2015

Nurse Work Environment and Hospital Outcomes

Jae Hyuk (John) Han
University of Pennsylvania

Nurse Work Environment and Hospital Outcomes

Abstract
The research question is as follows: To what extent do nurse work environments affect a hospital's performance? The purpose of this paper is to identify potential relationships between a popular measure of nurse work environments called the Practice Environment Scale of the Nursing Work Index (PES-NWI) and patient outcomes, specifically 30 day mortality. While previous research on the same topic have largely utilized traditional regression approaches to study the relationship, this paper will use a pre-processing technique called matching to reduce the imbalance between observations in the treated and control groups. Matching enables us to reduce biases frequently present in many social science studies and strengthen the validity of the conclusions drawn. The resulting comparison of patient outcome showed no statistically significant difference between high and low PES-NWI hospitals.

Keywords
Nursing work environment, Hospital outcomes, Patient Environment Scale-Nurse Work Index, Statistical Matching

Disciplines
Business

Nurse Work Environment and Hospital Outcomes

Han, Jae Hyuk (John)

University of Pennsylvania – The Wharton School

Summer Program for Undergraduate Research 2015

Dr. Matthew McHugh – Professor in the School of Nursing

Dr. Dylan Small – Professor in the Department of Statistics

hanjh@wharton.upenn.edu

Statistics, Healthcare

ABSTRACT

The research question is as follows: To what extent do nurse work environments affect a hospital's performance? The purpose of this paper is to identify potential relationships between a popular measure of nurse work environments called the Practice Environment Scale of the Nursing Work Index (PES-NWI) and patient outcomes, specifically 30 day mortality. While previous research on the same topic have largely utilized traditional regression approaches to study the relationship, this paper will use a pre-processing technique called matching to reduce the imbalance between observations in the treated and control groups. Matching enables us to reduce biases frequently present in many social science studies and strengthen the validity of the conclusions drawn. The resulting comparison of patient outcome showed no statistically significant difference between high and low PES-NWI hospitals.

Key words: Nursing work environment, Hospital outcomes, Patient Environment Scale-Nurse Work Index, Statistical Matching

INTRODUCTION

Background

With the expansion of pay for performance and value-based purchasing programs in the healthcare industry, many hospitals across the United States are increasingly challenged to improve outcomes under tighter budget constraints. Unlike past payment systems that generally paid hospitals on a fee-for-service basis, current payment models put greater emphasis on performance and treatment quality (Tanenbaum 2009). As a result, an increasing number of hospital administrators are seeking ways to refine their strategic allocation of resources to boost performance, while keeping costs down.

Research has repeatedly shown that investments in nursing resources can significantly improve patient outcomes, as measured by patient satisfaction scores, readmission rates and frequency of hospital acquired infections (Aiken et al. 2002, 2010, 2014; Cimiotti, Aiken, Sloane and Wu 2012; McHugh and Ma 2013; McHugh, Berez and Small 2013; Needleman et al. 2011; Kutney-Lee et al. 2009). But from a business perspective, whether an investment in nursing generates positive returns still remains uncertain (Needleman, Buerhaus, Stewart, Zelevinsky and Mattke 2006). Considering that nursing labor costs account for as much as 40% of a hospital's inpatient expenses, increasing nurse staffing without matching increases in revenues often incurs significant costs. Another business strategy is to focus on improving the work environments of nurses. A large body of evidence demonstrates that patient outcomes are significantly better in hospitals where nurses have managerial support, greater autonomy, healthy relationships with physicians and adequate resources (Kutney-Lee et al. 2009; McHugh et al. 2013; Kutney-Lee, Wu, Sloane and Aiken 2013; Kelly, Kutney-Lee, Lake, Aiken 2013; Aiken et al. 2011a, 2011b; Aiken, Clarke, Sloane,

Lake and Cheney 2008; Friese, Lake, Aiken, Silber and Sochalski 2008; Lake 2007). The Institute of Medicine's report Keeping Patients Safe highlighted that a satisfactory work environment consequently encourages nurses to better fulfill the critical tasks that are assigned to them, including bedside care, accident prevention, care coordination and patient education (IOM 2003).

Research Question

This project will attempt to corroborate the relationship between nurse work environment and patient outcomes. The research question is as follows: To what extent do nurse work environments affect a hospital's performance? To answer this question, the paper will use a sophisticated pre-processing method to create a comparable set of observations.

DATA

Variables

Nurse work environment

The Practice Environment Scale of the Nursing Work Index (PES-NWI) has been studied extensively and is generally considered to be an accurate representation of a hospital's nurse work environment (Lake 2002). It has been endorsed by the National Quality Forum as a quality measure (National Quality Forum 2004). The measure consists of scores on a 4-point Likert scale for 5 different subscales: Staffing and Resource Adequacy, Nurse Participation in Hospital Affairs, Nursing Foundations for Quality of Care, Collegial Nurse-Physician Relations and Nurse Manager Ability, Leadership and Support of Nurses. Given the reliability of the measure in representing the quality of nurse work environments, the PES-NWI is used as the predictor for this

paper. These data come from a survey of nurses fielded as part of the Multistate Nursing Care and Patient Safety Study (Aiken et al.2011).

Patient outcomes

Many different metrics are used to estimate hospital performance. The measure used in this paper are 30-day inpatient mortality for surgical patients. This measure is appealing for two reasons. First, mortality rates generally have very little room for ambiguity, ensuring consistency across hospitals. Second, there are established approaches for risk-adjustment in order to account for the differences in patient severity of illness and case mix so that these differences do not confound the findings (Silber et al. 2007, 2009). Data on patient outcomes will be from hospital discharge databases for all adult, non-federal, acute care hospitals from the four states (CA, FL, NJ, PA) where the nurse survey was conducted.

Hospital characteristics

In order to use matching analysis, data for each hospital's structural characteristics are required. The following lists hospital characteristics that are used in the study: hospital size, teaching status, geographic location, patient population characteristics, market competition, and profit vs. non-profit status. Data on hospital structural characteristics will be from the American Hospital Association Annual Survey and hospital Medicare Cost Reports. Dr. McHugh, who supervised the completion of this project, has made the data sets available and has extensive experience working with these data.

METHODOLOGY

Traditional Methods

In research, researchers often want to estimate the causal effect of a treatment. For example, one may want to study the effect of a smoking cessation program on quit rates among participants. In this case, the treatment is participation in the program and the outcome is quit rate. One might also want to estimate the effect of unions on wages, the effects of compulsory school attendance on academic performance or the effect of attending college on future income.

In any of these studies, the preferred method to estimate causal effect is a randomized experiment. Randomization is an ideal method for research because it guarantees the absence of two major biases. The first one is selection bias, which occurs when a certain group is selected from a population. The second is omitted variable bias, which occurs when those who receive treatment and those who do not are dissimilar on average. Unfortunately, in social science research, it is difficult to achieve complete randomization. For instance, in the smoking cessation example, there is no ethical or legal way to randomly select people from the population and randomly assign them to the program (Ho, Imai, King and Stuart 2007).

Because of such limitations with conducting randomized experiments, social science research frequently relies on observational data. These studies analyze pre-existing data from events that have happened in the past. However, for any given set of observational data, it is unrealistic to expect that observations in the treated group will be similar to those in the control group, a condition that is met in randomized experiments. Therefore, traditional regression approaches with observational data often result in biased conclusions. In order to avoid this problem, researchers use a pre-processing method called matching.

Matching

Matching attempts to “simulate” a randomized experiment for observational data by pairing up each member of the treated group with a member from the control group that is most similar to it (Austin 2011, 402). This allows us to make an apples to apples comparison, since the confounding effects of pretreatment variables are controlled for and the imbalance between the treated and control groups is reduced. The benefits of matching is compounded by the fact that it can be done with a single estimated measure called the propensity score. The propensity score is the probability that an observation is assigned to treatment conditional on observed covariates (Rosenbaum and Rubin 1983). By projecting any number of covariates, variables that describe an observation, into a single scalar measure, propensity scores make matching easier, since the matching algorithm only has to minimize the sum difference of propensity scores between matched pairs rather than take into account an entire array of covariates. Research has widely shown that conditional on the propensity score, the distribution of covariates is similar between treated and untreated subjects (Austin 2011, 402). In addition, because a balance in propensity score guarantees a balance in covariates, only the distribution of propensity scores between the treated and control groups have to be compared to assess the quality of matches.

Application of Matching

The application of matching of this paper is a two-step process. First, ‘good’ hospitals are matched with ‘bad’ hospitals (a ‘good’ hospital being one with a high PES-NWI) using what’s called nonbipartite matching. Second, for each matched pair of hospitals, the patients of the two hospitals are matched using mixed integer programming with an R package called mipmatch.

Nonbipartite Matching of Hospitals

In bipartite matching, the treatment is a binary variable. For example, a study involving the effectiveness of smoking cessation programs will use bipartite matching, since one either does or does not participate in the program. In this paper, however, the treatment is the PES-NWI, a continuous variable. Therefore, nonbipartite matching is used to match hospitals. In nonbipartite matching, the treatment is first classified into quantiles. Then, an ordinal logit model is used to estimate a propensity score for each hospital. Finally, units that belong in different quantiles are matched. As diagram 1 illustrates, the PES-NWI is divided into five quantiles, each containing 20% of the hospitals and hospitals that belong in different quantiles are matched based on their propensity scores. The goal of the matching algorithm is to minimize the differences in propensity scores between each matched pair and maximize the difference of the PES-NWI. In order to achieve this, the following formula is used to create a distance matrix:

$$Distance = \frac{((p1 - p2)^2 + 0.001)}{(l1 - l2)^2}$$

p1 and *p2* are the propensity scores of each hospital and *l1* and *l2* are their respective PES-NWI quantiles. Note that distance decreases (i.e. hospitals are more likely to be matched) as difference of propensity scores decreases and as difference of PES-NWI increases. Also note that when *l1 = l2*, the distance is infinite, which prevents any hospitals within the same quantile from being matched.

Mipmatch of Patients

For each matched pair of hospitals, the hospital with the higher PES-NWI is the 'good' hospital and its matched counterpart is the 'bad' hospital. The next step is to match patients in the 'good' hospital with those in the 'bad' hospital. Note that only patients from matched hospitals can be matched. This paper uses an R package called mipmatch, which conducts matching using mixed integer programming. Mipmatch

was chosen because it allows a great degree of flexibility in choosing the specifications for matching. For example, instead of only using the propensity score, the matching algorithm also takes into account other patient covariates, such as comorbidity and surgical groups. Once patients are matched, the difference of 30 day mortality between patients in the ‘good’ hospital and those in the ‘bad hospital’ is calculated and its statistical significance is assessed.

RESULTS

The distribution of propensity scores of the nonbipartite matching of hospitals is shown in plot 1. As shown, the distribution of propensity scores for the ‘good’ hospitals resembles that of the ‘bad’ hospitals. The distribution of propensity scores of matched patients is shown in plot 2. When compared to the distributions before matching, the post-matching distribution of propensity scores clearly has a better balance. In other words, a valid comparison of these two groups of patients can be made. A t-test is conducted to test the hypothesis that the true mean of 30 day mortality is different for patients in ‘good’ hospitals and those in ‘bad’ hospitals. The calculations and results have been attached to the end of this paper. The statistical difference of outcomes between patients in “good” hospitals and patients in “bad” hospitals, which had a p-value of 0.34, was not significant.

DISCUSSION

The ultimate goal of the two-step matching process is to isolate the effect of nurse work environment on patient outcomes. Because matching produces hospitals that are different in their PES-NWI, which is indicative of the quality of the hospital’s nurse work environment, but are otherwise similar, the risk of comparing patient outcomes

of two very different hospitals is eliminated. Second, by matching patients with similar DRGs, comorbidity and surgical groups, the risk of comparing the outcomes of two patients with very different characteristics is avoided. The quality of matches was assessed by comparing the distributions of propensity scores before and after matching. Plots 1 and 2 illustrate the balanced propensity scores desired.

After matches had been made, the patient outcomes of matched patients were compared in order to assess statistical significance. In this study, the statistical difference of 30 day mortality rates between patients in "good" hospitals and those in "bad" hospitals had a p-value of 0.34 and therefore, was not significant. There may be multiple interpretations of this result. One may conclude that the effect of the nurse work environment on this particular patient outcome is not statistically significant. The validity of this conclusion is challenged by previous research that has repeatedly shown that better nurse work environments is associated with better hospital performance on a variety of measures, including patient satisfaction, mortality, failure to rescue and numerous patient safety indicators (McHugh, Berez and Small 2013). Another interpretation may look towards the limitations of this study. In this paper, two of them will be mentioned.

Limitations

One limitation is manifested in the data set for patient outcomes. First, there were numerous missing entries in the data set from which the 30 day mortality rates were obtained. Although they were ignored in the analysis, this may have introduced biases in the results and inaccuracies in the calculation. The problem is compounded by the fact that the proportion of patients that experienced 30 day mortality was small, less than 1%. Therefore, even if a small number of those missing entries turned out to be patients that experienced 30 day mortality, it could have a significant influence on the

proportion of 30 day mortality patients. Also, although failure to rescue rates were originally planned to be included in the patient outcome analysis, due to technical difficulties, they were not included. Inclusion of this variable in the analysis may have changed the resulting conclusions. Another limitation was with computing power. The data set of patients contains over 600,000 entries, far more than what can be reasonably handled by a personal computer for matching analysis. In order to overcome this problem, this paper implements a couple of circumventions that reduces computational complexity. For instance, the matching algorithm was designed so that only patients with the same DRGs are matched. This reduces computing time since it significantly reduces the number of attempts at matching patients. However, it may also have introduced biases that could have affected the analysis.

CONCLUSION

Numerous studies have used traditional regression methods to predict hospital performance using predictors such as staffing levels, teaching status and work environments. This project uses a more sophisticated statistical approach to help corroborate or refute conclusions drawn from such studies. The process of matching has allowed us to obtain groups of comparable patients. Each pair of matched patients received treatment from similar hospitals and are similar to each other in terms of the available patient data. The goal was to identify differences in the medical outcomes of these patients and if any difference does exist, it would be reasonable to attribute it to differences in the PES-NWI of the hospitals. The result showed that this relationship was not statistically significant. Considering previous literature on this topic corroborating the relationship between nurse work environments and patient

outcomes, it is possible that limitations in the data sets as well as computational power introduced inaccuracies in the results. Further review of the code and specifications of the matching algorithm may identify the underlying problem. This will also serve as a motivation for further analysis that utilizes better data sets with minimal error and to develop optimized matching algorithms that can process the data more efficiently.

References

Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA.* Oct 23-30 2002;288(16):1987-1993.

Aiken LH, Clarke SP, Sloane DM, Lake ET, Cheney T. Effects of hospital care environment on patient mortality and nurse outcomes. *J. Nurs. Adm.* May 2008;38(5):223-229.

Aiken LH, Sloane DM, Cimiotti JP, et al. Implications of the California nurse staffing mandate for other states. *Health Serv. Res.* 2010;45(4):904-921.

Aiken LH, Cimiotti JP, Sloane DM, Smith HL, Flynn L, Neff DF. Effects of nurse staffing and nurse education on patient deaths in hospitals with different nurse work environments. *Med. Care.* 2011;49(12):1047-1053.

Aiken LH, Sloane DM, Clarke S, et al. Importance of work environments on hospital outcomes in nine countries. *Int. J. Qual. Health Care.* May 11 2011.

Aiken LH, Sloane DM, Bruyneel L, et al. Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. *Lancet.* 2014;383(9931):1824-1830.

Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Psychology Press.* 2011;46:399-424.

Cimiotti JP, Aiken LH, Sloane DM, Wu ES. Nurse staffing, burnout, and health care–associated infection. *Am. J. Infect. Control.* 2012;40(6):486-490.

Friese CR, Lake ET, Aiken LH, Silber JH, Sochalski J. Hospital nurse practice environments and outcomes for surgical oncology patients. *Health Serv. Res.* 2008;43(4):1145–1163.

Ho DE, Imai K, King G and Stuart EA. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis.* 2007;15:199-236.

Institute of Medicine. *Keeping Patients Safe: Transforming the Work Environment of Nurses.* Washington, DC: National Academies Press; 2003.

Kelly D, Kutney-Lee A, Lake ET, Aiken LH. The critical care work environment and nurse-reported health care–associated infections. *Am. J. Crit. Care.* 2013;22(6):482-488.

Kutney-Lee A, McHugh MD, Sloane DM, et al. Nursing: a key to patient satisfaction. *Health Aff.* Jun 12 2009;28(4):w669-w677.

Kutney-Lee A, Wu ES, Sloane DM, Aiken LH. Changes in hospital nurse work environments and nurse job outcomes: An analysis of panel data. *Int. J. Nurs. Stud.* 2013;50(2):195-201.

Lake ET. Development of the Practice Environment Scale of the Nursing Work Index. *Res. Nurs. Health.* Jun 2002;25(3):176-188.

Lake ET. The nursing practice environment: measurement and evidence. *Med. Care Res. Rev.* Apr 2007;64(2 Suppl):104S-122S.

Lu B, Greevy R, Xu X, Beck C. Optimal nonbipartite matching and its statistical applications. *The American Statistician.* 2011;65(1):21-30.

McHugh MD, Berez J, Small DS. Hospitals with higher nurse staffing had lower odds of readmissions penalties than hospitals with lower staffing. *Health Aff.* 2013;32(10):1740-1747.

McHugh MD, Kelly LA, Smith HL, Wu ES, Vanak J, Aiken LH. Lower mortality in Magnet hospitals. *Med. Care.* 2013;51(5):382-388.

McHugh MD, Ma T. Hospital nursing and 30-day readmissions among Medicare patients with heart failure, acute myocardial infarction, and pneumonia *Med. Care.* 2013;51(1):52-59.

National Quality Forum. *National Voluntary Consensus Standards for Nursing-Sensitive Care: An Initial Performance Measure Set.* Washington D.C.: National Quality Forum;2004. NQFCR-08-04.

Needleman J, Buerhaus P, Pankratz VS, Leibson CL, Stevens SR, Harris M. Nurse staffing and inpatient hospital mortality. *N. Engl. J. Med.* 2011;364(11):1037-1045.

Needleman J, Buerhaus PI, Stewart M, Zelevinsky K, Mattke S. Nurse staffing in hospitals: is there a business case for quality? *Health Aff.* Jan-Feb 2006;25(1):204-211.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70, 41–55.

Silber JH, Romano PS, Rosen AK, Wang Y, Even-Shoshan O, Volpp KG. Failure-to-rescue: comparing definitions to measure quality of care. *Med. Care.* Oct 2007;45(10):918-925.

Silber JH, Rosenbaum PR, Romano PS, et al. Hospital teaching intensity, patient race, and surgical outcomes. *Arch. Surg.* Feb 2009;144(2):113-120.

Tanenbaum SJ. Pay for performance in Medicare: evidentiary irony and the politics of value. *J. Health Polit. Policy Law.* Oct 2009;34(5):717-746.

Diagrams and Plots

Diagram 1. Match hospitals using non-bipartite matching

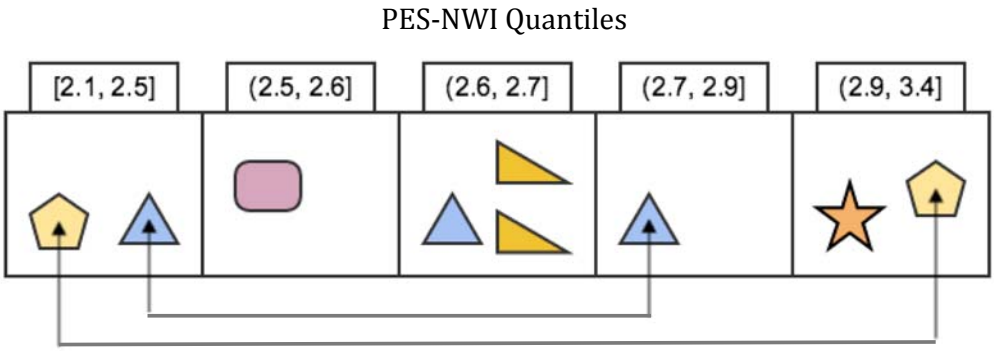
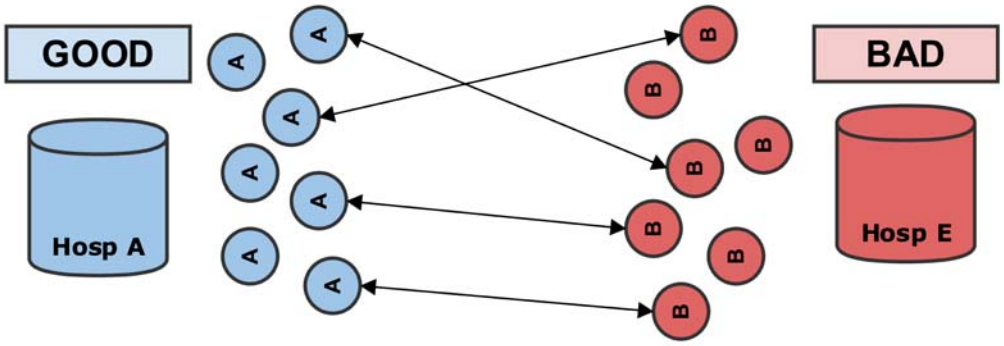
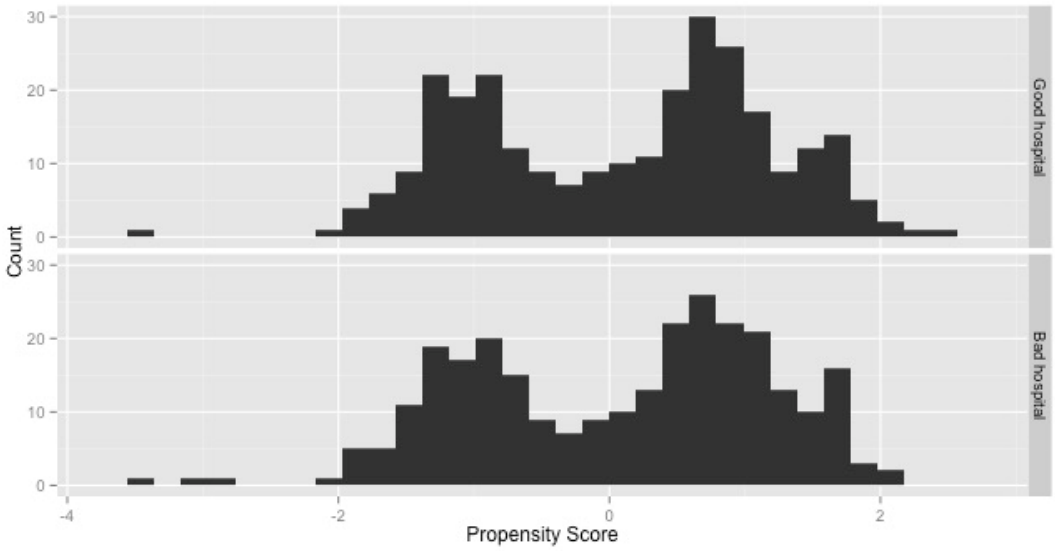


Diagram 2. Match patients for each matched pair of hospitals using mipmatch

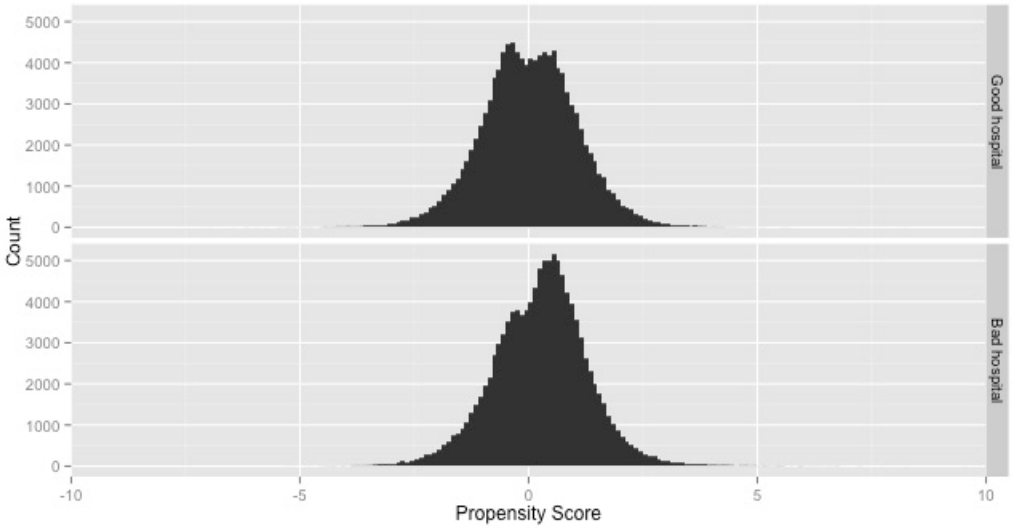


Plot 1. Distribution of propensity scores after nonbipartite matching of hospitals

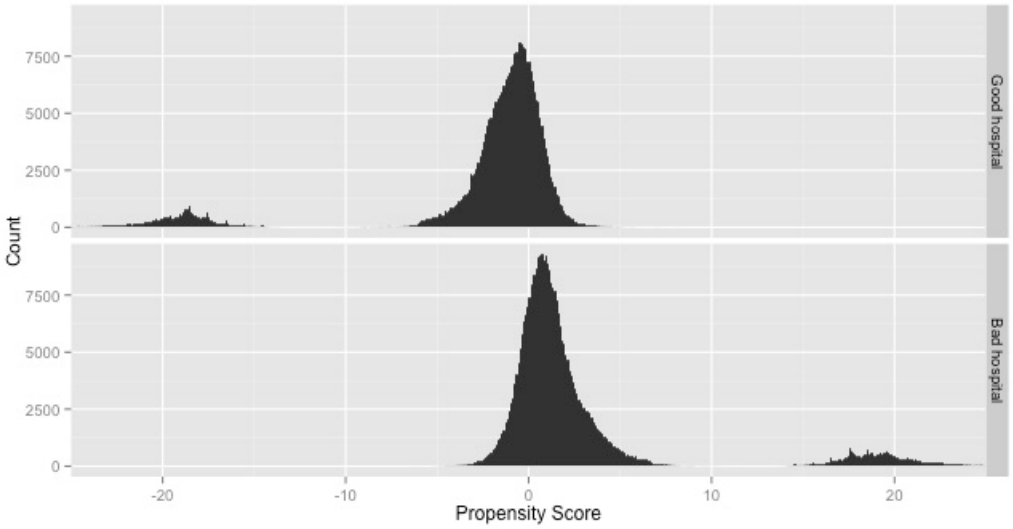


Plot 2. Distribution of propensity scores after mipmatch matching of patients

After Matching



Before Matching



Calculations and R Code

t-test for 30 Day Mortality

$$p1 = proportion\ of\ 'good'\ hospital\ patients\ with\ 30\ day\ mortality$$

$$= 0.009597469$$

$$p2 = proportion\ of\ 'bad'\ hospital\ patients\ with\ 30\ day\ mortality$$

$$= 0.01023027$$

$$n1 = number\ of\ 'good'\ hospital\ patients = 113780$$

$$n1 = number\ of\ 'bad'\ hospital\ patients = 113780$$

$$p\ = pooled\ sample\ proportion = \frac{(p1\ *\ n1\ +\ p2\ *\ n2)}{(n1\ +\ n2)}$$

$$= \frac{(0.0096\ *\ 113780\ +\ 0.01\ *\ 113780)}{(113780\ +\ 113780)} = 0.0098$$

$$SE\ = \sqrt{p\ *\ (\ 1\ -\ p\ )\ *\ (\frac{1}{n1}\ +\ \frac{1}{n2})}$$

$$= \sqrt{0.0098\ *\ (\ 1\ -\ 0.0098\ )\ *\ (\frac{1}{113780}\ +\ \frac{1}{113780})} = 0.0004$$

$$Null\ Hypothesis: P1\ -\ P2\ =\ 0$$

$$Alternative\ Hypothesis: P1\ -\ P2\ \neq\ 0$$

$$t\ = \frac{(p1\ -\ p2)}{SE} = \frac{(0.0096\ -\ 0.01\ )}{0.0004} = -0.9685$$

$$p - value = 0.34$$

Code for matching hospitals using Bo Lu's optimal nonbipartite matching R package (nbpMatching)

```
library(nbpMatching)
library(dplyr)
```

```
library(gridExtra)

# set wd
setwd('/Users/jaehyukhan/Desktop/CS/R/wd/HospitalMatching/nbp_matches')

# Use JMP for ordinal logistic regression
hosps <-
read.csv("/Users/jaehyukhan/Desktop/CS/R/wd/HospitalMatching/nbp_matches/hosp
s_data_ordlgt.csv")

################ nbpmatching #################
# create distance matrix
rn <- nrow(hosps)
dist_mat <- matrix(nrow=rn, ncol=rn)
for(row in seq(rn)) {
  for(row2 in row:rn) {
    p1 <- hosps$Linear[row]
    l1 <- hosps$pes_quant[row]
    p2 <- hosps$Linear[row2]
    l2 <- hosps$pes_quant[row2]
    dist <- as.numeric(((p1 - p2)^2 + 0.001) / (l1 - l2)^2)
    if(!is.finite(dist)) {
      dist <- 10^8
    } else {
      dist <- floor(dist * 10^6)
    }
    dist_mat[row, row2] <- dist
    dist_mat[row2, row] <- dist
  }
}

# Add phantoms
num_phantoms <- 6
dist_mat <- make.phantoms(dist_mat, num_phantoms)

cov <- hosps %>% select(magnet, final_beds, final_tech, ownership,
                        hhi, permedicare, permedicaid, teach, cbsadum, stdum)
ignored <- NULL
weights <- NULL
prevent <- NULL
mates <- NULL
rankcols <- NULL
missing.weight <- NULL
ndiscard <- 0
df.dist <- list(dist=dist_mat, cov=cov, ignored=ignored, weights=weights,
prevent=prevent,
                mates=mates, rankcols=rankcols, missing.weight=missing.weight,
ndiscard=ndiscard)

# create distancematrix object
```

```
df.mdm <- distancematrix(df.dist)
# create matches
df.match <- nonbimatch(df.mdm)
# review quality of matches
df.qom <- qom(df.dist$cov, df.match$matches)

n_mat <- nrow(df.match$halves)
m.combined_matches <- cbind(rbind(hosps[df.match$halves$Group1.Row,],
hosps[df.match$halves$Group2.Row,]), subclass=rep(seq(1, n_mat)), good=1)
rows_with_phantoms <- c()

for(row_num in seq(1, n_mat)) {
  p1 <- m.combined_matches[row_num,]$pes_quant
  p2 <- m.combined_matches[row_num+n_mat,]$pes_quant
  if(is.na(p1) | is.na(p2)) {
    m.combined_matches[row_num,]$good <- NA
    m.combined_matches[row_num+n_mat,]$good <- NA
    rows_with_phantoms <- c(rows_with_phantoms, row_num, row_num+n_mat)
    next
  }
  if(p1 > p2) {
    m.combined_matches[row_num+n_mat,]$good <- 0
  } else{
    m.combined_matches[row_num,]$good <- 0
  }
}

# remove rows with phantoms
m.combined_matches <- m.combined_matches[-rows_with_phantoms,]
m.combined_matches <- m.combined_matches[order(m.combined_matches$good,
decreasing = TRUE) , ]

# m.combined matches contains the following columns:
# [1] "Column.1"             "hospid"               "location"
# [4] "magnet"              "final_beds"           "final_tech"
# [7] "ownership"            "hhi"                 "permedicare"
# [10] "permedicaid"          "teach"               "cbsadum"
# [13] "stdum"               "pes"                 "pes_quant"
# [16] "Linear"              "Cum.1."               "Cum.2."
# [19] "Cum.3."              "Cum.4."               "Prob.1."
# [22] "Prob.2."             "Prob.3."             "Prob.4."
# [25] "Prob.5."             "Most.Likely.pes_quant" "Ord.Expected"
# [28] "subclass"            "good"
```

Code for matching patients using José R. Zubizarreta's mixed integer

programming R package (mipmatch)

```
library(mipmatch)

# read in data
pts <-
read.csv("/Users/jaehyukhan/Desktop/CS/R/wd/HospitalMatching/pts4match.csv")
pts_copy <- pts

# create drg variable
pts['drg'] <- factor(apply(pts %>% dplyr::select(drg110:drg537), 1,
                           function(x) which(x == 1)),
                     labels = colnames(pts %>% dplyr::select(drg110:drg537)))

# excluded sets
excluded = list(r3=c(113), r6=c(218, 263), r19=c(113, 493), r23=c(150), r32=c(209),
                r42=c(148), r78=c(209), r86=c(110), r101=c(209), r126=c(292),
                r128=c(216), r149=c(209), r152=c(503))

# create empty vectors of record_ids
treated_ids = c()
control_ids = c()

# Cut patient data by matched hospitals
for(row in seq(1, nrow(m.combined_matches) / 2)) {
  # Get patient list for matched pair of hospitals
  hosp1 <- m.combined_matches[row,]
  hosp_id1 <- hosp1$hospid
  subclass1 <- hosp1$subclass
  pts1 <- pts[pts$hospid == hosp_id1,]

  hosp2 <- m.combined_matches[m.combined_matches$subclass == subclass1,][2,]
  hosp_id2 <- hosp2$hospid
  subclass2 <- hosp2$subclass
  pts2 <- pts[pts$hospid == hosp_id2,]

  # Combine 'treated' and 'control' patients and add 'good'
  combined_pts <- cbind(rbind(pts1, pts2), good=c(rep(hosp1$good, nrow(pts1)),
rep(hosp2$good, nrow(pts2))))

  # add 'prop_score' and select necessary variables
  combined_pts_for_reg <- combined_pts %>% dplyr::select(age:htn_c, drg, good)
  glm.out <- glm(good ~ ., na.action=na.exclude, family=binomial(logit),
data=combined_pts_for_reg)
  combined_pts$prop_score <- predict(glm.out)
  combined_pts <- combined_pts %>% dplyr::select(record_id:htn_c,
d30hosp:prop_score)
```

```
  # identify drg common to pts1 and pts2
  common_drgs <- Reduce(intersect, list(pts1$drg,pts2$drg))

  # find number of rows for each drg
  drg_rows <- data.frame(drg_id=c(0), num_row=c(0))
  for(drg in common_drgs) {
    drg_rows <- rbind(drg_rows, c(drg, sum(pts1$drg == drg, pts2$drg == drg)))
  }
  drg_rows <- drg_rows[2:nrow(drg_rows),]

  for(drg in common_drgs) {
    # if in excluded list, skip
    if(as.numeric(gsub("drg", "", drg)) %in% excluded[[paste('r', row, sep="")]]) {
      next
    }

    num_row <- as.numeric(drg_rows[drg_rows[,1]==drg, 2])

    # select pts with this drg
    reduced_pts <- combined_pts[combined_pts$drg == drg,] %>% dplyr::select(-drg)

    # if number of treated > number of control, switch the two
    switched <- FALSE
    if(sum(reduced_pts$good) > sum(reduced_pts$good == 0)) {
      reduced_pts$good <- reduced_pts$good * -1 + 1
      switched <- TRUE
      # print("Switched 'good'")
    }

    # if too little treated, too many treated compared to total or too small of a pool, skip
matching
    if(sum(reduced_pts$good) < 2 || sum(reduced_pts$good) / num_row < 0.05 ||
sum(reduced_pts$good) / num_row > 0.45 || num_row < 10) {
      # print(paste("Skipped bc too little treated", drg))
      next
    }

    # if all prop_score is neg, skip matching
    # Error: Error in (c_index + n_t)[(1:length(out$xopt[-aux])) * out$xopt[-aux]] :
only 0's may be mixed with negative subscripts
    if(sum(reduced_pts$prop_score >= 0, na.rm=TRUE) < 1) {
      # print(paste("Skipped bc of neg prop scores", drg))
      next
    }

    ################ mipmatch #################
    # IMPORTANT: allmatch needs the data to be sorted in decreasing order by the
treatment indicator
    reduced_pts <- reduced_pts[order(reduced_pts$good, decreasing = TRUE) , ]
```

```
# Treatment indicator
t_ind <- reduced_pts$good

# Matrix of covariates
X_mat <- reduced_pts %>% dplyr::select(age:htn_c)
calip_cov <- reduced_pts$prop_score
calip_size <- 0.2*sd(calip_cov)
calip_penalty <- 2
# near_exact_covs <- cbind(final_tech, ownership, teach, cbsadum, stdum)
# near_exact_penalties <- c(2,1.5,1.5)
digits <- 6

# Distance matrix
dist_mat <- distmat(t_ind, X_mat, calip_cov, calip_size, calip_penalty, digits=6)

# Number of matches
n_matches = 1

# Moment covariates:
mom_covs = NULL
# Weights for the moment covariates
mom_weights = NULL
# Tolerances for the moment covariates
mom_tols = NULL

# Kolmogorov-Smirnov covariates
ks_covs = reduced_pts %>% dplyr::select(age)
# Number of grid points for the Kolmogorov-Smirnov statistic
ks_n_grid = 10
# Weights for the Kolmogorov-Smirnov covariates
ks_weights = c(1)
# Tolerances for the Kolmogorov-Smirnov covariates
ks_tols = NULL

# Covariates for near-exact matching, fine and near-fine balance
exact_covs = NULL
near_exact_covs = NULL
near_exact_devs = NULL
fine_covs = NULL
near_fine_covs = NULL
near_fine_devs = NULL

# Whether specific controls need to be used
use_controls_mat = NULL
use_controls_totals = NULL
use_controls_signs = NULL

# Enforce all the constraints
enforce_constraints = FALSE
```

```
   # Find matches
   out = allmatch(dist_mat, t_ind, n_matches,
                  mom_covs, mom_weights, mom_tols,
                  ks_covs, ks_n_grid, ks_weights, ks_tols,
                  exact_covs,
                  near_exact_covs, near_exact_devs,
                  fine_covs,
                  near_fine_covs, near_fine_devs,
                  use_controls_mat, use_controls_totals, use_controls_signs,
                  enforce_constraints)

   # Indices of the treated units and matched controls
   t_id = which(t_ind==1)
   c_id = out$c_id

   if(switched) {
     treated_ids <- c(treated_ids,reduced_pts[c_id,"record_id"])
     control_ids <- c(control_ids,reduced_pts[t_id,"record_id"])
   } else {
     treated_ids <- c(treated_ids,reduced_pts[t_id,"record_id"])
     control_ids <- c(control_ids,reduced_pts[c_id,"record_id"])
   }

   print(paste("length of treated=", length(treated_ids), " control=",
length(control_ids)))
#     # export summary to csv file
#     write.table(meantab(X_mat, t_ind, t_id, c_id, digits = 2), file = paste(drg,
"_summary.csv", sep=''), sep = ",",
#                 qmethod = "double", col.names=NA)
#
   #   # use following code to import
   #   # read.table("foo.csv", header = TRUE, sep = ",", row.names = 1)
  }
}

# create combined matches data set
m.combined_matches_pts <-
rbind(cbind(pts[treated_ids,],subclass=seq(length(treated_ids)), good=1),
                    cbind(pts[control_ids,],subclass=seq(length(treated_ids)), good=0))
```