

## Research Paper

## Evaluation of the Causal Association of Risk Factors for Death From COVID-19 Patients Admitted to Golestan Hospitals by Propensity Score Estimation Method

Hassan Khorsha<sup>1</sup>, Manoochehr Babanezhad<sup>1\*</sup>, Naser Behnampour<sup>2</sup>

1. Department of Statistics, Faculty of Sciences, Golestan University, Gorgan, Iran.

2. Department of Biostatistics and Epidemiology, School of Health, Golestan University of Medical Sciences, Gorgan, Iran.



**Citation** Khorsha H, Babanezhad M, Behnampour N. Evaluation of the Causal Association of Risk Factors for Death From COVID-19 Patients Admitted to Golestan Hospitals by Propensity Score Estimation Method. *Iranian Journal of Health Sciences*. 2023; 11(4):279-288. <http://dx.doi.org/10.32598/ijhs.11.4.567.3>

<http://dx.doi.org/10.32598/ijhs.11.4.567.3>



## ABSTRACT

**Background and Purpose:** Evaluating the causal association effect of risk factors is essential in estimating the mortality rate among COVID-19 patients. Much research has been conducted to assess the impact of COVID-19 on death in various countries worldwide. However, few studies have addressed the effect of causal association of risk factors. This study aims to address this gap by estimating the impact of COVID-19 on death by evaluating the causal association with the risk factors.

**Materials and Methods:** The research population included all inpatients with initial COVID-19 symptoms, confirmed by their PCR test results. They were admitted to hospitals affiliated with Golestan University of Medical Science, Golestan, Iran, in 2020. We employed the propensity score method, an effective statistical technique for evaluating the causal association effect of risk factors in observational studies. We also used the student and chi-squared tests to compare the differences between the two study groups.

**Results:** We used the propensity score matching estimation approach and logistic regression analysis for comparison. Of 6379 inpatients, 5581 (87.5%) were discharged or recovered, and 798 (12.5%) died. The causal association between treatment results (discharged vs died) and variables of PCR test, SpO<sub>2</sub>, gender, age, and hospitalization duration in ICU were statistically significant.

**Conclusion:** The propensity score matching estimation method revealed a high risk of death in patients with PCR+ test diagnosis. Specifically, using this approach, the above-measured risk factors increased the chance of death in patients with PCR+ to 72%. However, the traditional multiple logistic regression model estimated the risk of death at 46%, suggesting potential underestimation. This disparity might be due to better control of the effect of the above-measured risk factors by the propensity score matching. Therefore, the former estimating approach is more effective in assessing the impact of COVID-19 on death.

**Keywords:** COVID-19, Risk factors, Causal association, Propensity score, Propensity score matching, logistic regression model

### Article info:

**Received:** 15 Jul 2023

**Accepted:** 23 Sep 2023

**Available Online:** 01 Oct 2023

### \* Corresponding Author:

**Manoochehr Babanezhad, Associated Professor.**

**Address:** Department of Statistics, Faculty of Sciences, Golestan University, Gorgan, Iran.

**Tel:** +98 (17) 32245882

**E-mail:** [m.babanezhad@gu.ac.ir](mailto:m.babanezhad@gu.ac.ir)

## 1. Introduction

In December 2019, an emerging disease called COVID-19 originated in Wuhan, China, and rapidly spread worldwide. The resulting global pandemic impacted all continents [1, 2]. In Iran, the first case of COVID-19 was reported on February 19, 2020, in Qom city [3, 4]. Various diagnostic methods have been used to detect COVID-19, including laryngoscopy, initial blood tests, and polymerase chain reaction (PCR) tests [5-7]. Both the World Health Organization (WHO) and the Infectious Disease Controlling Center recommended PCR as an accurate way to diagnose the disease [8, 9].

When trying to determine the mortality rate of COVID-19 in a scenario where a higher number of hospitalized patients are likely to succumb to the disease, it becomes imperative to assess the relationship between risk factors and deaths attributed to COVID-19 [10, 11]. Epidemiological studies have indicated that certain factors will likely influence mortality risk in patients with a positive PCR test diagnosis [5, 12-14]. For example, Li et al. reported that age over 60 was one of the leading mortality factors in patients with PCR<sup>+</sup> test diagnosis [5]. Older age weakens the body's immune system, acting as a barrier against viral infections. Also, background diseases such as blood pressure, diabetes, and cardiovascular diseases increase the risk factors for illness and mortality from COVID-19 in older patients.

However, traditional statistical models like logistic regression may not adequately account for the impact of factors such as age and underlying medical conditions on COVID-19-related deaths. Failing to account for these factors in statistical models can lead to biased estimates of the association between COVID-19 and mortality [13-15]. The potential bias in estimating the association between COVID-19 and mortality could arise from improperly matching risk factors among various groups of COVID-19 patients. This discrepancy in risk factors between these groups may introduce an association between COVID-19 and mortality influenced by the variations in these risk factors [16-19].

Several estimation methods have been recently developed to enhance our ability to generate unbiased estimators from data derived from improperly matched risk factors among different COVID-19 patient groups [18-22]. The propensity score estimation (PSM) method has been introduced in certain studies as a novel approach for matching risk factors [14, 18]. This method effectively balances the distribution of risk factors among various

groups of COVID-19 patients. The propensity score (PS) is defined as the conditional probability of an individual  $i$  with a positive PCR (PCR<sup>+</sup>) in a specific condition, given the risk factors for that individual. Therefore, according to the Equation 1, the likelihood of:

$$1. \hat{P} = \Pr(Z_i = 1 | X_i = x_i)$$

for the individual  $i$  with PCR<sup>+</sup>,  $Z_i=1$ , and for those with PCR<sup>-</sup>,  $Z_i=0$ , where  $X_i$  denotes the vector of risk factors for the individual  $i$  are measured before measuring PCR, or their scores are not affected by the PCR results. In administrating the PSM in practice, the logistic regression model and matching approaches are used [15, 16]. In the PSM approach, the main variables observed between the treated (PCR<sup>+</sup>,  $Z_i=1$ ) and control (PCR<sup>-</sup>,  $Z_i=0$ ) groups are balanced on Equation 2:

$$2. d = \frac{(\hat{P}_{treatment} - \hat{P}_{control})}{\sqrt{\frac{\hat{P}_{treatment}(1 - \hat{P}_{treatment}) + \hat{P}_{control}(1 - \hat{P}_{control})}{2}}}$$

This approach, named nearest neighbor matching (NNM), or PSM, is a quasi-experimental method where the researcher uses statistical techniques to construct an artificial control group by matching each treated unit with a control unit of similar characteristics. The PS is used for estimating the results before and after matching. The outliers can be shown by the standard mean difference reflecting the balance between the treated and control groups [23, 24].

This study aims to assess COVID-19 mortality by examining the causal association of risk factors among COVID-19 patients admitted to hospitals in Golestan Province, Iran. The data and statistics presented in this article were obtained from the Hospital Information System (HIS) software installed in hospitals affiliated with Golestan University of Medical Sciences. Among the 30114 hospitalized patients suspected of having the coronavirus, PCR tests were conducted on 6379 patients. It should be noted that during the initial months of the pandemic, testing availability was limited, and not all patients underwent PCR testing.

The study focuses on risk factors such as SpO<sub>2</sub> levels, gender, age, and duration of ICU hospitalization. We investigated their influence on the risk of death in COVID-19 patients with a positive PCR diagnosis. Initially, logistic regression analysis was used to estimate the effect of COVID-19 on the death rate among patients. However, it was observed that the underlying risk fac-

tors were not adequately adjusted for. To address this, PS matched deceased patients with recovered patients with similar scores. This matching approach was mainly utilized when a higher number of hospitalized patients were expected to die. The study also compared the results obtained from logistic regression analysis and the PSM estimation method.

## 2. Materials and Methods

### Study design and sample size

The current study participants were 6379 patients with COVID-19 symptoms and their PCR test results. They were referred to hospitals with different service qualities in Golestan Province, north-east of Iran, in 2020. Golestan Province has 26 hospitals. The main hospitals for admitting COVID-19 patients are Sayyad Shirazi Gorgan and Payambar Azam Gonbad hospitals. It can be stated that the other hospitals admitted patients in emergencies. All hospitals were under the supervision of [Golestan University of Medical Science](#). Note that not all patients were tested with the PCR test from the beginning of the COVID-19 outbreak in Golestan Province until the end of 2019. However, a total of 30114 people were hospitalized in all the hospitals of this province.

### Inclusion criteria

The inclusion criteria were patients diagnosed and registered as COVID-19 in the study hospitals, and their PCR tests were positive.

### Exclusion criteria

The exclusion criteria were the patients discharged from the hospital with partial recovery or died.

### Questioners and data collection

The hospital admission registers the basic information of each patient at admission. Using the hospital HIS software, the collected information was transferred to the central data of the University of Medical Sciences. Then, another application used in the laboratory of the provincial health center recorded the results of the PCR test for coronavirus patients. Next, the patient's information was processed based on the patient's profile with a specific code. This information can be identified and tracked.

Their gender, age, SpO<sub>2</sub> (pulse oximeter), hospitalization, and ICU (intensive care unit) stay duration were recorded by referring to the hospital HIS. The results of

the PCR test and treatment as died or discharged were conceived as independent and dependent variables, respectively.

### Statistical data analysis

This study used the logistic regression analysis as an initial step to estimate the effect of COVID-19 on the rate of death among patients. After the initial logistic regression analysis, PS were used to adjust for underlying risk factors such as age, gender, and SpO<sub>2</sub> levels. PS were used to match deceased patients with recovered patients who had similar PS. The approach helped to create more comparable groups and reduce bias in estimating the effect of COVID-19 on the rate of death among patients. PS estimation was applied to estimate the probability scores of the underlying risk factors with probability in expression 1. PSM method relies on the distance  $d$  in expression 2. The value  $d$  was calculated for estimating results before and after matching. Then, the means of the treatment group and the control group were assimilated. We also used the traditional multiple logistic regression model to compare with the PSM method results. R software, version 4.3.2 and the SPSS software, version 22 were used for statistical analysis.

## 3. Results

Of 6379 inpatients, 5581(87.5%) were discharged/recovered, and 798(12.5%) died. Moreover, 1954 patients (30.6%) had positive and 4425(69.4%) negative PCR test results. Further, their average age was  $53.91 \pm 23.61$  years; 3130(49.1%) were male, and 3249(50.9%) were female. [Table 1](#) presents the characteristics of the patients under the study based on their treatment results (discharged vs died).

The chi-square test showed a significant association between the results of treating patients and that of PCR results and ICU stay. Moreover, the t-test showed that the average rate of age among dead ones was significantly higher than that of recovered ones, and the mean rate of SpO<sub>2</sub> in dead ones was considerably lower than that of recovered ones ( $P < 0.001$ ).

[Table 2](#) presents the causal association between the underlying covariates and treatment results (discharged vs dead) for 6379 subjects before matching and 3960 after matching. Results showed a considerable change in the percentage of deaths of patients with PCR<sup>+</sup> after matching. [Table 2](#) also indicates that PCR<sup>+</sup> between dead has increased from 38% to 59.1% before and after matching. Moreover, male deaths have changed from

**Table 1.** Descriptive statistics on the underlying covariates studied in COVID-19 inpatients by the treatment result

Variables		No. (%) / Mean $\pm$ SD		
		Total	Discharged	Deceased
PCR	Positive	1954(30.6)	1651(29.6)	303(38.0)
	Negative	4425(69.4)	3930(70.4)	495(62.0)
Gender	Male	3130(49.1)	2697(48.3)	433(54.3)
	Female	3249(50.9)	2884(51.7)	365(45.7)
ICU	Hospitalized	1295(20.3)	822(14.7)	473(59.3)
	Non-hospitalized	5084(79.7)	4759(85.3)	325(40.7)
Age (y)		53.91 $\pm$ 23.61	52.02 $\pm$ 23.9	67.12 $\pm$ 16.27
SpO <sub>2</sub>		91.49 $\pm$ 9.33	92.48 $\pm$ 8.2	84.6 $\pm$ 12.88

Abbreviations: PCR: Polymerase chain reaction; ICU: Intensive care unit; SpO<sub>2</sub>: Oxygen saturation.

54% to 56%, and females from 45% to 43% after matching. This finding implies that the number of deaths due to COVID-19 in men is higher than that in women. In contrast, there were no significant differences between the dead groups before and after matching regarding their average age and SpO<sub>2</sub>.

The results of binary regression goodness-of-fit showed that before matching, PCR<sup>+</sup>, low SpO<sub>2</sub>, older

age, male gender, and hospitalization in ICU resulted in an increased risk of death among the patients (Table 3).

After matching, binary logistic regression results revealed a positively increased score based on the PS approach and an increased death ratio among patients with PCR<sup>+</sup>. Matching also reflected the influence of hospitalization in the ICU (Table 4).

**Table 2.** Comparing measures of central and distribution tendencies of variables before and after matching

Basic Variables		Before			After		
		Dead	Discharged	P	Dead	Discharged	P
PCR	+	38.0	29.6	<0.001	59.1	48.6	<0.001
	-	62.0	70.4		40.9	51.4	
Gender	Male	54.3	48.3	<0.002	56.1	47.5	<0.001
	Female	45.7	51.7		43.9	52.5	
ICU	Yes	59.3	14.7	<0.001	61.6	13.1	<0.001
	NO	40.7	85.3		38.4	86.9	

Basic Variables		Mean $\pm$ SD					
		Before			After		
		Dead	Discharged	P	Dead	Discharged	P
Age		67.1 $\pm$ 16.2	52.0 $\pm$ 23.9	<0.001	67.8 $\pm$ 14.82	56.92 $\pm$ 18.25	<0.001
SpO <sub>2</sub>		84.6 $\pm$ 12.8	92.4 $\pm$ 8.2	<0.001	83.92 $\pm$ 13.02	92.05 $\pm$ 8.14	<0.001

Abbreviations: PCR: Polymerase chain reaction; ICU: Intensive care unit; SpO<sub>2</sub>: Oxygen saturation.

**Table 3.** The results of the multiple logistic regression analysis before matching data

Variables	Coefficient	Exp (Coefficient)	95% CI	P
Constant	-0.654	0.520		
PCR (+)	0.377	1.459	(1.224-1.737)	<0.001
SpO <sub>2</sub>	-0.044	0.957	(0.950 - 0.964)	<0.001
Age	0.029	1.029	(1.024-1.034)	<0.001
Gender (male)	0.226	1.254	(1.060-1.483)	<0.001
ICU (+)	1.891	6.625	(5.599-7.840)	<0.001

Abbreviations: PCR: Polymerase chain reaction; ICU: Intensive care unit; SpO<sub>2</sub>: Oxygen saturation; Exp (coefficient): Exponential function calculator to find its value.

**Table 4.** The results of the multiple logistic regression analysis after matching data

Variables	Coefficient	Exp (Coefficient)	95% CI	P
Constant	-1.028	0.358		
PCR (+)	0.544	1.724	(1.386-2.143)	<0.001
SpO <sub>2</sub>	-0.046	0.955	(1.026-1.041)	<0.001
Age	0.033	1.033	(1.026-1.041)	<0.001
Gender (male)	0.315	1.371	(1.105-1.700)	<0.001
ICU (+)	2.081	8.010	(6.451-9.944)	<0.001

Abbreviations: PCR: Polymerase chain reaction; ICU: Intensive care unit; SpO<sub>2</sub>: Oxygen saturation; Exp (coefficient): Exponential function calculator to find its value.

According to [Table 5](#), before matching, the estimation of the mean effect of propensity (distance) in the PCR<sup>+</sup> group and PCR<sup>-</sup> group reached 0.3161 and 0.3020, respectively. After matching, the indicator reached 0.3161 in the PCR<sup>+</sup> group and 0.3160 in the PCR<sup>-</sup> group, reflecting an unchanged amount in the former. However, the standardized mean difference before matching (0.2968) and after matching (0.0007) reflected a considerable decrease. The variance ratio before matching (0.6148) and after matching (1.0011) reflected more homogeneity of variances in the groups after matching.

[Table 5](#) summarizes the PSM results for the death rate of COVID-19 patients based on the hospitalization of having a PCR test. It indicates that the patient's initial diagnosis at the admission was based on the clinical symptoms, and then the desired test was performed. In the initial data, there were 6379 patients. [Figure 1](#) illustrates that matching considerably affected the groups' homogeneity in other studied variables. Outlier observations in [Figure 1](#) present estimates of the probability of receiving treatment close to 0 or 1; therefore, the

PSM method with limited overlap can generate estimates approximately unchanged in bias and precision.

#### 4. Discussion

This study aimed to examine the impact of various risk factors on the mortality of COVID-19 patients treated in Golestan hospitals with varying healthcare capabilities. The measured risk factors in this study included gender, age, SpO<sub>2</sub> (pulse oximeter), hospitalization, and duration of ICU stay. These factors were documented by referencing the HIS.

Usually, logistic regression procedures are employed in routine practice to address such analyses; however, they come with inherent flaws. For example, the association of certain risk factors, such as age, may confound the effect of COVID-19 on mortality. Younger patients with shorter hospitalization and ICU stays exhibit a lower mortality risk than older patients with longer stays.

**Table 5.** The results of the propensity scores before and after matching

Variables	Summary of Balance for All Data					
	Means Treated	Means Control	Std. Mean Diff	Var. Ratio	eCDF Mean	eCDF Max
Distance	0.3161	0.3020	0.2968	0.6148	0.0609	0.1531
SpO <sub>2</sub>	90.5218	91.9304	-0.1494	1.0349	0.0201	0.1613
Age	57.8132	52.1835	0.3128	0.4972	0.0615	0.1379
Gender	1.5118	1.5082	0.0070	1.0000	0.0018	0.0035
ICU	1.8050	1.7934	0.0292	0.9580	0.0058	0.0116

Variables	Summary of Balance for Matched Data					
	Means Treated	Means Control	Std. Mean Diff	Var. Ratio	eCDF Mean	eCDF Max
Distance	0.3161	0.3160	0.0007	1.0011	0.0003	0.0036
SpO <sub>2</sub>	90.5218	91.5230	0.1062-	1.0385	0.0146	0.1100
Age	57.8132	58.8971	0.0602-	0.9541	0.157	0.0512
Gender	1.5118	1.5154	0.0072-	1.0004	0.0018	0.0036
ICU	1.8050	1.8050	0.0000	1.0000	0.0000	0.0000

Variables	Percent Balance Improvement					
	Std. Pair Dist.	Std. Mean Diff	Var. Ratio	eCDF Mean	eCDF Max	
Distance	0.0014	99.8	99.8	99.6	97.7	
Spo <sub>2</sub>	0.5442	28.9	10.1-	27.2	31.8	
Age	0.3909	80.7	93.3	74.4	62.9	
Gender	0.6889	-1.7	10334.5-	1.7-	1.7-	
ICU	0.2446	100.0	100.0	100.0	100.0	

Variables	Sample Size	
	Control	Treated
All	4425	1954
Matched	1954	1954
Unmatched	2471	0
Discarded	0	0

Abbreviations: PCR: Polymerase chain reaction; ICU: Intensive care unit; SpO<sub>2</sub>: Oxygen saturation.

Consequently, researchers during the COVID-19 pandemic aimed to determine whether certain risk factors, such as age and gender, impact COVID-19-related deaths. These risk factors can introduce bias in estimat-

ing the mortality rate of COVID-19, especially in individuals with underlying diseases. While several studies have been conducted in our country to evaluate the influence of risk factors on COVID-19-related mortality

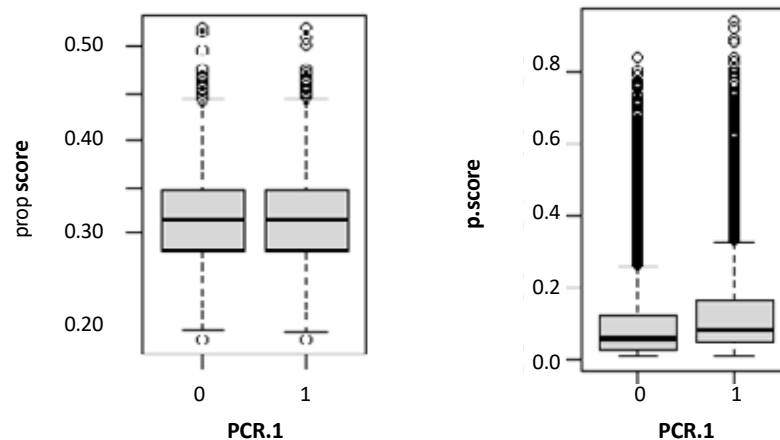


Figure 1. Box plot of mean propensity score (distance) before (in left) and after (in right) matching

outcomes, they may not comprehensively account for the role and effect of all potential risk factors for COVID-19-related deaths [1-3].

In our study, we observed that the risk of mortality increased in all age groups after the age of 53, with a significant rise in individuals aged 65 and older. However, the mortality risk among elderly patients was comparable to that of the general population, suggesting that although age is a significant risk factor for COVID-19-related mortality outcomes, other factors may also contribute.

We also observed that traditional multiple logistic regression methods fail to estimate unbiased mortality rates. This finding underscores the need for more advanced statistical methods to determine the causal association between COVID-19 and mortality rate accurately. Specifically, the traditional logistic regression output in Table 3 indicates interference with the data when analyzing the effect of PCR<sup>+</sup> on death. In contrast, the results in Table 4 demonstrate the advantages of using new data over the previous data, as evidenced by an increase in Means Treated and a decrease in Std. Diff (Table 5). According to Table 5, after removing the excess data, the mean values of the treatment and control groups in the assimilated data became closer. Consequently, the standard average difference reduced from 0.2968 to 0.0007. Moreover, the variance ratio increased from 0.6148 to 1.0011. These indicators specify better balance in the treatment and control data groups.

Furthermore, the average standard deviation tends towards 0, and a variance ratio of 1 suggests improved matching for the PCR<sup>+</sup> group. This improvement and reduction were achieved by equalizing the standard deviation between the control and treatment groups, a

task that logistic regression fails to reduce this standard deviation. Additionally, the box plot diagram below illustrates the index or average tendency score. According to our findings, the measured risk factors mentioned earlier resulted in a 72% increase in mortality risk among patients with a positive PCR result when employing the PSM estimation approach. However, utilizing the multiple logistic regression model approach yielded an observed increase in mortality risk of 46%. This outcome suggests that the latter approach is more effective in controlling measured risk factors.

Our findings indicated that the age at which individuals expired due to COVID-19 was significantly higher than those who recovered from it. This finding could be attributed to various physiological and anatomical changes in the human body associated with aging. These age-related changes render older individuals more susceptible to severe outcomes and complications related to COVID-19, leading to a higher mortality rate among this population. This finding is consistent with the findings of Sheikhi et al. [2]. In our study, we also observed that patients whose PCR test was positive had a higher risk of mortality than those with negative test results. These findings align with the previous research conducted by Azizmohammad Looha et al. [3], further supporting the association between a positive PCR test and an increased risk of mortality from COVID-19.

In our study, older age, male gender, and dyspnea were risk factors. Li et al. concluded that after matching, the effect of these factors had been well-reflected. Moreover, they found that without matching, the rate of mortality can increase in old male patients [5]. We also found such a finding for male gender and older patients. The higher risk of death observed in men could

be attributed to lifestyle factors such as higher rates of smoking and alcohol consumption.

The study conducted by Kim et al. demonstrated an increase in the mortality rate from COVID-19 [6]. Before matching, the difference in mortality rate was not significant. According to our observations in Table 5, females exhibited a significantly lower risk of death than males. These findings align with the previous research conducted by Her et al. in South Korea, as mentioned in some studies [7].

However, Elze et al. argued that utilizing moderating covariates and matching through the PS approach yielded better and more accurate results [10]. Our findings align with these studies, further supporting the notion that matching by PS and considering moderating covariates can enhance the accuracy of mortality rate analysis.

Our findings revealed that the SpO<sub>2</sub> levels of deceased patients were significantly lower than those who were discharged. This outcome may be due to the difference in SpO<sub>2</sub> levels between deceased patients and those discharged, which could reflect the severity of COVID-19 infection. Patients with more severe symptoms and complications tend to have lower SpO<sub>2</sub> readings due to impaired lung function caused by inflammation or fluid accumulation.

Additionally, the risk of death from the disease was found to be higher among patients hospitalized in the ICU compared to those who were not. The higher mortality may be because patients requiring ICU hospitalization are often those with more severe symptoms and complications related to COVID-19. These individuals may have compromised respiratory function, organ failure, or other critical conditions that increase their chance of mortality. These results are consistent with a study by Martinez-Martinez et al. [11].

In this study, after matching risk factors by applying the PS approach, the effect of PCR<sup>+</sup> increased significantly. This finding agrees with findings in other studies [8]. One advantage of this study is its use of the PSM approach, which has rarely been used in previous studies. This method correctly models the nature of the relationship between the PS and the outcome.

However, the PSM approach is yielding an amended estimation. Studies also reveal that methods such as G-estimation can, in principle, be adopted to show the impact of the causal association of the risk factors.

## 5. Conclusion

Using the PSM estimation method showed the high risk of death in patients with the PCR<sup>+</sup> test. Specifically, using the PSM estimation approach, the above-measured confounding factors increased the risk of death in patients with PCR<sup>+</sup> to 72%. However, the multiple logistic regression model revealed the risk of death at 46%. This discrepancy might be due to better control of the impact of above-measured risk factors. Therefore, the former estimating approach is more effective in controlling the impact of confounding factors.

### Study limitations

The study encountered certain limitations. The first one was related to the hospitals under the supervision of Golestan University of Medical Science. They had different levels of quality in services. The second limitation was that the exact time of being infected with COVID-19 was the inability to adjust for the exact time of COVID-19 infection precisely. These factors may limit the interpretation of our results.

### Ethical Considerations

#### Compliance with ethical guidelines

This work was approved by the Ethics Committee of Medical Research in Iran (Code: IR.GOUMS.REC.1401.185). Although the proposal has been approved by the Biomedical Research Ethics Committee, meeting the professional and legal requirements is the sole responsibility of the PI and other project collaborators.

### Funding

The paper was extracted from the first author's Master thesis under supervisor of the second author, Department of Statistics, Faculty of Sciences, Golestan University, Gorgan, Iran.

### Authors contributions

Conceptualization: Hassan Khorsha; Initial idea, methodology and writing original draft: Manoochehr Babanezhad; Editing: Naser Behnampour; Investigation: All authors.

### Conflict of interest

The authors declared no conflict of interest.



## Acknowledgments

The authors would like to thank the Golestan University authorities.

## References

- [1] Hadinejad Z, Saleh Tabari Y, Sajadi Z, Talebi Ghadicolaei H. Clinical characteristics and risk factors of COVID-19 and related deaths in elderly patients. *Salmamd*. 2021; 16(1):128-39. [DOI:10.32598/sija.16.1.658.2]
- [2] Sheikhi F, Mirkazehi Rigi Z, Azarkish F, Kalkali S, Seid Abadi M, Mirbaloochzahi A. [Clinical and demographic characteristics of patients with COVID-19 in Iranshahr hospitals, southeastern Iran in 2020 (Persian)]. *Journal of Marine Medicine*. 2021; 3(1):46-52. [DOI:10.30491/3.1.46]
- [3] Azizmohammad Looha M, Rezaei-Tavirani M, Rostami-Nejad M, Janbazi S, Zarean E, Amini P, et al. Assessing sex differential in COVID-19 mortality rate by age and polymerase chain reaction test results: an Iranian multi-center study. *Expert Review of Anti-Infective Therapy*. 2022; 20(4):631-41. [DOI:10.1080/14787210.2022.2000860] [PMID] [PMCID]
- [4] Benksim A, Ait Addi R, Cherkaoui M. Vulnerability and fragility expose older adults to the potential dangers of COVID-19 pandemic. *Iranian Journal of Public Health*. 2020; 49(Suppl 1):122. [DOI:10.18502/ijph.v49iS1.3682]
- [5] Li G, Liu Y, Jing X, Wang Y, Miao M, Tao L, et al. Mortality risk of COVID-19 in elderly males with comorbidities: A multi-country study. *Aging*. 2020; 13(1):27-60. [DOI:10.18632/aging.202456] [PMID] [PMCID]
- [6] Kim S, Choi H, Sim JK, Jung WJ, Lee YS, Kim JH. Comparison of clinical characteristics and hospital mortality in critically ill patients without COVID-19 before and during the COVID-19 pandemic: A multicenter, retrospective, propensity score-matched study. *Annals of Intensive Care*. 2022; 12(1):57. [DOI:10.1186/s13613-022-01028-2] [PMID] [PMCID]
- [7] Her AY, Bhak Y, Jun EJ, Yuan SL, Garg S, Lee S, et al. Sex-specific difference of in-hospital mortality from COVID-19 in South Korea. *Plos One*. 2022; 24;17(1):e0262861. [DOI:10.1371/journal.pone.0262861] [PMID] [PMCID]
- [8] Parmar H, Montovano M, Banada P, Pentakota SR, Shiau S, Ma Z, et al. RT-PCR negative COVID-19. *BMC Infectious Diseases*. 2022; 22(1):149. [DOI:10.1186/s12879-022-07095-x] [PMID] [PMCID]
- [9] Asghari E, Rahmani F, Hosseinzadeh M, Mahdavi N, Praskova A, Targhaq S, et al. the clinical characteristics of early cases with 2019 novel coronavirus disease in Tabriz, Iran. *Iranian Journal of Health Science*. 2021; 9(1):1-11. [DOI:10.18502/jhs.v9i1.5968]
- [10] 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. [DOI:10.1016/j.jacc.2016.10.060] [PMID]
- [11] Martínez-Martínez M, Plata-Menchaca EP, Nuvials FX, Roca O, Ferrer R. Risk factors and outcomes of ventilator-associated pneumonia in COVID-19 patients: A propensity score matched analysis. *Critical Care*. 2021; 25(1):235. [DOI:10.1186/s13054-021-03654-x] [PMID] [PMCID]
- [12] Austin PC. A comparison of 12 algorithms for matching on the propensity score. *Statistics in Medicine*. 2014; 33(6):1057-69. [DOI:10.1002/sim.6004] [PMID] [PMCID]
- [13] Shiba K, Kawahara T. Using propensity scores for causal inference: Pitfalls and tips. *Journal of Epidemiology*. 2021; 31(8):457-463. [DOI:10.2188/jea.JE20210145] [PMID] [PMCID]
- [14] Jalali N, Babanezhad M, Yaghmaie F. Impact of measured confounders in estimating the causal effect of an exposure: A monte carlo study. Paper presented at: 2th Workshop on Reliability and its Applications. 9 June 2011; Tehran, Iran. [Link]
- [15] Ganji A, Mosayebi Gh, Khaki M, Ghazavi A. [a review on immunopathogenesis, molecular biology and clinical aspects of the 2019 novel coronavirus (COVID-19) (Persian)]. *Journal of Arak University of Medical Sciences*. 2020; 23(1):8-21. [DOI:10.32598/JAMS.23.1.51.5]
- [16] Janani L, Moradian S, Ahmadi H, Mohammad K, Golestan B, Forouzanfar M. Propensity score analysis for comparing primary vitrectomy with and without encircling band in Pseudophakic/Aphakic retinal detachment. 2011; 16(2):100-7. [Link]
- [17] 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. [DOI:10.1080/0273171.2011.568786] [PMID] [PMCID]
- [18] Vansteelandt S, Daniel RM. On regression adjustment for the propensity score. *Statistics in Medicine*. 2014; 33(23):4053-72. [DOI:10.1002/sim.6207] [PMID]
- [19] Justus J R, Falbe K, Manuel AK, Balloun JL. A step-by-step guide to propensity score matching in R. *Practical Assessment Research and Evaluation*. 2014; 19(18):1-6. [DOI:10.7275/n3pv-tx27]
- [20] Dukes O, Vansteelandt S. How to obtain valid tests and confidence intervals after propensity score variable selection? *Statistical Methods in Medical Research*. 2020; 29(3):677-694. [DOI:10.1177/0962280219862005] [PMID]
- [21] Morgan CJ. Reducing bias using propensity score matching. *Journal of Nuclear Cardiology*. 2018; 25(2):404-6. [DOI:10.1007/s12350-017-1012-y] [PMID]
- [22] Zhao S, van Dyk DA, Imai K. Propensity score-based methods for causal inference in observational studies with non-binary treatments. *Statistical Methods in Medical Research*. 2020; 29(3):709-27. [DOI:10.1177/0962280219888745] [PMID]
- [23] Taquet M, Dercon Q, Luciano S, Geddes JR, Husain M, Harrison PJ. Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. *Plos Medicine*. 2021; 18(9):e1003773. [DOI:10.1371/journal.pmed.1003773] [PMID] [PMCID]
- [24] Nguyen TL, Collins GS, Spence J, Daurès JP, Devereaux PJ, Landais P, et al. Double-adjustment in propensity score matching analysis: Choosing a threshold for considering residual imbalance. *BMC Medical Research Methodology*. 2017; 17(1):78. [DOI:10.1186/s12874-017-0338-0] [PMID] [PMCID]

This Page Intentionally Left Blank

---