American Journal of Epidemiology
© The Author(s) 2018. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.
Vol. 188, No. 2
DOI: 10.1093/aje/kwy252
Advance Access publication:
November 26, 2018

Practice of Epidemiology

Risk-Set Matching to Assess the Impact of Hospital-Acquired Bloodstream Infections

David Watson*, Alicen B. Spaulding, and Jill Dreyfus

* Correspondence to Dr. David Watson, Children's Minnesota Research Institute, Children's Hospitals and Clinics of Minnesota, 2525 Chicago Avenue South, Minneapolis, MN 55404 (e-mail: dave.watson2@gmail.com).

Initially submitted July 18, 2018; accepted for publication November 5, 2018.

Hospital-acquired bloodstream infections have a definite impact on patient encounters and cause increased length of stay, costs, and mortality. However, methods for estimating these effects are potentially biased, especially if the time of infection is not incorporated into the estimation strategy. We focused on matching patient encounters in which a hospital-acquired infection occurred to comparable encounters in which an infection did not occur. This matching strategy is susceptible to a selection bias because inpatients that stay longer in the hospital are more likely to acquire an infection and thus also are more likely to have longer and more costly stays. Instead, we have proposed risk-set matching, which matches infected encounters to similar encounters still at risk for infection at the corresponding time of infection. Matching on the one-dimensional propensity score can create comparable pairs for a large number of characteristics; an analogous propensity score is described for risk-set matching. We have presented dramatically different estimates using these 2 approaches with data from a pediatric cohort from the Premier Healthcare Database, United States, 2009–2016. The results suggest that estimates that did not incorporate time of infection exaggerated the impact of hospital-acquired infections with regard to attributed length of stay and costs.

bloodstream infections; hospital-acquired infections; matching; propensity score

Abbreviation: CI, confidence interval; HAI, hospital-acquired infections.

Hospital-acquired infections (HAI), particularly bloodstream infections, are expensive and potentially life-threatening events. Various published estimates show that HAI lead to longer hospital stays, more expensive visits, and higher mortality. Much of this work has compared patient encounters in which HAI occurred with encounters without HAI, typically either matched on clinical and demographic characteristics or regression adjusted in order to ensure patients are comparable (1–4). If the number of characteristics on which to match is large, finding exact or even close matches is difficult, but tools like propensity scores are useful for reducing the dimension of the matching problem (5, 6).

However, a limitation of many of these studies is they do not account for the time during the hospitalization at which the HAI occurred. This shortcoming occurs because detailed information on time of HAI is often unavailable in large databases of inpatient encounters (i.e., claims data). Failure to account for time of infection might lead to a selection bias because inpatients that stay longer in the hospital are more likely to acquire HAI and thus also are more likely to have longer and more costly stays. The aforementioned matching methods that attempt to control for the severity of a patient's condition cannot completely remove this bias.

If information on the time of infection is known, then a reasonable methodological variation is to match a patient who experiences a HAI event to comparable inpatients still at risk of infection at the time of that infection. This approach is called risk-set matching and effectively minimizes the selection bias (7). Additionally, this approach has an analogous propensity score that can produce well-matched cohorts for a large number of potential confounders, as well as time-varying factors (8). For this application of risk-set matching, the propensity score is estimated via Cox proportional hazards regression on the time to infection.

The purposes of this work were to: 1) highlight the severity of the selection bias that can occur when time of infection is

Downloaded from https://academic.oup.com/aje/article-abstract/188/2/461/5210259 by Karolinska Institutet user on 06 February 2019

left unaccounted for, and 2) to demonstrate risk-set matching as a tool to produce a well-designed study to assess the impact attributable to HAI.

METHODS

Data source and variables of interest

Data for this study come from the Premier Healthcare Database, which represents approximately 20% of annual discharges in the United States and includes over 6 million hospital discharges per year from more than 750 hospitals (9). This service-level, all-payer, administrative database contains detailed information on patients' demographics, diagnoses, treatments, and costs as well as hospital characteristics related to their inpatient stays. All data in the Premier Healthcare Database are statistically de-identified and compliant with the Health Insurance Portability and Accountability Act (HIPAA). This study was determined to be exempt from review by the institutional review board at Children's Hospitals and Clinics of Minnesota.

For this study, the specific population of interest was pediatric (aged less than 19 years) inpatient encounters from 2009 to 2016. Neonates were excluded from the analysis because of the difficulty in clearly defining HAI for this high-risk population. Only encounters with a length of stay over 2 days were included, in order to ensure all patient encounters were at risk of HAI. The study was limited to 160 hospitals that reported data from their microbiology labs during the study period. Of particular interest for this study was data on blood cultures, the gold standard for diagnosing bloodstream infections.

Several potential confounding factors are related to both risk of HAI and the outcomes of interest. The potential confounding variables included patient demographics (i.e., age, sex, race, and insurance type), provider characteristics, clinical description of the patient, geographic region (i.e., 9 census divisions), and year of encounter. Descriptions of the providers included whether they were pediatric specialists as well as information on the hospital size (i.e., number of beds), location (urban or rural), and teaching status. Clinical characteristics included a description of the primary diagnosis (classified into 25 Major Diagnostic Categories provided by the Center for Medicaid and Medicare Services), presence of a pediatric complex chronic disease (as previously defined by Feudtner et al. (10)), and treatments known to be associated with infections. Treatments of interest for this analysis included whether the patient was in the intensive care unit or had a central line and/or catheter in place. These features were identified as treatments present or occurring on admission based on charge codes and/or *International Classification of Diseases* (*Ninth* or *Tenth Revision*, as applicable) procedure codes recorded within the first 2 days of admission. Treatments were limited to the first 2 days of the encounter because they could be determined to have occurred prior to any potential HAI, and not after (or as a result of) the infection.

From the microbiology data, a laboratory-confirmed bloodstream infection was identified via a positive blood culture with the isolate being a known organism and not a common commensal according to the list published by the Centers for Disease Control and Prevention's National Healthcare Safety Network (11, 12). For this study, to be considered a HAI event, the first positive blood culture had to be drawn on day 3 or later of the hospital admission without the patient having a primary diagnosis of an infectious disease present on admission. This definition differs from the National Healthcare Safety Network's definition, which includes some diagnoses that do not require a positive blood culture. Time of infection was defined as the day the first positive blood culture was drawn. This definition might overestimate the time of infection, which naturally had to have occurred before the blood draw.

The outcomes of interest for this study were length of stay, cost, and mortality. Total costs for the hospitalization included all costs recorded by the hospital during the inpatient encounter (e.g., room and board, pharmacy, laboratory, etc). All costs were inflation adjusted to (December) 2016 US dollars using the US Department of Labor Consumer Price Index (13) and Winsorized at the 0.1 and 99.9 percentiles; encounters with $0 costs were excluded. Mortality was determined from discharge status classified as expired or discharged to hospice.

Conventional matching with propensity scores

A matched cohort of patients with and without HAI was constructed based on propensity scores. The propensity score was estimated with a logistic regression model that had HAI status as the dependent variable and potential confounding factors as the independent variables (14). We emphasize that the propensity score model does not incorporate the time of infection; it models only the occurrence of HAI. Moreover, all confounding variables are based on information available at time of admission (or within the first 2 days), and these variables do not change over time.

In total, the propensity score model included over 70 independent variables. The theory of propensity scores shows that if a HAI patient and non-HAI patient have the same value of the propensity score, then these 2 patients are comparable, on average, with respect to the variables in the propensity score. Thus, instead of trying to match on over 70 variables at the same time, the matching problem is reduced to matching on only 1 variable (or dimension): the propensity score. Patient encounters with HAI were matched to 3 encounters without HAI based on nearest-neighbor matching on the propensity score (on the log odds scale). We refer to patients matched on the logistic propensity score as the "conventionally matched cohort," because it stems from standard applications of matched designs.

Risk-set matching with propensity scores

For risk-set matching, a patient experiencing a HAI event on a specific day is matched to similar patients who have not experienced HAI up to that point in their hospital admission. The idea, as the name suggests, is to match HAI patients to comparable patients still at risk of HAI. Multivariate matching on a large number of confounding factors is even more difficult in this setting because the set at risk of HAI continually decreases (i.e., as patients are discharged). However, analogous propensity score methods effectively reduce the dimension of the matching variable to one.

For risk-set matching, the propensity score is the hazard of HAI occurring as a function of all confounding variables. Heuristically, just as the propensity score is the probability of HAI,

Downloaded from https://academic.oup.com/aje/article-abstract/188/2/461/5210259 by Karolinska Institutet user on 06 February 2019

the propensity score for risk-set matching is the instantaneous probability of HAI given that a patient is at risk. Extending the analogy, just as logistic regression is one tool for modeling the propensity score, the risk-set propensity score can be estimated using standard survival analyses, namely Cox proportional hazards regression (15). Under certain assumptions, using the modeled hazard has similar balancing properties as the propensity scores. A rigorous treatment of this theory is beyond the scope of this work; however, see Lu (8) for theoretical details.

The risk-set propensity score is estimated with a Cox regression with the survival outcome of either HAI (i.e., an event) occurring and its corresponding time or HAI not occurring (i.e., a censored event) and the length of stay in the hospital (i.e., censoring time). The same independent variables used in the logistic regression propensity score model were included in the proportional hazards model, with the exception of treatments on admission. Instead, these variables were allowed to vary over the duration the patient was at risk during their hospital admission. Stay in an intensive care unit could be determined for every day of the encounter, however; the presence of a catheter or central line could be determined only from procedure or billing codes on the day of the procedure. As such, a patient who had a catheter or central line procedure code was assumed to have this in place for the remainder of the hospital stay.

The estimated risk-set propensity score is the linear prediction from the Cox regression model, which, when exponentiated, is proportional to the hazard. Each HAI event and time was matched to 3 patient encounters that were still at risk of HAI at the time of infection based on nearest-neighbor matching. We refer to patients matched on the risk-set propensity score as the "risk-set matched cohort."

A subtlety of risk-set matching is that a patient can be at risk at one point in their hospital stay, and thus matched to a HAI event, but eventually go on to experience a HAI event later in their stay. That is, an encounter could contribute as both an uninfected and infected patient.

Statistical analysis

The balance of the matched pairs design was assessed by comparing the distribution of a potential confounder between the patient with a HAI event and (matched) patients without a HAI event. A plot of the standardized difference in percentage for each covariate was used to assess of the degree of imbalance between the HAI cohort and non-HAI cohorts (16, 17). For variable x, the standardized difference is defined as:

$$\text{Percent standardized difference} = 100 \times \frac{\bar{x}_{\text{HAI}} - \bar{x}_{\text{no HAI(match)}}}{\sqrt{(s^2_{\text{HAI}} + s^2_{\text{no HAI}})/2}}$$

where $\bar{x}_{\text{HAI}}$ and $\bar{x}_{\text{no HAI(match)}}$ are the sample means of the HAI patients and matched non-HAI patients respectively and s^2_{HAI} and $s^2_{\text{no HAI}}$ are the sample variances of the HAI and all non-HAI patients. Values near zero suggest the distributions are well balanced, whereas values far from zero suggest imbalance. The magnitude of the standardized difference follows the interpretation offered for Cohen's d.

The outcomes of length of stay, costs, and mortality were compared between HAI and non-HAI patients. Matched pairs were analyzed using a paired t interval of the HAI patient's outcome and the average of the 3 matched patients' outcomes. Comparisons with all patients with no HAI are made for the sake of comparison; 2 sample t intervals were used for these inferences. In addition, a sensitivity analysis was performed in which the time of infection was defined as 2 days prior to the positive blood culture in order to account for 48 hour incubation period.

RESULTS

In total, 237,625 inpatient encounters from 160 hospitals were included in the analysis, with 374 (0.16%) encounters experiencing a laboratory-confirmed HAI. Web Figure 1 (available at https://academic.oup.com/aje) plots the standardized difference comparing HAI and non-HAI patients for all covariates. Before matching, there was severe imbalance on several factors, including whether the patient was in the intensive care unit or had a central line. After 3:1 conventional matching on the (logistic) propensity score, the balance was improved considerably. Risk-set matching also improved the balance relative to the no-HAI sample, but not as well as conventional matching. For example, patients in the HAI group were more likely to have hematologic conditions than were their risk-set matched counterparts (18% versus 8%), but these groups were similar in the conventional matched cohort (18% versus 18%).

Table 1 summarizes the distribution of time-varying factors for the different cohorts. Among the encounters with no HAI, only 34% were even at risk when averaging over the times of infection. Conventional matching increased the percentage at risk but only to 46%. That is, 54% of these conventionally matched patients had already been discharged by the time

Table 1. Percentages of Time-Varying Factors at the Time of Hospital-Acquired Infection Among Pediatric Encounters at Hospitals With Microbiology Data, Premier Healthcare Database, United States, 2009–2016

Variable	HAI (n = 374)	No HAI[a] (n = 237,251)	Conventionally Matched (n = 1,122)	Risk-Set Matched (n = 1,122)
At risk of HAI	100	34	46	100
Central line in place during admission	32	2	12	24
Catheter in place during admission	9	3	4	15
In intensive care unit	35	4	14	27

Abbreviation: HAI, hospital-acquired infections.
[a] Percent of all pairwise comparisons of HAI and no-HAI encounters.

Downloaded from https://academic.oup.com/aje/article-abstract/188/2/461/5210259 by Karolinska Institutet user on 06 February 2019

Table 2. Comparison of Outcomes Among Pediatric Encounters at Hospitals With Microbiology Data, Premier Healthcare Database, United States, 2009–2016

Outcome	HAI	No HAI			Conventionally Matched			Risk-Set Matched		
	Average	Average	Difference[a]	95% CI	Average	Difference[a]	95% CI	Average	Difference[a]	95% CI
Length of stay, days	43.5	7.3	36.2	28.1, 44.3	12.8	30.7	22.7, 38.7	31.5	11.9	6.1, 17.8
Cost, $ (thousands)	103.8	13.5	90.3	79.8, 100.8	37.4	66.4	56.6, 76.1	71.9	31.9	23.3, 40.4
Died, %	9.4	0.3	9.1	6.1, 12.1	1.3	8.0	5.0, 11.0	2.4	7.0	3.9, 10.0

Abbreviations: CI, confidence interval; HAI, hospital-acquired infections.
[a] Difference calculated as HAI cohort minus non-HAI cohort.

their match pair got infected in the hospital. In contrast, the risk-set matched cohort were all at risk by design. The time-varying covariates were also better balanced after risk-set matching than conventional matching. Of the matched encounters at risk of infection, 66 (5.9%) went on to incur HAI, and these encounters contribute to both the HAI and non-HAI groups.

Table 2 provides estimates of the impact of HAI on the outcomes of interest. For those conventionally matched, the estimated attributable length of stay was an additional 31 days on average (95% confidence interval (CI): 23, 39). This estimate is larger than the risk-set matched estimate of 12 additional days (95% CI: 6, 18). A similar result was obtained for costs to the hospital, with the conventional approach estimating an additional $66,400 (95% CI: $56,600, $76,100) due to HAI, whereas the risk-set matching estimated $31,900 (95% CI: $23,300, $40,400) in additional costs. The estimates for mortality are comparable, but the risk-set estimate suggests a slightly larger effect.

Results were similar for the sensitivity analysis that defined time of infection as 2 days prior to the first positive blood culture. For conventional matching, the additional length of stay due to HAI was 33 days (95% CI: 25, 40), and the additional costs were $60,000 (95% CI: $52,600, $67,400). These estimates were again much larger than estimates using risk-set matching, with an average of 16 additional days (95% CI: 10, 22) and $35,100 in additional costs to the hospital (95% CI: $28,900, $41,300) due to HAI.

DISCUSSION

The estimates of the impact attributable to HAI on length of stay and costs were dramatically larger using conventional matching compared with risk-set matching. Our explanation for this difference is the selection bias that can occur because longer encounters are more likely to incur HAI. The selection bias is made apparent when considering that 54% of the conventionally matched patients with no HAI were already discharged by the time their matched counterpart was infected. Risk-set matching reduces this selection bias by matching patients who are at risk for HAI (i.e., still in the hospital) at the time the infection occurred.

The selection bias due to HAI naturally occurring in longer hospital stays has been documented in other studies (18, 19). Most similar to our work is an analysis by Vrijens et al. (20), which showed large differences in estimates that account for the time of infection and those that do not. Even though the matched patients were at risk of infection, the approach of Vrijens (and others (21, 22)) is not quite risk-set matching because matched patients were confirmed to not have HAI. Recall, risk-set matching includes all patients still at risk of infection as potential matches at a given time, not just those that go on to have no infection. The reason for this approach is to match only on information from the past or present with respect to the time of infection. Matching on future information like whether HAI occurs can introduce bias analogous to the bias found in survival analyses that compare groups based on response at the end of a trial (23, 24).

As an example of how this bias can occur, consider finding a match for patient C in Figure 1. Risk-set matching uses all information up to day 5, when the infection occurred. Patient A was discharged on day 3 and thus ineligible, but patients B and D are both at risk and eligible matches. Excluding patient D because she went on to incur an infection on day 16 would bias the group at risk on day 5 to have favorable outcomes in the sense that patient D had longer length of stay and died. Moreover, patient D should also contribute to the HAI group because to exclude patient D (i.e., after being matched to patient B) would bias the HAI group to infections that occurred earlier.

The reason to match on propensity scores—for both conventional and risk-set strategies—is that it reduces the problem of

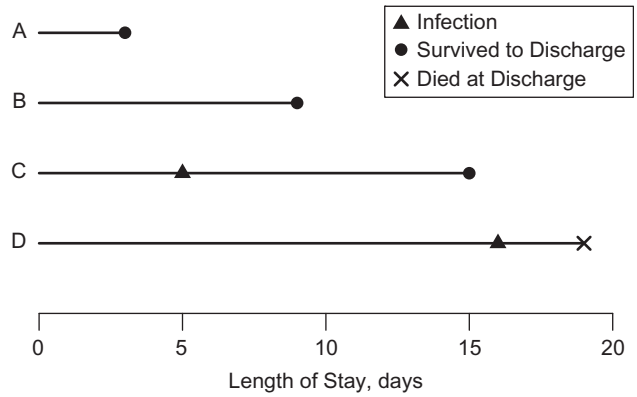

Figure 1. Diagram demonstrating risk-set matching for patient C; both patients B and D are eligible matches.

Downloaded from https://academic.oup.com/aje/article-abstract/188/2/461/5210259 by Karolinska Institutet user on 06 February 2019

matching on several variables down to only 1 variable. For example, Vrijens et al. required exact matching on all potential confounders, which led to including only half of HAI cases. This restriction on the sample size only to successfully matched cases might bias results (4). In this study, we performed risk-set matching on the propensity score, which was estimated via standard survival analysis techniques (i.e., Cox proportional hazards regression) on the time to infection. The theory of Lu (8) suggests that under certain assumptions the matched cohort is comparable to the HAI cohort with respect to any observed variables used in the propensity score model. Our results showed improvement in balance relative to the no-HAI cohort and suggest that the risk-set matched group was comparable to the HAI patients across several covariates.

Modeling time to infection introduces additional complications to propensity-score estimation for risk-set matching. Our model demonstrated how to incorporate time-varying covariates into the propensity score (25). This task is computationally more challenging than conventional propensity-score estimation and matching because each encounter is broken up into multiple disjoint time intervals over which the covariates are constant. Another potential complication is violations of the proportional hazards assumption, for example, if associations with the outcome change over time (26). Andersen et al. (27) offer an application of risk-set matching in which the time to tracheal intubation is modeled with time-varying coefficients.

Although risk-set matching improved balance of the baseline covariates, conventional matching produced a cohort with smaller (absolute) standardized differences. This better balance was because all patients without HAI were available to match whereas the risk set of potential matches dwindles as infections occur later in the stay. For example, the reservoir of encounters without HAI was over 600 times the number of HAI encounters. In contrast, 50% and 25% of HAI occurred, respectively, after days 8 and 16, which corresponded to at-risk reservoirs of only 217 and 138 times the number of remaining HAI encounters. Although these reservoir sizes are quite large, matching on rare conditions is still challenging. We view the additional imbalance of risk-set matching as inherent to the problem of finding comparable patients still at risk of infection as opposed to a disadvantage of the method.

More complex matching or inferential techniques could address additional imbalance from risk-set matching. Modeling outcome variables via mixed effects regressions or generalized estimating equations (to account for matched pairs) could adjust for any remaining imbalances as well as examine potential effect modification. Additionally, more advanced propensity-score matching algorithms would likely improve the balance between the cohorts (e.g., (28) and (29)), although there might be challenges to implementing them for risk-set matching. For example, a multivariate distance metric (i.e., Mahalanobis distance (30)) should change as the composition of the risk set changes over time. In this study, we chose simple matching and inferential techniques for clarity of exposition.

Both risk-set and conventional matching rely on observing all potential confounders. Neither can adjust for unobserved confounding. Sensitivity analysis can assess the impact of any unobserved confounding and determine to what degree an unobserved variable might explain the results (31). There is a rich literature on these techniques for observational studies as well as conventional matched-pair designs (32–34). See Li et al. (7) for a sensitivity analysis specific to risk-set matching.

Conventional matching still has its place when there is no clear equivalent of time to infection that is an obvious confounding factor. However, that scenario is not the same as when time of infection is unknown or undocumented, which is often the case for studies of HAI using large databases. With no information on the approximate time of infection, risk-set matching cannot be implemented and conventional matching is not recommended. Exploring how to proceed in the absence of time to infection is an area for future research.

In conclusion, the results of this work suggest that studies that do not take into account time of infection might overestimate the impact of HAI on outcomes. Most research on the impact of HAI does not incorporate time of infection (35), and there is a need to both develop and implement statistical methods such as risk-set matching that do incorporate time of infection. Moreover, the implications of this work go beyond infections. The selection bias shown here also applies to other hospital-associated outcomes that are more likely to occur in longer inpatient stays: for example, venous thromboembolism, pressure ulcers, or falls (6, 36). A broad reassessment of the impact of hospital-associated events might be in order.

ACKNOWLEDGMENTS

Author affiliations: Children's Minnesota Research Institute, Children's Hospitals and Clinics of Minnesota, Minneapolis, Minnesota (David Watson, Alicen B. Spaulding); and Premier Applied Sciences, Premier, Inc., Charlotte, North Carolina (Jill Dreyfus).

This study was supported by a grant from Children's Hospitals and Clinics of Minnesota.

Conflict of interest: J.D. is an employee and shareholder of Premier, Inc. The other authors report no conflicts.

REFERENCES

1. Orsi GB, Di Stefano L, Noah N. Hospital-acquired, laboratory-confirmed bloodstream infection: increased hospital stay and direct costs. *Infect Control Hosp Epidemiol*. 2002;23(4):190–197.
2. Payne NR, Carpenter JH, Badger GJ, et al. Marginal increase in cost and excess length of stay associated with nosocomial bloodstream infections in surviving very low birth weight infants. *Pediatrics*. 2004;114(2):348–355.
3. Elward AM, Hollenbeak CS, Warren DK, et al. Attributable cost of nosocomial primary bloodstream infection in pediatric intensive care unit patients. *Pediatrics*. 2005;115(4):868–872.
4. De Angelis G, Murthy A, Beyersmann J, et al. Estimating the impact of healthcare-associated infections on length of stay and costs. *Clin Microbiol Infect*. 2010;16(12):1729–1735.
5. Goudie A, Dynan L, Brady PW, et al. Attributable cost and length of stay for central line-associated bloodstream infections. *Pediatrics*. 2014;133(6):e1525–e1532.
6. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA*. 2003;290(14):1868–1874.

Downloaded from https://academic.oup.com/aje/article-abstract/188/2/461/5210259 by Karolinska Institutet user on 06 February 2019

7. Li YP, Propert KJ, Rosenbaum PR. Balanced risk set matching. *J Am Stat Assoc*. 2001;96(455):870–882.
8. Lu B. Propensity score matching with time-dependent covariates. *Biometrics*. 2005;61(3):721–728.
9. Premier Inc., Premier Applied Sciences. Premier Healthcare Database White Paper: Data That Informs and Performs. https://learn.premierinc.com/pharmacy-and-research/premier-healthcare-database-whitepaper. Published February 18, 2018. Accessed March 15, 2018.
10. Feudtner C, Feinstein JA, Zhong W, et al. Pediatric complex chronic conditions classification system version 2: updated for ICD-10 and complex medical technology dependence and transplantation. *BMC Pediatr*. 2014;14:199.
11. Centers for Disease Control and Prevention. Bloodstream Infection Event (Central Line-Associated Bloodstream Infection and Non-central Line Associated Bloodstream Infection). Atlanta, GA: Centers for Disease Control and Prevention; 2018. (National Healthcare Safety Network Patient Safety Manual) http://www.cdc.gov/nhsn/pdfs/pscmanual/4psc_clabscurrent.pdf. Accessed March 27, 2018.
12. Centers for Disease Control and Prevention, National Healthcare Safety Network. NHSN Organisms List. Atlanta, GA: Centers for Disease Control and Prevention; 2018. https://www.cdc.gov/nhsn/xls/2018-NHSN-Organisms-List-Validation.xlsx. Accessed March 27, 2018.
13. US Bureau of Labor Statistics. Division of Consumer Prices and Price Index. Consumer Price Index: All Urban Consumers (Current Series)—US Medical Care. https://beta.bls.gov/dataViewer/view/timeseries/CUUR0000SAM. Accessed April 11, 2018.
14. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997;127(8 Part 2):757–763.
15. Cox DR. Regression models and life-tables. In: Kotz S, Johnson NL, eds. *Breakthroughs in Statistics*. New York, NY: Springer-Verlag; 1992:527–541.
16. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat*. 1985;39(1):33–38.
17. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28(25):3083–3107.
18. Barnett AG, Beyersmann J, Allignol A, et al. The time-dependent bias and its effect on extra length of stay due to nosocomial infection. *Value Health*. 2011;14(2):381–386.
19. De Angelis G, Allignol A, Murthy A, et al. Multistate modelling to estimate the excess length of stay associated with meticillin-resistant *Staphylococcus aureus* colonisation and infection in surgical patients. *J Hosp Infect*. 2011;78(2):86–91.
20. Vrijens F, Hulstaert F, Van de Sande S, et al. Hospital-acquired, laboratory-confirmed bloodstream infections: linking national surveillance data to clinical and financial hospital data to estimate increased length of stay and healthcare costs. *J Hosp Infect*. 2010;75(3):158–162.
21. Kaye KS, Marchaim D, Chen TY, et al. Effect of nosocomial bloodstream infections on mortality, length of stay, and hospital costs in older adults. *J Am Geriatr Soc*. 2014;62(2):306–311.
22. Nelson RE, Samore MH, Jones M, et al. Reducing time-dependent bias in estimates of the attributable cost of health care-associated methicillin-resistant *Staphylococcus aureus* infections: a comparison of three estimation strategies. *Med Care*. 2015;53(9):827–834.
23. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. *J Clin Oncol*. 1983;1(11):710–719.
24. Buyse M, Piedbois P. On the relationship between response to treatment and survival time. *Stat Med*. 1996;15(24):2797–2812.
25. Fisher LD, Lin DY. Time-dependent covariates in the Cox proportional-hazards regression model. *Annu Rev Public Health*. 1999;20:145–157.
26. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81(3):515–526.
27. Andersen LW, Granfeldt A, Callaway CW, et al. Association between tracheal intubation during adult in-hospital cardiac arrest and survival. *JAMA*. 2017;317(5):494–506.
28. Diamond A, Sekhon JS. Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. *Rev Econ Stat*. 2013;95(3):932–945.
29. Zubizarreta JR. Using mixed integer programming for matching in an observational study of kidney failure after surgery. *J Am Stat Assoc*. 2012;107(500):1360–1371.
30. Mahalanobis PC. On the generalized distance in statistics. *Proc Natl Inst Sci India*. 1936;2(1):49–55.
31. Rosenbaum PR. Discussing hidden bias in observational studies. *Ann Intern Med*. 1991;115(11):901–905.
32. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. *Epidemiology*. 2016;27(3):368–377.
33. Robins JM, Rotnitzky A, Scharfstein DO. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: Halloran ME, Berry D, eds. *Statistical Models in Epidemiology, the Environment, and Clinical Trials*. New York, NY: Springer-Verlag; 2000:1–94.
34. Rosenbaum PR. Sensitivity analysis for certain permutation inferences in matched observational studies. *Biometrika*. 1987;74(1):13–26.
35. Manoukian S, Stewart S, Dancer S, et al. Estimating excess length of stay due to healthcare-associated infections: a systematic review and meta-analysis of statistical methodology. *J Hosp Infect*. 2018;100(2):222–235.
36. Goudie A, Dynan L, Brady PW, et al. Costs of venous thromboembolism, catheter-associated urinary tract infection, and pressure ulcer. *Pediatrics*. 2015;136(3):432–439.

Downloaded from https://academic.oup.com/aje/article-abstract/188/2/461/5210259 by Karolinska Institutet user on 06 February 2019