potential confounding variables. By scrutinizing covariate balance and employing rigorous statistical methods, we aimed to provide a comprehensive understanding of the effect of a two dose vaccination on hospitalization outcome in COVID-19 patients.[4]

Result

A comparison of baseline characteristics between vaccinated and unvaccinated groups is presented in Table 1. We included 306630 patients in the analysis, 104115 (33.95%) unvaccinated and 202515 (66.05%) vaccinated patients.

After incorporating the variables of Table 1 into the propensity score model and estimating propensity scores, propensity score matching was employed.

After applying propensity score matching, the analysis included a total of 46,686 pairs in the vaccinated group and 26,871 pairs in the unvaccinated group.

Subsequently, the balance in the distribution of confounding variables between the two groups was examined.

After employing the optimal full matching method and investigating the balance of confounding variables, the logistic regression model was formed as the outcome model. The effect of vaccination was assessed using both propensity score-based methods and logistic regression, with the results presented in Table 2. The results of Multivariate logistic regression (MLR) showed that vaccinated patients had significantly lesser odds of hospitalization [adjusted odds ratio (OR), 95% CI 0.29 (0.26, 0.32), p <0.001] compared to the unvaccinated group adjusted for age, gender, occupation, education, marital status, place of residence,

Table 1. Baseline characteristics in the vaccinated and non-vaccinated population

Parameters	Category	All (n=306630)	Vaccinated (n=104115)	Unvaccinated (n= 202515)	P value
Age		33.77±18.34	35.89±14.46	29.67±23.64	<0.001
Body Mass Index		23.92±10.90	25.10±5.38	21.63±16.91	
Gender	Female	159768 (52.10)	107127 (52.90)	52641 (50.56)	<0.001
	Male	146862 (47.90)	95388 (47.10)	51471 (49.44)	
Marital Status	Single	112073 (36.55)	53895 (26.60)	58214 (50.51))	<0.001
	Married	194557 (63.45)	148656 (73.40)	45901 (44.09)	
Education	Cycle & Below	169548 (55.29)	95142 (46.98)	74406 (71.47)	<0.001
	Diploma & BA	127670 (41.64)	100547 (49.65)	27123 (26.05)	
	Master & PhD	5984 (1.95)	4905 (2.42)	1079 (1.04)	
	Others	3428 (1.12)	1921 (0.95)	1507 (1.45)	
Occupation	Unemployed	172226 (56.17)	100717 (49.73)	71509 (68.68)	<0.001
	Employee	14292 (4.66)	12442 (6.14)	1850 (1.78)	
	Laborer	30112 (9.82)	24465 (12.08)	5647 (5.42)	
	Freelancer	44682 (14.57)	32859 (16.23)	11823 (11.36)	
	Others	45318 (14.78)	32032 (15.82)	13286 (12.76)	
Place of Residence	Rural	65789 (21.46)	46520 (22.97)	19269 (18.51)	<0.001
	Urban	240841 (78.54)	84846 (77.03)	202515 (81.49)	
Number of Comorbidities	0	270559 (88.24)	183131 (90.43)	87428 (83.97)	<0.001
	1	23671 (7.72)	14289 (7.06)	9382 (9.01)	
	2	10277 (3.35)	4664 (2.30)	5613 (5.39)	
	3 and more	2123 (0.69)	431 (0.21)	1692 (1.63)	
Number of Clinical Symptoms	0	278008(90.67)	201333 (99.42)	76675 (73.64)	<0.001
	1	12905(4.21)	533 (0.26)	12372 (11.88)	
	2	10144(3.31)	453 (0.22)	9691 (9.31)	
	3 and more	5573(1.82)	196 (0.10)	5377 (5.16)	
History of COVID-19 infection	Yes	139189 (45.39)	86307 (37.99)	52882 (50.79)	<0.001
	No	167441 (54.61)	116208 (57.38)	51233 (49.21)	

Table 2. Absolute Standardized Mean Difference values before and after propensity score matching

Variable	Before Matching	After Matching Logistic Regression
Age	0.13	0.031
Gender	0.01	0.003
Occupation	0.25	0.042
Education	0.36	0.002
Marital Status	0.21	0.005
Place of Residence	0.04	0.008
Body Mass Index	0.19	0.005
Number of Comorbidities	0.31	0.081
Number of Clinical Symptoms	0.68	0.075
History of COVID-19 infection	0.12	0.053

Table 3. Comparison of the results of Crude Logistic regression, MLR and PSM analysis for Hospitalization

Parameters	Crude Logistic Regression	P-value	MLR	P-value	PSM	P-value
	OR (95% CI)		OR (95% CI)		OR (95% CI)	
Vaccine	0.016 (0.014, 0.018)	<0.001	0.29 (0.26, 0.32)	<0.001	0.32 (0.31, 0.33)	<0.001
Age	1.06 (1.05, 1.07)	<0.001	1.33 (1.32, 1.34)	<0.001		
Gender	0.97 (0.94, 0.99)	<0.001	1.01 (1.00, 1.02)	<0.001		
Occupation	1.16 (1.15, 1.17)	<0.001	1.01 (0.99, 1.00)	<0.001		
Education	0.56 (0.54, 0.57)	<0.001	0.84 (0.83, 0.85)	<0.001		
Marital Status	1.40 (1.36, 1.44)	<0.001	1.02 (1.01, 1.04)	<0.001		
Place of Residence	1.56 (1.51, 1.62)	<0.001	1.02 (1.02, 1.03)	<0.001		
Body Mass Index	1.022 (1.020, 1.023)	<0.001	1.00 (0.99, 1.01)	<0.001		
Number of Comorbidities	4.48 (4.40, 4.56)	<0.001	1.78 (1.76, 1.80)	<0.001		
Number of Clinical Symptoms	18.42 (18.12, 18.72)	<0.001	11.45 (11.18, 11.71)	<0.001		
History of COVID-19 infection	2.00 (1.97, 2.02)	<0.001	0.02 (0.01, 0.03)	<0.001		

body mass Index, Number of Comorbidities, Number of Clinical Symptoms and History of COVID-19 infection.

The Estimation of the vaccination effect by PSM analysis also showed that vaccinated patients had significantly lesser odds of hospitalization compared to unvaccinated patients (OR, 95% CI using PSM: 0.32 (0.31, 0.33), $p < 0.001$) as determined when the propensity scores were estimated with logistic regression. See Table 3.

Discussion

The findings of our study reveal a significant reduction in hospitalization among COVID-19 patients who have been vaccinated. Utilizing propensity score matching (PSM) enabled us to estimate the effect of vaccination with greater precision, exhibiting a lower standard error and a narrower confidence interval compared to traditional multivariable logistic regression (MLR) techniques.

Our analysis indicates that vaccinated patients tend to present with less clinical symptoms, as evidenced by a higher mean age, and a lower number of comorbidities.

This trend is corroborated by similar studies conducted in various regions. For instance, the population-based study from Israel showed two doses of the BNT162b2 vaccine reduced symptomatic as well as asymptomatic COVID-19 infections, and breakthrough infection was less severe with reduced hospitalisation and lower mortality.[5]

Another study in Germany reported that vaccination protected against severe disease for at least six months, with a vaccine effectiveness (VE) of 90% for two doses and 99% for three doses. The VE was significantly lower for adults with three or more pre-existing comorbidities compared to those with fewer comorbidities, but this reduction was compensated when a third dose was administered.[6]

Moreover, a study in Canada found that two-dose vaccine effectiveness against SARS-CoV-2 hospitalization was 93% during the Delta variant's dominance. However, the effectiveness was lower at 40% among adolescents during the circulation of the Delta variant, possibly due to vaccine waning and earlier vaccination dates in the United States compared to Ontario.[7]

A multistate analysis of over 34,000 hospitalizations for COVID-19–like illness among adults with immunocompromising conditions found that 2 doses of monovalent mRNA COVID-19 vaccine were 36% effective against COVID-19–associated hospitalization during a period of Omicron predominance.[8]

While our study underscores the efficacy of vaccination in reducing hospitalization, several limitations necessitate consideration. Most notably, the absence of data on anti-spike antibody titers, oxygen levels, and specific vaccine types hindered a comprehensive analysis of their impact on hospitalization. Additionally, the variability in vaccine effectiveness, timing of vaccination, and host immune responses underscore the complexity of managing breakthrough infections. A significant limitation of our study is the lack of consideration for the specific type of COVID-19 variant affecting patients. Different variants may have varying levels of virulence and vaccine resistance, which could influence hospitalization rates. Future studies should aim to include variant data to provide a more nuanced understanding of vaccine effectiveness against specific strains.

Furthermore, it's important to acknowledge the potential for unmeasured confounding factors in our analysis. While we accounted for several important demographic and clinical variables, there may be other unmeasured factors that could have influenced the observed outcomes. These might include socioeconomic status, healthcare access, adherence to public health measures, or genetic factors that could affect COVID-19 susceptibility or severity.

Conclusion

Our findings underscore the pivotal role of vaccination in mitigating COVID-19 hospitalization. Propensity score matching offers a robust methodological approach to assessing the impact of vaccination, highlighting its importance in pandemic control efforts. Future research endeavors should focus on vaccine characteristics, specific COVID-19 variants, and potential unmeasured confounding factors to optimize public health strategies in combating COVID-19. Additionally, efforts should be made to collect and analyze data on specific vaccine types and their effectiveness against different variants to provide more targeted vaccination strategies.

Abbreviations List

OR: Odds ratio;
PS: Propensity score;
PSM: Propensity score matching;
ATE: Average treatment effect;
MLR: Multivariate logistic regression;
VE: Vaccine effectiveness;
ASMD: Absolute standardized mean difference;

Conflicts of interests

The authors have no conflicts of interest to declare.

Acknowledgements

This article is the result of a research project approved by MUMS, Iran, under the ethics code IR.MUMS.FHMPM.REC.1403.097 and grant number 4030664. We extend our gratitude to the Health Deputy of MUMS for their invaluable assistance in data collection.

References

1. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate behavioral research. 2011;46(3):399-424.

2. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70(1):41-55.

3. Elze MC, Gregson J, Baber U, Williamson E, Sartori S, Mehran R, et al. Comparison of propensity score methods and covariate adjustment: evaluation in 4 cardiovascular studies. Journal of the American College of Cardiology. 2017;69(3):345-57.

4. Austin PC, Stuart EA. Estimating the effect of treatment on binary outcomes using full matching on the propensity score. Statistical methods in medical research. 2017;26(6):2505-25.

5. Haas EJ, Angulo FJ, McLaughlin JM, Anis E, Singer SR, Khan F, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. The Lancet. 2021;397(10287):1819-29.

6. Stoliaroff-Pepin A, Peine C, Herath T, Lachmann J, Perriat D, Dörre A, et al. Effectiveness of vaccines in preventing hospitalization due to COVID-19: A multicenter hospital-based case-control study, Germany, June 2021 to January 2022. Vaccine. 2023;41(2):290-3.

7. Simmons AE, Amoako A, Grima AA, Murison KR, Buchan SA, Fisman DN, et al. Vaccine effectiveness against hospitalization among adolescent and pediatric SARS-CoV-2 cases between May 2021 and January 2022 in Ontario, Canada: a retrospective cohort study. PLoS One. 2023;18(3):e0283715.

8. Britton A. Effectiveness of COVID-19 mRNA vaccines against COVID-19–associated

hospitalizations among immunocompromised adults during SARS-CoV-2 Omicron predominance—VISION Network, 10 states, December 2021—August 2022. MMWR Morbidity and Mortality Weekly Report. 2022;71.