The Impact of a Mouthwash Program on the Risk of Nosocomial Pneumonia at the ICU

Marcela Granados, Fernando Rosso, Álvaro J. Riascos, Natalia Serna*

June 5, 2017

Abstract

During June 2006 a high complexity hospital in Colombia underwent a mouthwash program in critical care patients to reduce the risk of nosocomial pneumonia. In this paper we measure the impact of the program on the probability of developing nosocomial pneumonia using two empirical approaches: machine learning methods to recover counterfactual probabilities assuming the program never went in force and a matching procedure that guarantees patients are comparable in all relevant characteristics. We find that the mouthwash program has reduced significantly the probability of nosocomial pneumonia by almost 100% relative to the observed scenario. We also find that diagnoses such as cancer, hepatic diseases, and neurologic diseases are some of the most relevant risk factors of nosocomial pneumonia.

Keywords: Counterfactual analysis, machine learning, healthcare, nosocomial pneumonia, matching techniques

1 Introduction

During June 2006, the Fundación Valle del Lili (FVL), a high complexity hospital in Colombia, underwent a mouthwash program in critical care patients to reduce the incidence of nosocomial pneumonia (also known as ventilator-associated pneumonia) at the Intensive Care Unit (ICU). In the late 2008, the program was extended to include a patient body wash treatment. The program was implemented because critical care patients are usually more exposed to sources of infections compared to patients in other areas of the hospital and doctors wanted to reduce the risk of patients developing endangering states during the ICU stay. Doctors at the FVL identified the use of mechanical ventilation in critical care patients as one of the most important risk factors for this type of pneumonia. However, there is no empirical evidence supporting this believe nor measuring the impact of the mouthwash program on patient health outcomes. Besides being an interesting outcome, the incidence of nosocomial pneumonia is also an indicator of hospital quality. Our purpose with this paper, then, is two-fold: first, measuring the impact of the mouthwash program on the probability of developing nosocomial pneumonia using machine learning techniques and matching techniques, and, second, finding the most important exogenous risk factors for this type of pneumonia.

The critical care literature has mostly revolved around our second objective but the first is novel both from the point of view of the empirical techniques used as well as from the point of view of the mouthwash program. Chastre et al. (1998) find that patients with Acute Respiratory Distress Syndrome are not more

*Álvaro Riascos: Quantil - Universidad de los Andes. e-mail: alvaro.riascos@quantil.com.co. Natalia Serna: Quantil. e-mail: natalia.serna@quantil.com.co. Marcela Granados: Fundación Valle del Lili, e-mail: mgranados@fcvl.org. Fernando Rosso: Fundación Valle del Lili, e-mail: frosso@fcvl.org

likely to develop nosocomial pneumonia compared to patients without the syndrome, thus suggesting respiratory diseases are not always a risk factor.factors for nosocomial pneumonia: comparing adult critical-care popoulations (1996) agree that prolongued mechanical ventilation is the strongest predictor of nosocomial pneumonia. They also add the Apache score as a potential risk factor, while Fagon et al. (1996) mentions the number of dysfunctional organs, Esperatti et al. (2010) the etiologic diagnoses, and Joshi et al. (1992) procedures such as bronchoscopy and the presence of a nasogastric tube as risk factors for nosocomial pneumonia. In terms of sociodemographic factors, Rello et al. (2002) finds men are more likely to develop this type pneumonia compared to women as well as patients with trauma admission compared to patients with medical or surgical admission. Most of these studies rely on randomized control trials for capturing the relevant risk factors and some use matching techniques to analyze the impact of nosocomial pneumoniaon certain outcomes like mortality, health costs, and length-of-stay. We build up on the matching techniques by adding machine learning to the impact evaluation methods. Using machine learning we build a counterfactual scenario and argue that features selected statistically are the only ones that matter for prediction of the probability of nosocomial pneumonia, thus, the counterfactual in the absence of the mouthwash program is an unbiased measure of the impact of the program on the risk of nosocomial pneumonia.

2 Methods

For the two purposes of this study, we use machine learning-based models on the data of patient admissions to the FVL's ICU from 1998 to 2015. For predicting the probability of nosocomial pneumonia, we estimate and compare the following models:

- Logit.
- Artificial Neural Networks (ANN).
- Random Forests (RF).
- Boosted trees (BT).

The parameters for the ANN (number of neurons in the hidden layer and weight decay), the RF (number of trees) and the BT (number of trees, shrinkage, and degree of interaction between variables) are chosen using 10-fold cross-validation on a grid of values, in order to maximize the area under the ROC curve in a train set. In the next section we explain how the train and test sets are built. These models are fitted on a subset of features X for each patient that are considered exogenous to the program and relevant to predict the event of nosocomial pneumonia, as well as a on dummy variable defined below which denotes the beginning of the mouthwash program:

$$T_i = \left\{ \begin{array}{ll} 1 & \text{if patient } i \text{ is admitted after June 2006} \\ 0 & \text{otherwise} \end{array} \right\} \tag{1}$$

The probability of nosocomial pneumonia under the observed scenario is then:

$$\widehat{Pr}[y_i = 1 | X_i, T_i] = f(X_i, T_i) \tag{2}$$

where f is modeled in accordance to each of the four machine learning techniques considered and y_i is the event of having nosocomial pneumonia. The counterfactual probability assuming the program never went in force is:

$$\widehat{Pr}[y_i = 1 | X_i, T_i] = f(X_i, 0) = f(X_i) \tag{3}$$

Thus, the impact of the mouthwash program on the probability of nosocomial pneumonia using machine learning techniques can be estimated as:

$$\frac{1}{N}\sum_{i=1}^{N}(f(X_i, T_i) - f(X_i)) \tag{4}$$

The second approach to measure the impact of the mouthwash program is a standard matching technique. Intuitively, for this approach we build clusters of patients based on the subset of selected features X. Within each cluster we will have patients that received the mouthwash ($T_i = 1$) and patients who didn't ($T_i = 0$). Since the variables in X are the only ones that are relevant for prediction of the probability of nosocomial pneumonia, guaranteeing patients are identical conditional on X, allows us to obtain an unbiased estimate of the impact of the mouthwash program. There are two obvious challenges with this approach: the first, is guaranteeing patients within each cluster are in fact identical to argue unbiasedness and, the second, is having enough patients in each cluster to argue robustness.

Before creating the clusters, each continuous variable in X, namely X_c, is dichotomized using a decision tree. We estimate as many trees as continuous variables there are in X. Each tree predicts the event of nosocomial pneumonia y minimizing a loss function by choosing a splitting value γ_c for each predictor X_c:

$$\hat{y} = B_c(X_c; \gamma_c) \tag{5}$$

Then, the variable is dichotomized as:

$$D_c = \left\{ \begin{array}{ll} 1 & \text{if } X_c \geq \gamma_c \\ 0 & \text{otherwise} \end{array} \right\} \tag{6}$$

With the full set of dichotomous variables D, we build clusters according to the combination of values of each variable in D. Let N^T be the number of patients who received the mouthwash (treated), N_i^M the number of patients who did not receive the mouthwash (controls) and who are contained within the same cluster as the treated patient i, y_i^T the probability of nosocomial pneumonia of treated patient j, and y_j^M the probability of nosocomial pneumonia of control patient j. The average treatment effect on the treated (ATT) is:

$$\tau = \frac{1}{N^T}\sum_{i \in T} y_i^T - \frac{1}{N^T}\sum_{j \in M}\frac{1}{N_i^M} y_j^M \tag{7}$$

3 Data and descriptive evidence

To predict the probability of nosocomial pneumonia and measure the impact of the mouthwash program at the FVL, we have a database of all patients admitted to the ICU from 1998 to 2015. The data is disaggregated at the level of admissions. Per patient admission we observe the admission date, patient age, gender, health insurer, municipality of residence, admission diagnosis, cause of admission, APACHE II score, length of stay, catheter days, bladder catheter days, procedures, complications, and other indicator of the patient's health state. Table (1) shows some summary descriptive statistics among the group of patients with nosocomial pneumonia and the group of patients without the infection. The mean and the standard deviation of each variable is reported as well as the differences in means between the two groups. These statistics show the proportion of males among patients with nosocomial pneumonia is significantly higher than the proportion of males without the infection. Infected patients are 4 years younger than uninfected patients and, overall, receive more blood transfusions, have higher platelet count, higher Apache score and remain more days with catheters compared to patients without nosocomial pneumonia. In particular, the table shows there is a difference of 266 hours of invasive ventilation between patients in both groups, which is the risk factor reported by the literature. Also, among patient with nosocomial pneumonia there is a higher incidence of shock, respiratory diseases, and neurologic diseases relative to the control group.

Table 1: Summary descriptive statistics

	No N. Pneumonia		With N. Pneumonia			
Variables	Mean (1)	S.D	Mean (2)	S.D	(1)-(2)	
Demographics						
Male	0.555	0.497	0.642	0.48	-0.087	***
Age	58.003	18.527	53.359	19.286	4.644	***
Hospital						
Red blood cells	0.031	0.174	0.07	0.255	-0.039	***
Other transfusions	0.021	0.144	0.045	0.207	-0.024	***
Platelets 150000	0.025	0.156	0.053	0.224	-0.028	***
Apache	13.221	6.383	16.545	5.994	-3.324	***
Length of stay	3.936	20.713	17.067	13.162	-13.131	***
Catheter days	1.583	3.669	9.447	10.172	-7.864	***
Bladder catheter days	2.163	4.229	13.742	14.343	-11.579	***
# of central catheters	0.402	1.045	1.253	3.363	-0.851	***
# of Swan Ganz	0.091	0.433	0.377	0.74	-0.286	***
# of arterial lines	0.496	0.593	1.171	1.665	-0.675	***
Hrs of inv. ventil.	27.502	92.114	294.091	301.764	-266.589	***
Hrs of non inv. lentil.	0.016	0.821	0	0	0.016	***
Cardiology admission cause	0.281	0.449	0.123	0.329	0.158	***
Medical admission cause	0.418	0.493	0.455	0.498	-0.037	**
Surgical admission cause	0.3	0.458	0.422	0.494	-0.122	***
From catheterism	0.021	0.144	0.005	0.073	0.016	***
From surgery	0.211	0.408	0.161	0.368	0.049	***
From other ICU	0.044	0.206	0.085	0.279	-0.04	***
Admission diagnoses						
Shock	0.057	0.232	0.095	0.294	-0.038	***
Respiratory	0.077	0.267	0.134	0.34	-0.057	***
Major post-op	0.232	0.422	0.184	0.388	0.048	***
Neurologic	0.073	0.261	0.198	0.399	-0.125	***
Trauma	0.055	0.229	0.187	0.39	-0.131	***
Gastrointestinal	0.035	0.183	0.017	0.13	0.018	**
Cardiac	0.29	0.454	0.09	0.286	0.2	***
Cardiac risk	0.009	0.096	0.001	0.036	0.008	**
Renal	0.021	0.142	0.009	0.096	0.011	**
Electrolyte imbalace	0.025	0.155	0.012	0.109	0.013	**
Multiorgan failure	0.005	0.07	0.024	0.153	-0.019	***

This table shows the mean of the predictors in our database conditioned on the event of nosocomial pneumonia. The column "diff" shows the differences in means between the base group (patients without pneumonia) and the treated group (patients with pneumonia). ***p<0.01, **p<0.05, *p<0.1.

Figure (1) shows the proportion of patients with nosocomial pneumonia of those admitted to the ICU each month. First of all, the proportion of patients with nosocomial pneumonia is highly volatile but, second of all, there is evidence of reduction of the proportion just after June 2006 when the mouthwash program began. In the early 2009 there is another reduction in the proportion which could be due to the extension of the program to body wash treatment. Our purpose is to test if the observed reductions are actually explained by the program or not.

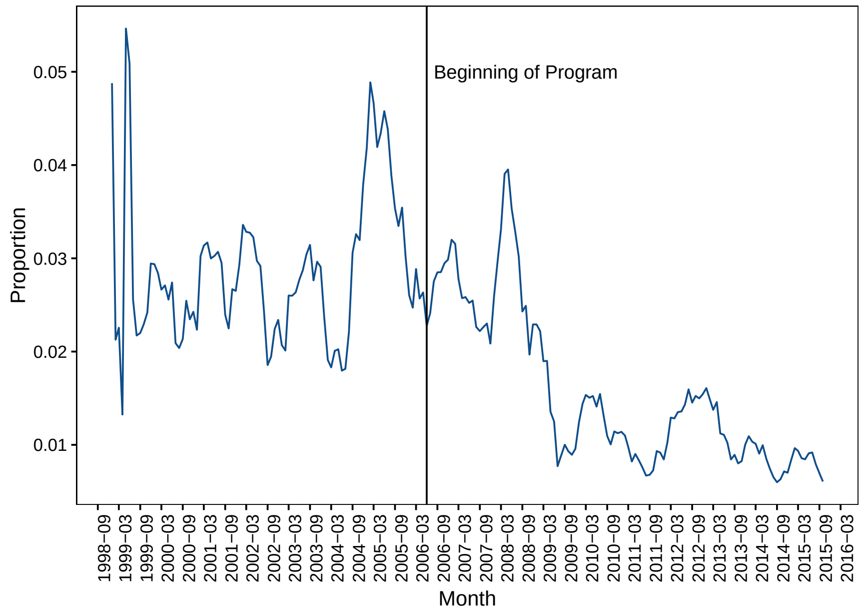


Figure 1: Proportion of patients with nosocomial pneumonia (6-month moving average)
FVL Data. Authors' calculations.

The data processing procedure consists of deleting outliers in the following variables: hours of invasive ventilation, catheter days, length-of-stay, bladder catheter days and number of central catheters. We define outliers as those observations where the variable takes a value greater than the 99th percentile of its distribution.We also delete observations associated to an Apache score greater than 71 because those scores are due to misregistrations.

The full database of patients is then split into two mutually exclusive datasets, train and test, comprising 70 and 30% of the total number of patients, respectively. Notice the datasets are built from randomly choosing patients but not admissions, this way, first, we can have all admissions associated to a single patient in the same database and, second, we avoid overfitting predictions because the characteristics of the admissions for a single patient may be correlated.

Before training the machine learning-based models we use two feature selection techniques to choose the variables that most accurately predict the probability of nosocomial pneumonia: logit regressions with forward and backward stepwise selection and a boosted tree model with thresholds for variable relative influence. We restrict the set of variables for feature selection to those that are measured at the moment of admission since variables measured afterwards can be endogenous to the mouthwash program and therefore could bias the estimated effect. In the case of the logit regressions we estimate 50 models and then select those features that are kept in the final subset of variables of at least 20 models. In the case of the boosted tree model, we set the threshold for variable relative influence to 3 such that variables whose relative influence is greater than or equal to the threshold are selected for the final subset of predictors. We did some robustness checks on the threshold but found no significant gains in predictive power by reducing the threshold or increasing it. The result of this process yielded the following risk factors:

- Patient coming from surgery
- Patient coming from another ICU
- Patient coming from intermediate ICU
- Admission during July

- Shock
- Respiratory diseases
- Major post-op
- Trombosis
- Neurologic diseases
- Trauma
- Gastrointestinal diseases
- Cardiac diseases
- Infections
- Renal diseases
- Electrolyte imbalance
- Multiple organ failure
- Cancer
- Poisoning
- Apache
- Age
- Number of admitted patients per day
- Indicator of mouthwash program

4 Results

The four machine learning-based models where estimated in the train set on the subset of variables listed in the previous section. We compare the accuracy of the models using the out-of-sample area under the ROC curve (AUC) computed in the test set. Figure (2) shows the AUC of the four models. The artificial neural network with 3 neurons in the hidden layer achieved the highest AUC, 75.6%, while the random forest model with 100 trees achieved the lowest. However, there are no significant differences in accuracy between the ANN and the BT models. We select the BT model for computation of the counterfactual probabilities because even if a patient has a missing value in any of the variables in the final subset, the model is able to predict the probability of nosocomial pneumonia with the remaining non-missing features.

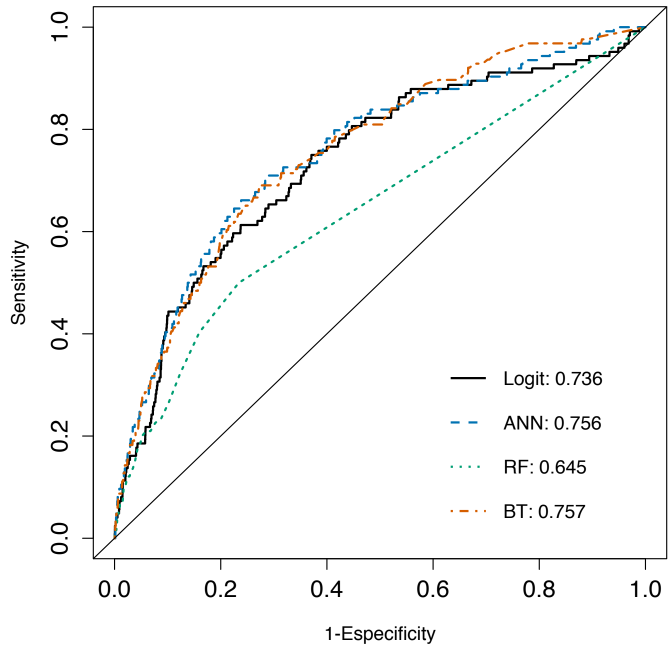


Figure 2: ROC curves of the logit, ANN, RF, and BT models
FVL Data. Authors' calculations.

Using the BT model we predict the probability of nosocomial pneumonia in the full sample (train + test) and then compute the counterfactual probabilities assuming the mouthwash program never went in force or assuming $T_i = 0 \; \forall \; i$. Figure (3) shows observed proportion of patients with nosocomial pneumonia each month, the average predicted probability, and the average counterfactual probability. Notice the BT model adjusts the observed proportion of patients with nosocomial pneumonia very well: it accurately predicts the mean although is less precise in the deviations from the mean. This suggests the model could recreate the probabilities under different situations also quite accurately. The average counterfactual probability assuming no mouthwash program is the orange line in the figure. In the absence of a mouthwash program the average probability of nosocomial pneumonia would have been significantly greater than the observed one, going from a 1% in the observed scenario to a 2% in the counterfactual. This difference represents a 100% increase in the risk of nosocomial pneumonia for the observed case. Conversely, this result shows the mouthwash program has accounted for a 100% decrease in the probability of nosocomial pneumonia at the FVL's ICU, and the effect is relatively stable in the time series.

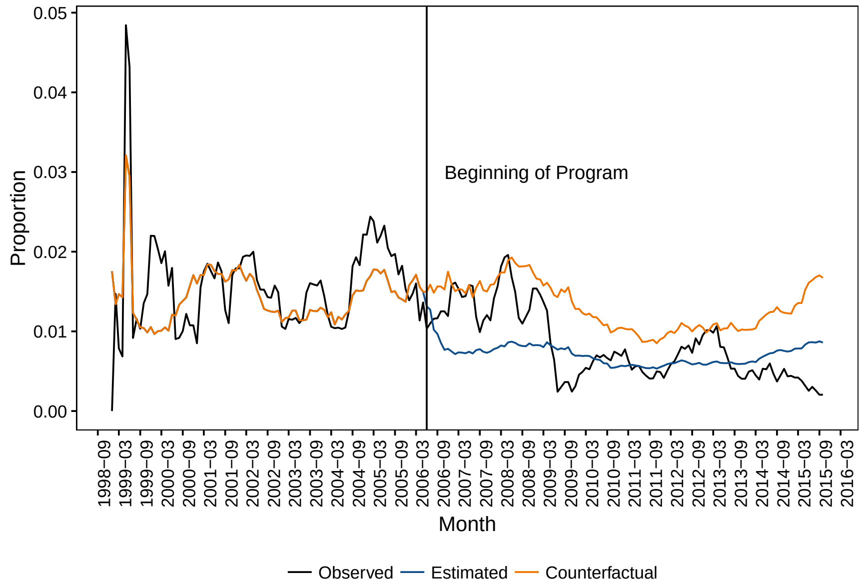


Figure 3: Counterfactual probability of nosocomial pneumonia
FVL Data. Authors' calculations.

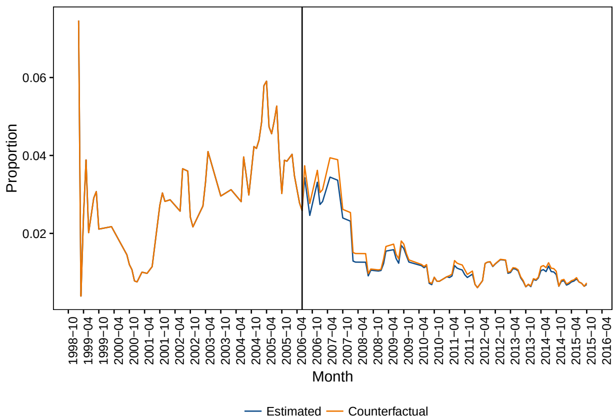
Figure (4) shows the average estimated probability and the average counterfactual probability of the BT model fitted on the subset of patients with a particular disease. Evidence from this figure suggests the mouthwash program had a significant effect in patients with multiple organ failure, poisoning, neurologic diseases, pathologies of the aorta, major post-op, respiratory diseases, shock, and trauma in reducing their probability of developing hospital-acquired pneumonia at the ICU. However, the figure also shows the program had no effect in patients with trombosis, renal diseases, infections, hepatic diseases, gastrointestinal diseases, and cardiac diseases.

Table (2) shows the average treatment effect based on the machine learning model following equation (4) and the ATT following the matching procedure and equation (7). Both the machine learning (ML) approach to computation of counterfactual probabilities and the matching technique conclude the mouthwash program had at least a negative effect on the probability of nosocomial pneumonia reducing it in 0.4 percentage points in the first case and 1 percentage point in the second case. Relative to the average observed probability, these numbers represent a 80 and a 100% decrease in the risk of nosocomial pneumonia, respectively. However for the ML approach the effect is indistinct from zero while for the matching approach it is significantly negative. Table (3) shows the balancing property of the matching procedure is satisfied. The table reports no significant differences in the observed characteristics on which we matched treated individuals to control individuals, thus ensuring they are comparable.

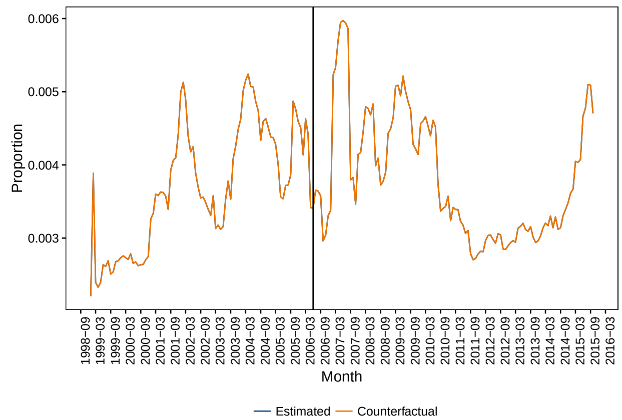
Table 2: Effect of the mouthwash program

Variable	ML (ATT)	Matching (ATT)
Probability of N. Pneumonia	-0.004	-0.017***
	(0.010)	(0.001)
Observations	30,474	31,765†

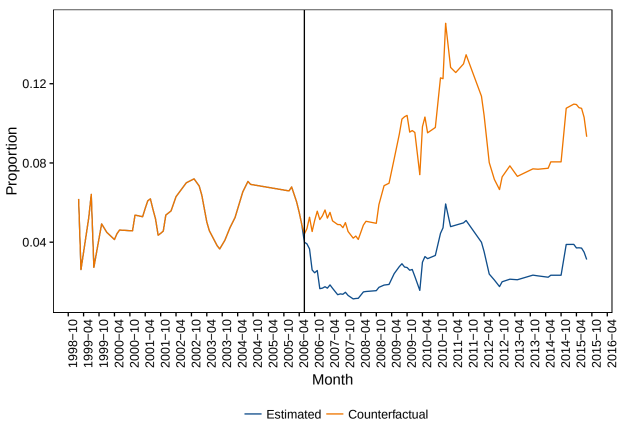
This table shows average treatment effect estimated with machine learning techniques and the average treatment on the treated using a matching technique. Standard deviation of the difference reported in parenthesis for the ML technique and bootstrap standard error reported in parenthesis for the matching technique. † correspond to matched observations.
***p<0.01, **p<0.05, *p<0.1.



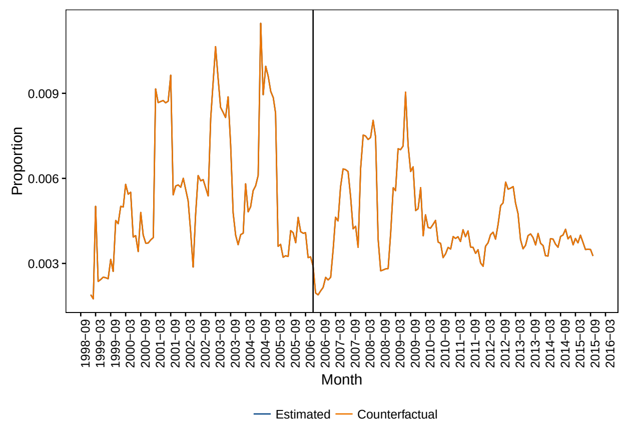
(a) Cancer



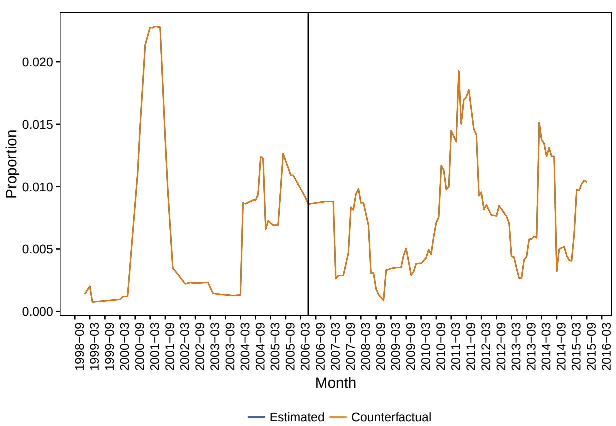
(b) Cardiac diseases



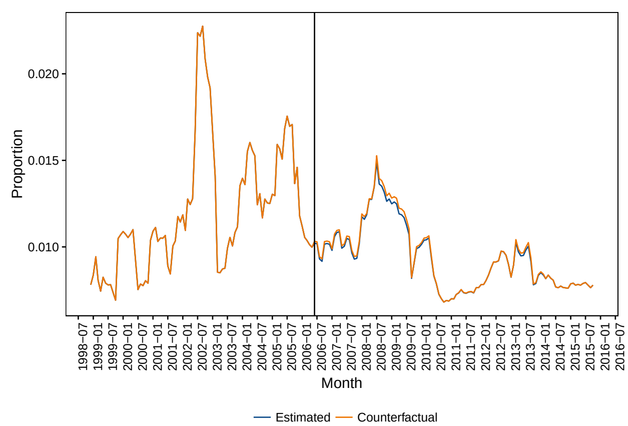
(c) Multiorgan failure



(d) Gastrointestinal disease

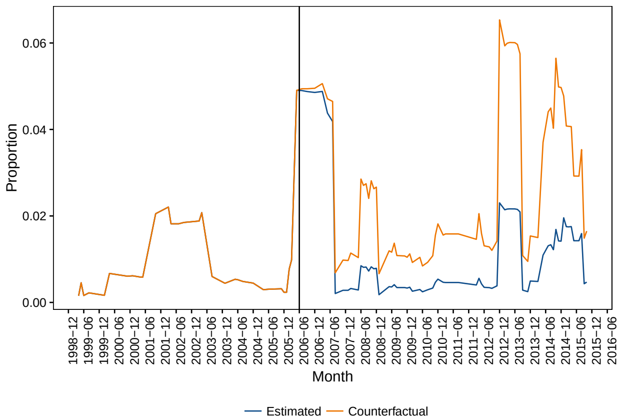


(e) Hepatic

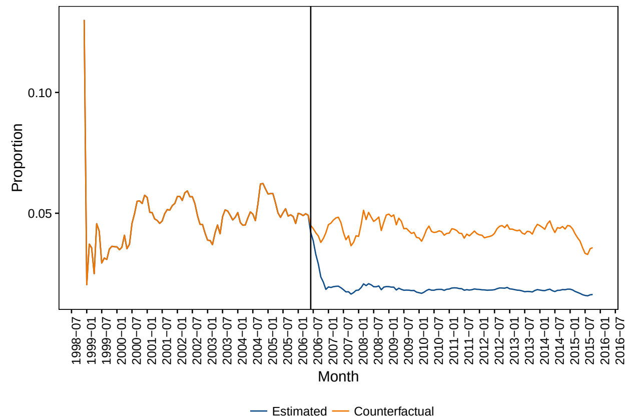


(f) Infections

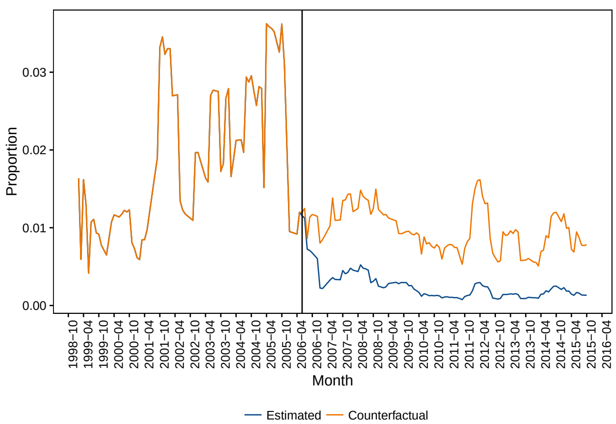
Figure 4: Counterfactual probability of nosocomial pneumonia
FVL Data. Authors' calculations.



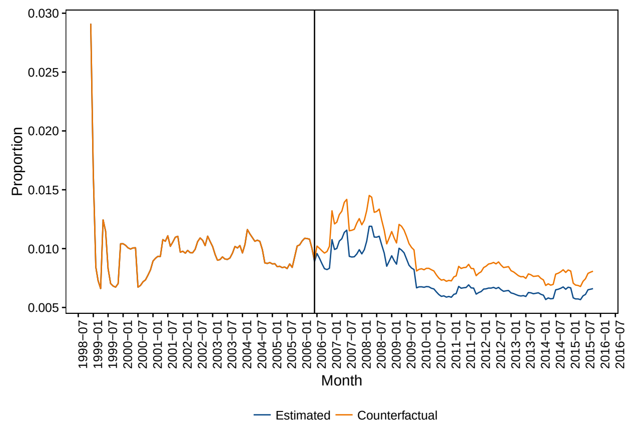
(g) Poisoning



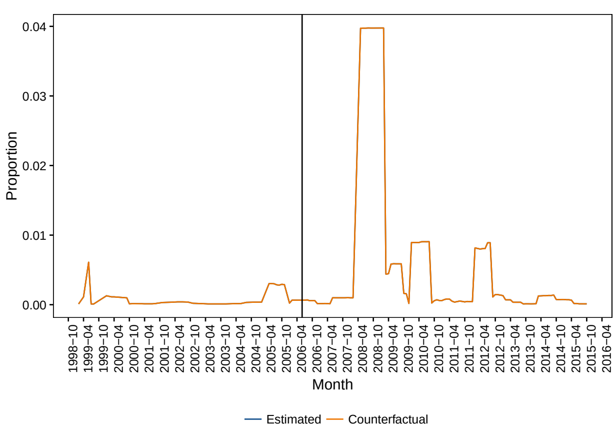
(h) Neurologic diseases



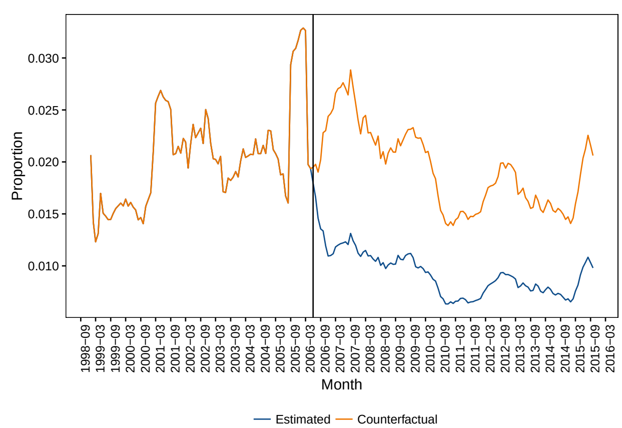
(i) Pathologies of the aorta



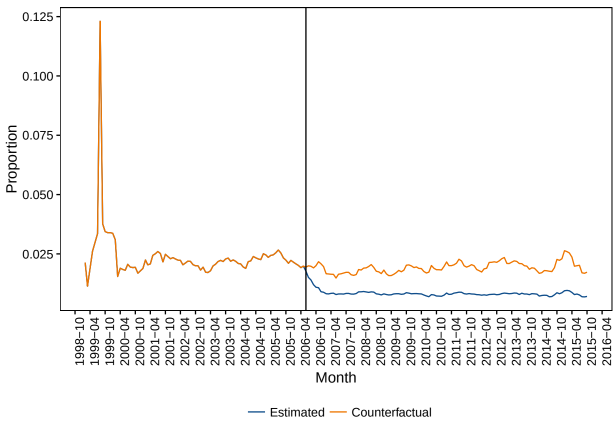
(j) Major post-op



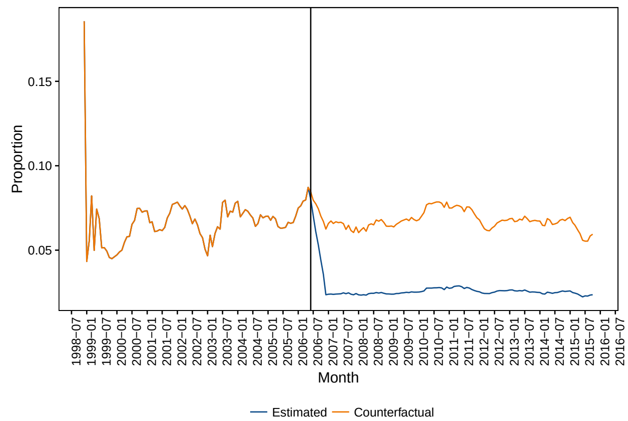
(k) Renal diseases



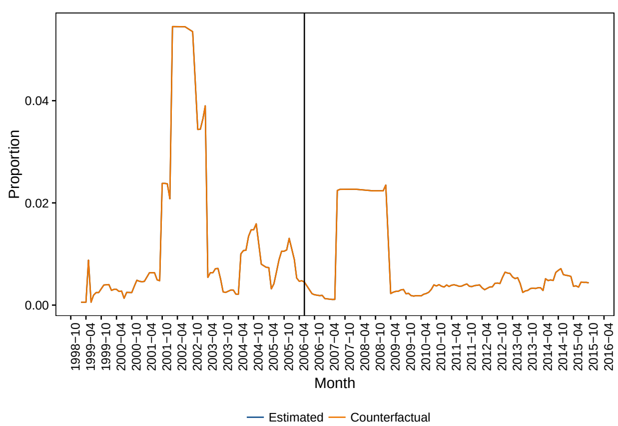
(l) Respiratory diseases



(m) Shock



(n) Trauma



(o) Trombosis

Table 3: Balancing property

	Treated (1)	Control (2)	diff (1)-(2)
From surgery	0.2045	0.2040	0.0005
From another ICU	0.0397	0.0397	0.0000
From intermediate ICU	0.0039	0.0000	0.0039
Shock	0.0572	0.0572	0.0000
Respiratory	0.0776	0.0776	0.0000
Major Post-op	0.2078	0.2079	-0.0001
Trombosis	0.0168	0.0168	0.0000
Neurologic diseases	0.0751	0.0751	0.0000
Trauma	0.0604	0.0604	0.0000
Gastrointestinal diseases	0.0385	0.0385	0.0000
Cardiac diseases	0.2684	0.2684	-0.0001
Infections	0.0980	0.0980	0.0000
Rena diseases	0.0262	0.0262	0.0000
Electrolyte imbalance	0.0321	0.0321	0.0000
Multiple organ failure	0.0046	0.0046	0.0000
Hepatic diseases	0.0082	0.0082	0.0000
Cancer	0.0102	0.0102	0.0000
Poisoning	0.0038	0.0038	0.0000
Admission in July	0.0946	0.0942	0.0003
Age	0.7446	0.7455	-0.0009
Apache	0.2706	0.2657	0.0049
Number of admitted patients	0.7580	0.7404	0.0176

This table shows the variable means among the treated group (patients subject to the mouthwash program) and among the control group (patients not subject to the mouthwash program) in the matched observations. The third column reports the differences in means and whether the difference is significant. ***p<0.01, **p<0.05, *p<0.1.

5 Robustness check

In this section we report some robustness checks for the matching procedure. The tests consist of varying the time window around the beginning of the mouthwash program in which patients are admitted. We perform three exercises: one with patients admitted two months before and after the beginning of the program, another one with patients admitted one month around the beginning of the program, and the last with patients admitted just 2 weeks before and after the beginning of the program. This tests allow us to control for factors that change over time and that could affect the conditions of the ICU or the characteristics of the patients, biasing the effect reported in the previous section. Table (4) shows the ATT on these subsets of patients. Although the estimated effect is indifferent from zero in the three exercises because of the small number of matched observations, results suggest the smaller the time window the greater the effect of the program on the probability of nosocomial pneumonia. The mouthwash program accounts for a 0.5 percentage point reduction in the probability of nosocomial pneumonia within a 2-month and 1-month time window and then increases to a 1 percentage point reduction in a 2-week time window. This is consistent with the effect reported in the previous section were we used the full sample of patients.

Table 4: Robustness check of the impact of the mouthwash program

	2 months	1 month	2 weeks
ATT	-0.0051	-0.0052	-0.0101
s.e	(0.0102)	(0.0103)	(0.0111)
Matched obs	392	385	296

This table shows the ATT on the matched subset of patients admitted two months before and after the beginning of the program in column 1, one month before and after the beginning of the program in column 2, and 15 days before and after the beginning of the program in column 3. Standard errors in parenthesis.

6 Conclusions

Nosocomial pneumonia is one of the most prevalent hospital-acquired infection in Colombia. The Fundación Valle del Lili, a hospital in this country, implemented a mouthwash program for patients in the intensive care unit in order to reduce the risk of nosocomial pneumonia. In this paper we measure the impact of the program on the probability of developing nosocomial pneumonia using machine learning techniques and matching techniques. We find the risk factors that explain most of the variation in the probability of such infection. Results show diagnoses like neurologic diseases, cardiac diseases and cancer are relevant risk factors as well as the Apache II score and demographic characteristics such as the patient's age. We found no significant differences in the accuracy of the different machine learning-based models but used the boosted tree model to compute counterfactual probabilities. In the counterfactual scenario we assumed the mouthwash program never went in force and calculated the resulting probabilities per patient. This empirical approach showed the program reduced the average probability of nosocomial pneumonia by 80% at most. In order to guarantee the patients who received the mouthwash program and the patients that did not are in fact comparable, we estimated the average treatment on the treated (ATT) using a matching technique. The matching consisted of creating clusters of patients with the exact same characteristics (based on the obtained risk factors), then computing the difference in the average probability of nosocomial pneumonia of treated patients and control patients in each cluster, and finally averaging the differences among clusters. This empirical approach showed the mouthwash program reduced the probability of nosocomial pneumonia by almost 100% relative to the observed scenario and this effect is significant. Robustness checks choosing patients that were admitted 2 months, 1 month and just 2 weeks before and after the beginning of the program also showed a similar effect. Overall, these results have important policy implications for health management. For instance, a policy of patient mouth hygiene conditional on the type of admission diagnosis can improve the patient's quality of life by reducing the risk of developing endangering states and can improve the measures of hospital quality.

References

Chastre, J., Trouillet, J., Vuagnat, A., Joly-Guillou, M., Clavier, H., Dombret, M., and Gibert, C. (1998). Nosocomial pneumonia in patients with acute respiratory distress syndrome. *American Journal of Respiratory and Critical Care Medicine*, 157:1165–1172.

Esperatti, M., Ferrer, M., Theesen, A., Liapikou, A., Valencia, M., Saucedo, L., Zaval, E., Welte, T., and Torres, A. (2010). Nosocomial pneumonia in the intensive care unit acquired by mechanically ventilated versus nonventilated patients. *American Journal of Respiratory and Critical Care Medicine*, 182:1533–1539.

factors for nosocomial pneumonia: comparing adult critical-care popoulations, R. (1996). Cunnion, k., and weber, d., and broadhead, e., and hanson, i., and pieper, c., and rutala, w. *American Journal of Respiratory and Critical Care Medicine*, 153:158–162.

Fagon, J., Chastre, J., Vuagnat, A., Trouillet, J., Novara, A., and Gibert, C. (1996). Nosocomial pneumonia and mortality among patients in the intensive care units. *JAMA*, 275(11):866–869.

Joshi, N., Localio, R., and Hamory, B. (1992). A predictive risk index for nosocomial pneumonia in the intensive care unit. *The American Journal of Medicine*, 93:135–142.

Rello, J., a. O. D., Oster, G., and Vera-Llonch, M. (2002). Epidemiology and outcomes of ventilator-associated pneumonia in a large US database. *CHEST*, 122(6):2115–2121.