Assignment

Martijn Koster, 6234119 Jurrian van de Kraats, student nr
Tim Poorthuis, student nr

2023-02-15

Introduction

In this paper the

In this paper we will examine whether causal inference can be drawn about the association elderly people (i.e., 65+) who have received an influenza vaccine and the likelihood of hospitalization.

Write here a short motivation with a RQ

- well articulated RQ

The aim of the study is to assess whether annual influenza vaccination reduces the risk of hospitalization among elderly (i.e., people aged >=65 years).

Methods

Selection of confounders

Based on the Directed Acyclical Graph (DAG) model in Figure 1 confounders were selected to adjust the statistical model.

Statistical Methods

Baseline Characteristics

Baseline characteristics are important, because they provide essential information on the participant and help to ensure that the study is conducted in a fair and unbiased manner. The baseline characteristics of the study were summarised. The categorical variables were summarised as frequencies and means, whereas the continuous variables as means SD. The baseline characteristics will be stratified by both the exposure as the outcome variable.

Table 1 presents the baseline characteristics of 40,000 individuals who were included in the study. The mean age 75.65 and the majority of the study population 61.91% were female. Approximately half of the respondent have Cardiovascular disease, whereas 12.34 percent have Pulmonary disease and 6.54% have Diabetes Mellitus.

A total of 29616 (74.04%) respondents were vaccinated, whreas 10384 (25.96%) were not vaccinated. People who received a vaccination had on average more often pulmonary disease (14.33% vs 6.67%), cardiovascular disease (53.02% vs 39.05%) and diabetes (7.51% vs 3.8%).

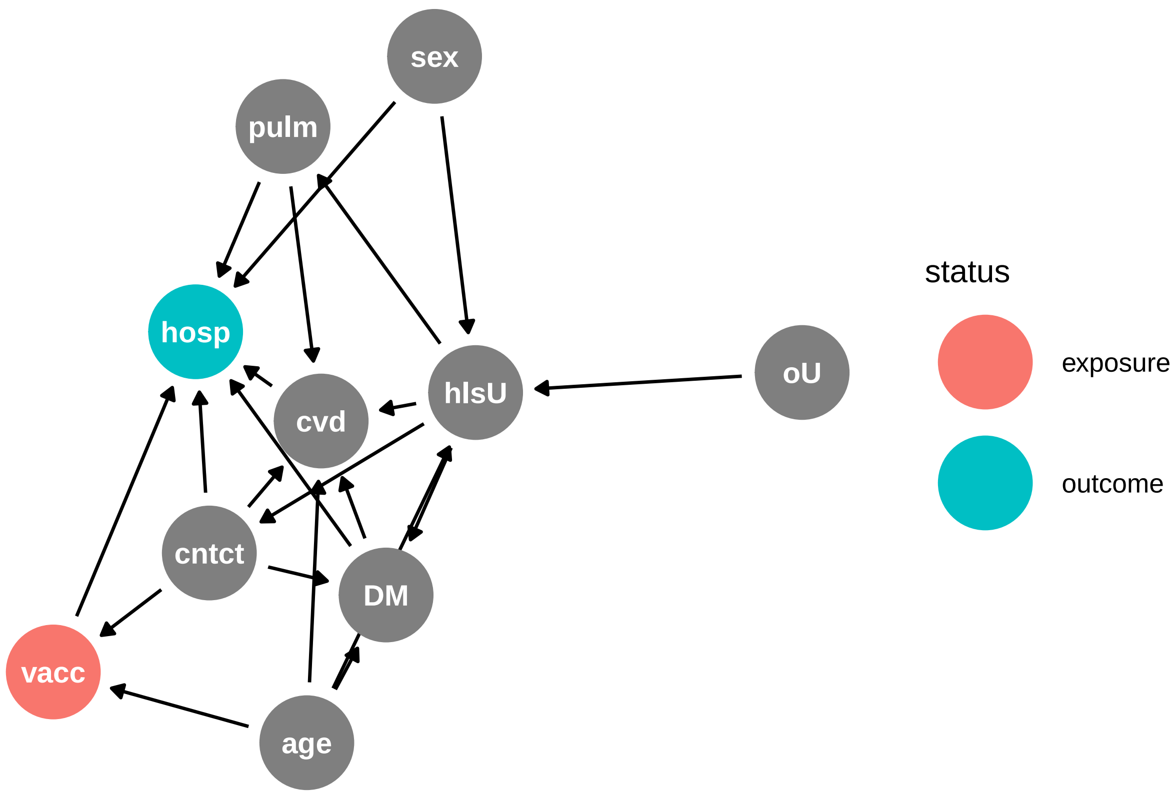

Figure 1: DAG model with Vaccine as exposure and hospitalization as outcome

Baseline characteristics stratified by the study outcome indicate that 254 of the respondents were hospitalized as opposed to 39746 who were not. Respondents who were hospitalized were older (79.68 vs 75.63), had more often contact with the GP (21.08 vs 14.71), were on average less often a female (48.43% vs 61.99%), had on average more often cardiovascular disease (72.05% vs 49.25%) and on average more often Diabetes (11.42% vs 6.51%).

Propensity Scores (PS)

In order to control for confounding (observed and unobserved), a PS was estimated by fitting a logistic regression. A PS gives the probability being vaccinated for the respondents. The variables used to calculate the propensity score are derived from the aforementioned DAG, namely: age, sex, cardiovascular disease, pulmonary disease, diabetis mellitus, and GP contact in 12 months prior to start of study. For the variables age and contact a spline is used, because it has shown favorable performance compared to other propensity score methods (Tian, Baro & Zhang, 2019).

In Figure 2, the PS for vaccinated and unvaccinated individuals appear to be well-balanced, indicating that the distribution of covariates between the two groups is similar. This plot also supports the positivity assumption, which means that both vaccinated and unvaccinated individuals are represented in all sub-populations defined by different combinations of covariates (Westreich & Cole, 2010).

PS as Covariate

The first adjusted model is to use the PS as a covariate. By including PS as a covariate, the effect of treatment or exposure can be estimated while adjusteed for potential confounders. This method can provide mroe statistical power and precision than matching, particularly when the sample size is large (Austin, 2011)

Belangrijk Wat moet hier nog bij?

Inverse Probability Weighting (IPW)

With the aforementioned propensity scores, a

We see that the mean propensity score for the ones who are vaccinated is higher (As expected)

Assignment These are the unstabilised weights of the inverse probability score.

We can use stabilised weights as well

What's the difference compared to unstabilized weights? *The difference is in the numerator of the weights. Here, the results are equal, because the numerator is just a constant. However, the numerator model can include confounders too, in which case the stabilized weight yield more stable estimates.*

There is non-positivity

- Checking assumptions of method to control for confounding
- Implementation of methods to control for (observed/unobserved) confounding

Results

Baseline Characteristics

- Reporting characteristics of study population
- Reporting crude/adjusted effect measures

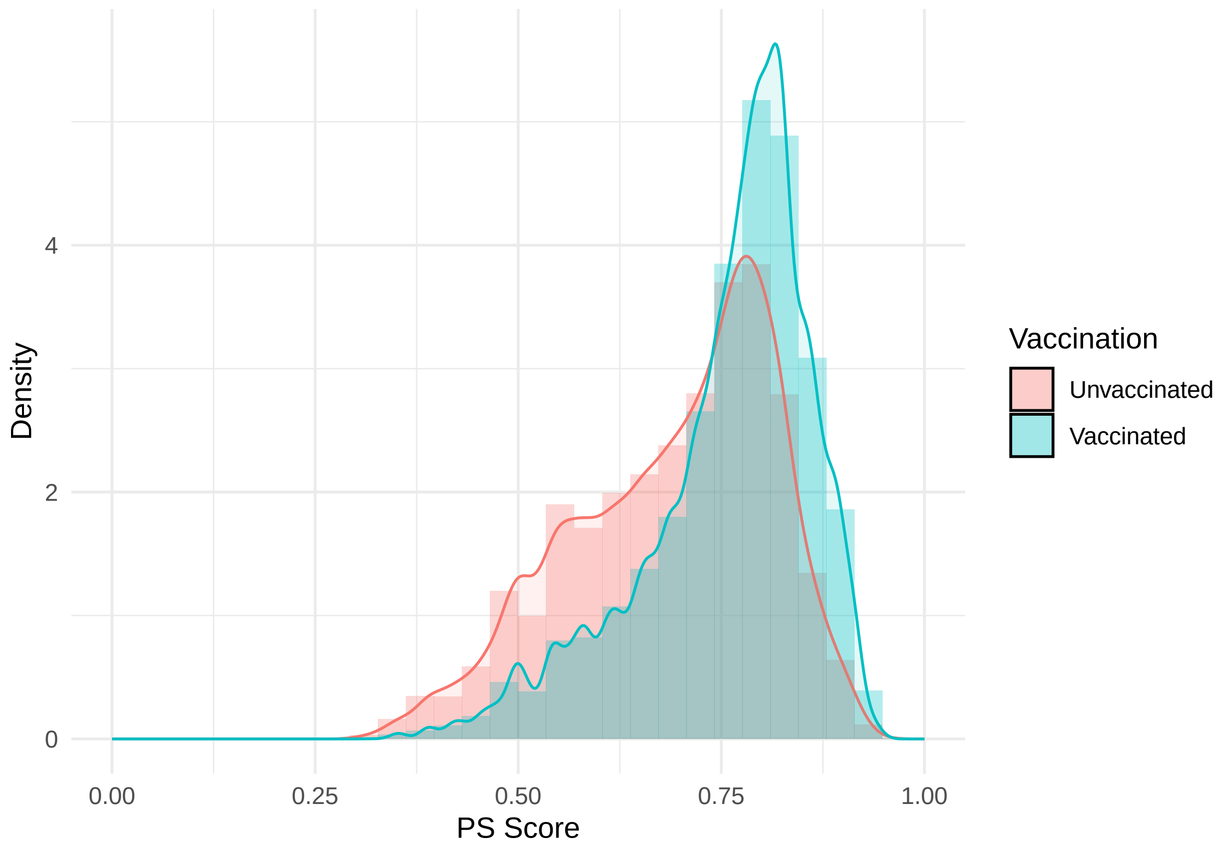

Figure 2: Distribution of the Propensity score for participants who received a vaccination compared with those who did not receive a vaccination

Table 1: Baseline Characteristics stratified by Influenza vaccination received and Hospitalisation

		Influenza vaccination received		Hospitalized	
Characteristics	Total	Yes	No	Yes	No
N	40000	29616	10384	254	39746
Age, mean (SD)	75.65 (6.97)	75.9 (6.83)	74.97 (7.32)	79.68 (7.2)	75.63 (6.97)
Contact, mean (SD)	14.75 (11.54)	15.85 (11.73)	11.64 (10.38)	21.08 (15.59)	14.71 (11.5)
Female, n (%)	24763 (61.91)	18022 (60.85)	6741 (64.92)	123 (48.43)	24640 (61.99)
Pulmonary disease, n (%)	4937 (12.34)	4244 (14.33)	693 (6.67)	60 (14.33)	4877 (12.27)
Cardiovascular disease, n (%)	19757 (49.39)	15702 (53.02)	4055 (39.05)	183 (72.05)	19574 (49.25)
Diabetes mellitus, n(%)	2618 (6.54)	2223 (7.51)	395 (3.8)	29 (11.42)	2589 (6.51)
Received Influenza Vaccination, n (%)				184 (72.44)	29432 (74.05)

Crude measures:

As shown in Table 2, the crude association between vaccination status and hospitalization was examined using logistic regression, and the odds ratio was found to be 0.921 (95% CI: 0.699, 1.214), indicating a non-significant association (p = 0.562). The C-statistic, a measure of discrimination, was 0.508, suggesting poor predictive performance of the model.

Considering the second model in Table 2, which presents the PS score as covariate, the Odds Ratio suggests a significant negative association (adjusted OR: 0.635. 95% CI 0.478 to 0.843; p <.001). The C-Statistic of 0.679 suggests that the model has moderate discriminatory power.

The IPW model with unstabilised weights yields significant negative association between the exposure and outcome variable (adjusted OR: 0.617, 95%CI 0.523 to 0.728; p <.001). The model presents moderate discriminatory power, where C = 0.573.

The last model in Table 2 is an IPW model with stabilised weights. This model also presents a negative association between vaccination and hospitalisation (Adjusted OR: 0.617, 95% CI 0.509 to 0.748; p <.001). The C-statistic is

Table 2: Association between influenza vaccination and hospitalization (n=40000)

Model Specification	OR (95% CI)	P-Value	C-Statistic
Unadjusted	0.921 (0.699 to 1.214)	0.562	0.508
PS score as covariate	0.635 (0.478 to 0.843)	<.001	0.679
Unstabilised IPW	0.617 (0.523 to 0.728)	<.001	0.573
Stabilised IPW	0.617 (0.509 to 0.748)	<.001	0.573

Adjusted measures

Conclusion / Discussion

- Conclusions supported by data
- Other issues (both positive and negative) Maybe write something about matching here?

References

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behavioral Research, 46(3), 399-424.

Shiba K, Kawahara T. Using Propensity Scores for Causal Inference: Pitfalls and Tips. J Epidemiol. 2021 Aug 5;31(8):457-463. doi: 10.2188/jea.JE20210145. Epub 2021 Jun 12. PMID: 34121051; PMCID: PMC8275441.

Yuxi Tian, Elande Baro & Rongmei Zhang (2019) Performance evaluation of regression splines for propensity score adjustment in post-market safety analysis with multiple treatments, Journal of Biopharmaceutical Statistics, 29:5, 810-821, DOI: 10.1080/10543406.2019.1657138

Westreich, D., & Cole, S. R. (2010). Invited commentary: positivity in practice. American journal of epidemiology, 171(6), 674-677.