



Published in final edited form as:

*Infect Control Hosp Epidemiol.* 2016 April ; 37(4): 404–410. doi:10.1017/ice.2015.340.

## Hospital *Clostridium difficile* Infection Rates and Prediction of Length of Stay in Patients Without *C. difficile* Infection

Aaron C. Miller, PhD<sup>1</sup>, Linnea A. Polgreen, PhD<sup>2</sup>, Joseph E. Cavanaugh, PhD<sup>2</sup>, and Philip M. Polgreen, MD<sup>2</sup>

<sup>1</sup>Cornell College, Mount Vernon, Iowa

<sup>2</sup>University of Iowa, Iowa City, Iowa

### Abstract

**Background**—Inpatient length of stay (LOS) has been used as a measure of hospital quality and efficiency. Patients with *Clostridium difficile* infections (CDI) have longer LOS.

**Objective**—To describe the relationship between hospital CDI incidence and the LOS of patients without CDI.

**Design**—Retrospective cohort analysis.

**Methods**—We predicted average LOS for patients without CDI at both the hospital and patient level using hospital CDI incidence. We also controlled for hospital characteristics (eg, bed size) and patient characteristics (eg, comorbidities, age).

**Setting**—Healthcare Cost and Utilization Project Nationwide Inpatient Sample, 2009–2011.

**Patients**—The Nationwide Inpatient Sample includes patients from a 20% sample of all nonfederal US hospitals.

**Results**—Inpatient LOS was significantly longer ( $P < .001$ ) at hospitals with greater CDI incidence at both the hospital and individual level. At a hospital level, a percentage point increase in the CDI incidence rate was associated with more than an additional day's stay (between 1.19 and 1.61 days). At the individual level, controlling for all observable variables, a percentage point increase in the CDI incidence rate at their hospital was also associated with longer LOS (between 0.6 and 1.05 additional days). Hospital CDI incidence had a larger impact on LOS than many other commonly used predictors of LOS.

**Conclusion**—CDI rates are a predictor of LOS in patients without CDI at an individual and institutional level. CDI rates are easy to measure and report and thus may provide an important marker for hospital efficiency and/or quality.

---

Address correspondence to Philip M. Polgreen, MD, University of Iowa Hospitals and Clinics, 200 Hawkins Dr, Iowa City, IA, 52242 (philip-polgreen@uiowa.edu).

*Potential conflicts of interest.* All authors report no conflicts of interest relevant to this article.

Supplementary Material

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/ice.2015.340>

Hospital length of stay (LOS) is an important contributor to healthcare expenditures.<sup>1</sup> Increased LOS is also a risk factor for adverse events.<sup>2,3</sup> Moreover, many factors used to measure healthcare quality have been linked to prolonged LOS,<sup>4-7</sup> and LOS has also been used to measure quality.<sup>8-10</sup> In addition, LOS has been used to measure efficiency in hospitals.<sup>11,12</sup> For these reasons, LOS is commonly used to study disease outcomes. However, LOS varies dramatically, not only for different procedures and diagnoses, but also among hospitals.<sup>13</sup> Thus, when using LOS as an outcome measure, adjusting for factors associated with hospital-level variation in LOS is important.

Many studies have analyzed excess LOS associated with adverse events, including postoperative hemorrhage or hematoma,<sup>5</sup> falls,<sup>14,15</sup> adverse drug events,<sup>16</sup> and decubitus ulcers.<sup>17</sup> Healthcare-associated infections are also a frequently studied source of excess LOS. Examples include bloodstream infections,<sup>18</sup> methicillin-resistant *Staphylococcus aureus* infection,<sup>19,20</sup> sepsis,<sup>21</sup> and surgical site infections.<sup>22</sup> A common healthcare-associated infection is *Clostridium difficile* infection (CDI), and CDI increases the LOS of patients with the disease.<sup>23-25</sup>

Although CDI may increase an infected patient's LOS, it is not known whether institution-level CDI is related to prolonged LOS in patients without CDI. To our knowledge, no study has described the association between CDI incidence and hospital-wide excess LOS in patients without CDI. However, a number of possible links exist between CDI incidence and excess LOS in patients without CDI. For example, hospital CDI rates may be a proxy for quality. Poor quality hospitals may foster more CDI and create conditions that lead to excess LOS (eg, other adverse events). Alternatively, patients may stay longer at hospitals that have administrative inefficiencies.<sup>26</sup> CDI cases acquired later in a patient's stay may be more likely to be captured in the discharge records at hospitals with excess LOS. Thus, CDI incidence could be a proxy for hospital efficiency. The purpose of this study is to explore the relationship between CDI incidence rates and LOS for patients who do not have CDI, controlling for both hospital-level and patient-level characteristics.

## Methods

### Data Source

We used the Healthcare Cost and Utilization Project Nationwide Inpatient Sample, 2009–2011. The Nationwide Inpatient Sample, maintained by the Agency for Healthcare Research and Quality, is the largest database of inpatient records in the United States. It contains records of roughly 8 million hospital stays each year and is a 20% stratified sample of US hospitals. The Nationwide Inpatient Sample contains data on patient demographic characteristics, diagnoses, and procedures, measures of comorbidity and severity, reasons for and sources of admission, discharge disposition, hospital characteristics, charges and payment sources, and LOS for each unique patient record.<sup>27</sup>

To estimate the relationship between a hospital's CDI rate and excess LOS, we excluded all patients with any CDI diagnosis (primary or secondary) from the analysis of LOS. Excluding such patients eliminates the direct connection that exists between CDI and the increased LOS associated with CDI. However, patients with CDI were used in calculating

CDI incidence rates. Because we were interested in analyzing excess LOS, patients were excluded if they were admitted and discharged on the same day. Analysis was conducted at both an aggregated hospital level and a patient level. At the hospital level, all patients without CDI who had nonmissing values for LOS were included. For the patient-level analysis, patients were excluded if records contained missing values for any of the predictor variables described below. Table 1 provides a summary of the total number of hospitals and patients included at each stage of the analysis.

### Outcome and Predictor Variables

Two outcome measures were used for this analysis. The first was each hospital's average inpatient LOS, calculated as the average LOS across all patients *without* a CDI diagnosis. We modeled this outcome as a function of hospital-level characteristics. Our second outcome measure was individual patient-level LOS. We used this outcome in order to control for both hospital- and patient-level characteristics (Table 2). We compared the estimated effects of CDI incidence on LOS between these 2 outcomes in order to determine how much of the relationship between CDI incidence and LOS was due to patient characteristics.

The primary explanatory variable of interest was each hospital's annual CDI incidence rate. The incidence rate at each hospital was calculated as the ratio of the number of patient discharges with CDI diagnosis to the total number of annual discharges. Patients with CDI were identified as those with either a primary or secondary diagnosis of CDI (*International Statistical Classification of Disease, Ninth Revision*, code 008.45). This code has been previously validated as a measure for overall hospital CDI burden.<sup>28–30</sup>

Two additional sets of explanatory variables were used for this analysis, which included hospital- and patient-level variables. Previous research has found LOS to be related to hospital-level factors such as bed size, teaching status,<sup>12,31</sup> structure,<sup>32</sup> and nurse staffing.<sup>33</sup> Thus, for the hospital-level-LOS analysis we controlled for a set of 12 hospital-level characteristics, including bed size, hospital ownership, region of the country, location (urban vs rural), and teaching status, along with the percentage of all licensed nurses who were registered nurses and the number of full-time nurses per 1,000 inpatient days. Note, the Nationwide Inpatient Sample contains a very limited number of hospital-level variables.

LOS has also been shown to be related to many patient-level characteristics, such as age,<sup>34</sup> comorbidities,<sup>35</sup> disease severity,<sup>36</sup> and insurance status.<sup>13</sup> For the patient-level-LOS analysis, we controlled for all of the hospital-level variables along with a number of patient-level variables. Patient-level factors included patient demographic characteristics (eg, age, sex, primary payer, and ZIP-code-level income), inpatient-stay characteristics (eg, admission type, discharge quarter, weekend-admission indicator, discharge disposition, hospital-mortality indicator, neonatal and maternal indicators, along with the number of procedures, diagnoses, and chronic conditions), and disease characteristics (eg, All Patient Refined Diagnosis Related Groups indicators, severity, and risk of mortality categories; and 29 specific comorbidities). Table 2 provides a complete list of covariates, along with a description of each. In total, the patient- and hospital-level factors resulted in 429 separate covariates in the patient-level-LOS analysis.

## Statistical Analysis

All statistical analyses were conducted using Stata SE, version 13.1 (StataCorp). Multivariate regression was used to estimate the effect of CDI incidence on inpatient LOS while controlling for the predictor variables described. Weighted least squares regression, with weights corresponding to the number of discharge records, was used to analyze average hospital-level LOS.

However, for the patient-level analysis, because patient LOS tends to be nonnormally distributed (ie, skewed), 5 different regression models were compared to estimate the effect of a hospital's CDI incidence on LOS. These included (1) ordinary least squares, and a generalized linear model using a log link along with a (2) Gaussian, (3) gamma, (4) Poisson, and (5) negative binomial distribution. We compared the fit of these models using the Akaike information criterion, and estimates from the model with the lowest Akaike information criterion values are presented.

## Results

### CDI Rates and Hospital-Level LOS

We first categorized hospitals into deciles on the basis of their CDI incidence and calculated the average LOS across hospitals in each of these deciles; Table 3 presents these results. Table 3 reports a significant ( $P < .001$ ) positive correlation between LOS and CDI incidence, with LOS increasing by more than a day with a percentage point increase in CDI incidence. We then used multivariate regression to predict hospital-level LOS while controlling for hospital-specific characteristics. Results from the hospital-level regression analysis are presented in Table 4. The regression results mimic the findings of the bivariate comparisons between hospital deciles. For each year, a percentage point increase in a hospital's CDI incidence rate was associated with an increase in average patient LOS of 1.19 to 1.61 days. For each year, CDI incidence was the strongest predictor of average LOS in terms of the absolute value of its coefficient estimate and test statistic. Furthermore, when CDI incidence was removed, the explanatory power of the model, as measured by the model's  $R^2$  value, dropped considerably: without CDI incidence, the  $R^2$  values decreased from .45 to .11, from .51 to .12, and from .42 to .12 for each year from 2009 through 2011, respectively. Thus, CDI incidence explained the greatest amount of variation in average LOS between hospitals, of all the hospital-specific characteristics.

### CDI Rates and Patient-Level LOS

Variation in average LOS between hospitals may reflect underlying differences in patient populations rather than excess LOS. Thus, to control for patient characteristics, we used multivariate regression to predict patient-level LOS as a function of both patient and hospital characteristics.

The top of Table 5 presents the results of the regression model with the smallest Akaike information criterion value, namely the generalized linear model using a log link and gamma distribution. Results for the additional models are not presented but mirror those of the gamma model. For every year from 2009 through 2011, CDI incidence was strongly

associated with an increase in an individual's LOS ( $P < .001$ ). On the basis of the coefficients of the chosen regression model, for 2009, 2010, and 2011, a percentage point increase in a hospital's CDI incidence was roughly associated with an increase in a patient's LOS by a factor of approximately 4.37%, 7.47%, and 5.87%, respectively. Moreover, in terms of  $P$  values, CDI incidence was one of the strongest predictors of LOS. Among the 429 covariates included in this model, CDI incidence had the eighth lowest  $P$  value in 2009 and the seventh lowest in 2010 and 2011. Only a small set of variables, such as the number of procedures a patient underwent or the number of diagnoses on a patient's record, were stronger predictors of LOS.

As a point of comparison with the hospital-level LOS results, the bottom of Table 4 also reports the results of the patient-level ordinary-least-squares model with untransformed LOS. In this model, a percentage point increase in a hospital's CDI incidence was associated with a 0.64-, 1.05-, and 0.84-day increase in a patient's LOS, on average, for 2009, 2010, and 2011, respectively. These estimates are smaller than those in the hospital-level model, suggesting that some of the variation in average LOS associated with CDI incidence is due to patient-disease characteristics (ie, non-excess LOS). However, the relative size of the estimates in the patient-level model suggests that CDI incidence rates capture a significant portion of the excess LOS between hospitals. After accounting for patient-level characteristics, the effect size of the coefficient estimates of CDI incidence in the patient-level model were still greater than 50% of their value in the hospital-level model.

## Discussion

Our results demonstrate that hospital-level CDI rates are highly correlated with increased LOS in patients without a CDI diagnosis after controlling for patient and hospital characteristics. These results suggest that factors associated with high CDI rates in hospitals are also associated with excess LOS. Indeed, a 1% increase in an institution's CDI rate, after controlling for all observable variables, is associated with an increase between 4.37% and 7.47% in a non-CDI patient's LOS. These findings translated to an increase in LOS between 0.64 and 1.05 days. Thus, we believe that CDI acts as a proxy for hospital quality, efficiency, or perhaps both.

We found hospital CDI incidence to be a highly significant predictor of LOS at both the hospital and patient level. In fact, CDI rates were one of the strongest predictors in each model, stronger than all other hospital characteristics and most patient characteristics. Moreover, many common patient-level characteristics used in studies of patient outcomes (eg, age) had a far smaller impact on LOS than did CDI incidence in our model. These results suggest that unmeasured hospital characteristics that are captured by CDI incidence may play a greater role in determining a patient's LOS than the patient's and hospital's underlying characteristics. In every model we estimated, the quality and fit (eg, Akaike information criterion or  $R^2$ ) of our model improved substantially when CDI incidence was included. This finding alone suggests that future research using LOS as an outcome measure should consider CDI incidence, or the factors it captures, as a proxy variable. Failure to account for CDI incidence may result in omitted variable bias leading to an incorrect interpretation of the effects of other variables related to LOS: the magnitudes, signs, and

significance of many of the estimated coefficients included in our model changed dramatically when CDI incidence was removed. Given the abundance of existing research that has used LOS as an outcome measure, it may be worth revisiting factors previously associated with LOS.

A useful feature of CDI rates in comparing LOS between hospitals is that CDI rates appear to be a good proxy for excess LOS rather than LOS due to disease characteristics. A potential limitation of any measure used to make hospital-level comparisons is the need to properly adjust for the type of patients a hospital treats. For example, hospitals that treat a more severely ill patient population may appear to have longer LOS on average, even though such increased LOS would not be considered excessive. Our results suggest that CDI rates may discriminate between excess and ordinary LOS. After individual patient characteristics were added to the hospital-level model, greater than 50% of the effect of CDI on LOS remained. These findings suggest that more than half of the variation in LOS between hospitals that can be explained by CDI rates may be due to excess LOS. Thus, CDI rates may be useful when comparing excess LOS between hospitals.

CDI rates are easy to compute and to compare across hospitals, especially in relation to many measures of quality or other markers of patient safety. The Agency for Healthcare Research and Quality indicators have frequently been used as measures of hospital quality and patient safety. However, such indicators require many variables, and the coding algorithm for these indicators has changed over time.<sup>37</sup> Less complex measures, such as adherence to guidelines for acute myocardial infarction, often focus on only specific patient populations. In contrast, CDI rates are simple to calculate and easy to compare across hospitals. CDI rates are inherently important to measure, and our results provide another reason for tracking and considering CDI rates.

Although CDI rates are associated with longer LOS in patients who do not have CDI, we do not claim that higher CDI rates are necessarily causing longer LOS. Instead, we think that CDI rates may act as a proxy for unobserved/unmeasured hospital characteristics that are related to hospital efficiency and/or quality (eg, environmental cleanliness, hospital crowding, and inappropriate and excess use of antibiotics). We hypothesize that 2 main factors play a role in driving this relationship. First, hospitals that are of lower quality or less efficient may tend to have longer LOS and generate more hospital-associated CDI. Second, hospitals with longer LOS, due to either efficiency or quality characteristics, may also observe more hospital-associated CDI before discharge of patients. Thus, there is no guarantee that efforts to reduce CDI may affect LOS in uninfected patients. In addition, CDI rates may be dependent upon connections with other hospitals via patient transfers that we are unable to observe in this analysis.<sup>38</sup> Future investigations should focus on analyzing potentially causative factors driving the relationship between CDI rates and LOS in patients without CDI. For example, the additional isolation rooms needed for hospitals with a higher CDI incidence may result in ineffective transitions of care or misallocation of staffing resources. Unfortunately, we are unable to perform this analysis with our data.

The connection between CDI incidence and LOS may occur because a hospital is generating more hospital-associated CDI. One limitation of our study is we cannot directly determine

whether a CDI diagnosis was hospital associated. Therefore, we performed a sensitivity analysis where we calculated CDI incidence using only CDI cases that were recorded as a secondary diagnosis. Although secondary CDI diagnoses have been shown to contain non-hospital-associated CDI cases,<sup>29</sup> by removing primary CDI cases, the calculated CDI incidence should contain a greater proportion of hospital-associated CDI cases. Results are reported in the Online Supplementary Appendix. When only secondary CDI cases were included, our findings became even stronger: both the estimated effect and significance of CDI incidence on LOS increased, and the fit of the model improved. Thus, the link between CDI incidence and LOS may occur via hospital-associated CDI.

There are other limitations to our study. First, we used administrative data rather than clinical microbiologic results to measure CDI. However, administrative codes for CDI have been demonstrated to be a relatively sensitive and specific marker for CDI.<sup>28–30</sup> Second, rates of CDI differ over time and region and depend upon different microbiologic testing approaches for CDI that undoubtedly differ across institutions. However, our study was conducted over a series of years and included many types of hospitals. Third, it would be ideal to have information regarding CDI cases attributable to hospital stays that occur after hospital discharge; the number of such cases may be nontrivial.<sup>39</sup> However, we have only inpatient data. Although this limitation does not detract from the effectiveness of CDI rates as a marker for hospitals with longer LOS, such information would be useful in further analyzing the connection between CDI rates and LOS. Fourth, in our patient-level analysis, we dropped a number of patients with missing values in the covariates. However, we are not concerned that this biased our results, given that our hospital-level model, which included all patients, showed a similar, yet stronger, association between CDI rates and LOS. Fifth, we included a large number of patient-level covariates in our analyses to avoid omitted variable bias. More parsimonious models, containing more patients but fewer patient-level variables, actually increased our observed effect (data not shown). Although the inclusion of a large number of variables could render the model susceptible to multicollinearity, we found no evidence of model instability or estimation inaccuracy. Finally, a limitation associated with all such modeling efforts is that the estimates we generate for excess LOS associated with higher CDI should be interpreted with caution: additional quality or efficiency measures are necessary to estimate the exact effect size.

In conclusion, CDI rates are an accurate predictor of LOS in patients without CDI, even after considering both individual- and institution-level factors. CDI incidence had greater explanatory power than any other hospital characteristic and almost all commonly used patient characteristics. Moreover, differences in CDI rates between hospitals appear to capture differences in excess, rather than ordinary, LOS. CDI rates are easy to measure and may provide an important marker for hospital efficiency and quality. Thus, our findings may provide another reason for policy makers, healthcare administrators, and clinicians to track CDI rates.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

*Financial support.* National Heart, Lung, and Blood Institute of the National Institutes of Health (grant K25HL122305); and the University of Iowa Health Care eHealth and eNovation Center.

## References

1. Polverejan E, Gardiner JC, Bradley CJ, et al. Estimating mean hospital cost as a function of length of stay and patient characteristics. *Health Econ.* 2003; 12:935–947. [PubMed: 14601156]
2. Hauck K, Zhao X. How dangerous is a day in hospital? A model of adverse events and length of stay for medical inpatients. *Med Care.* 2011; 49:1068–1075. [PubMed: 21945976]
3. Graffunder EM, Venezia RA. Risk factors associated with nosocomial methicillin-resistant *Staphylococcus aureus* (MRSA) infection including previous use of antimicrobials. *J Antimicrob Chemother.* 2002; 49:999–1005. [PubMed: 12039892]
4. Bankowitz RA, Doyle B, Duan M, et al. Identifying hospital-wide harm: a set of ICD-9-CM-coded conditions associated with increased cost, length of stay, and risk of mortality. *Am J Med Qual.* 2014; 29:373–380. [PubMed: 24081831]
5. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA.* 2003; 290:1868–1874. [PubMed: 14532315]
6. Kossovsky MP, Sarasin FP, Chopard P, et al. Relationship between hospital length of stay and quality of care in patients with congestive heart failure. *Qual Saf Health Care.* 2002; 11:219–223. [PubMed: 12486984]
7. Thomas JW, Guire KE, Horvat GG. Is patient length of stay related to quality of care? *Hosp Health Serv Adm.* 1997; 42:489–507. [PubMed: 10174462]
8. Southern WN, Bellin EY, Arnsten JH. Longer lengths of stay and higher risk of mortality among inpatients of physicians with more years in practice. *Am J Med.* 2011; 124:868–874. [PubMed: 21745651]
9. Scott I, Youlden D, Coory M. Are diagnosis specific outcome indicators based on administrative data useful in assessing quality of hospital care? *Qual Saf Health Care.* 2004; 13:32–39. [PubMed: 14757797]
10. Edwards WH, Morris JA Jr, Jenkins JM, et al. Evaluating quality, cost-effective health care: vascular database predicated on hospital discharge abstracts. *Ann Surg.* 1991; 213:433–438. [PubMed: 2025063]
11. Hollingsworth B. The measurement of efficiency and productivity of health care delivery. *Health Econ.* 2008; 17:1107–1128. [PubMed: 18702091]
12. McDermott C, Stock GN. Hospital operations and length of stay performance. *Int J Operations Production Management.* 2007; 27:1020–1042.
13. Yang J, Peek-Asa C, Allareddy V, et al. Patient and hospital characteristics associated with length of stay and hospital charges for pediatric sports-related injury hospitalizations in the United States, 2000–2003. *Pediatrics.* 2007; 119:e813–e820. [PubMed: 17403824]
14. Dunne TJ, Gaboury I, Ashe MC. Falls in hospital increase length of stay regardless of degree of harm. *J Eval Clin Pract.* 2014; 20:396–400. [PubMed: 24814338]
15. Wong CA, Recktenwald AJ, Jones ML, et al. The cost of serious fall-related injuries at three Midwestern hospitals. *Jt Com J Qual Patient Saf.* 2011; 37:81–87.
16. Classen DC, Pestotnik SL, Evans RS, et al. Adverse drug events in hospitalized patients: excess length of stay, extra costs, and attributable mortality. *JAMA.* 1997; 277:301–306. [PubMed: 9002492]
17. Graves N, Birrell F, Whitby M. Effect of pressure ulcers on length of hospital stay. *Infect Control Hosp Epidemiol.* 2005; 26:293–297. [PubMed: 15796283]
18. 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; 114:348–355. [PubMed: 15286215]



19. Macedo-Vinas M, De Angelis G, Rohner P, et al. Burden of meticillin-resistant *Staphylococcus aureus* infections at a Swiss University hospital: excess length of stay and costs. *J Hosp Infect.* 2013; 84:132–137. [PubMed: 23608003]
20. 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; 78:86–91. [PubMed: 21481492]
21. Rivard PE, Luther SL, Christiansen CL, et al. Using patient safety indicators to estimate the impact of potential adverse events on outcomes. *Med Care Res Rev.* 2008; 65:67–87. [PubMed: 18184870]
22. Monge Jodra V, Sainz de Los Terreros Soler L, Diaz-Agero Perez C, et al. Excess length of stay attributable to surgical site infection following hip replacement: a nested case-control study. *Infect Control Hosp Epidemiol.* 2006; 27:1299–1303. [PubMed: 17152026]
23. Lipp MJ, Nero DC, Callahan MA. Impact of hospital-acquired *Clostridium difficile*. *J Gastroenterol Hepatol.* 2012; 27:1733–1737. [PubMed: 22849881]
24. Dubberke ER, Butler AM, Reske KA, et al. Attributable outcomes of endemic *Clostridium difficile*-associated disease in nonsurgical patients. *Emerg Infect Dis.* 2008; 14:1031–1038. [PubMed: 18598621]
25. Kyne L, Hamel MB, Polavaram R, et al. Health care costs and mortality associated with nosocomial diarrhea due to *Clostridium difficile*. *Clin Infect Dis.* 2002; 34:346–353. [PubMed: 11774082]
26. White CM, Statile AM, White DL, et al. Using quality improvement to optimise paediatric discharge efficiency. *BMJ Qual Saf.* 2014; 23:428–436.
27. Agency for Healthcare Research and Quality. Overview of the National (Nationwide) Inpatient Sample (NIS). AHRQ website. <http://www.hcup-us.ahrq.gov/nisoverview.jsp>. Accessed November 16, 2015
28. Dubberke ER, Reske KA, McDonald LC, et al. ICD-9 codes and surveillance for *Clostridium difficile*-associated disease. *Emerg Infect Dis.* 2006; 12:1576–1579. [PubMed: 17176576]
29. Dubberke ER, Butler AM, Yokoe DS, et al. Multicenter study of surveillance for hospital-onset *Clostridium difficile* infection by the use of ICD-9-CM diagnosis codes. *Infect Control Hosp Epidemiol.* 2010; 31:262–268. [PubMed: 20100085]
30. Scheurer DB, Hicks LS, Cook EF, et al. Accuracy of ICD-9 coding for *Clostridium difficile* infections: a retrospective cohort. *Epidemiol Infect.* 2007; 135:1010–1013. [PubMed: 17156501]
31. Freitas A, Silva-Costa T, Lopes F, et al. Factors influencing hospital high length of stay outliers. *BMC Health Serv Res.* 2012; 12:265. [PubMed: 22906386]
32. Cots F, Mercade L, Castells X, et al. Relationship between hospital structural level and length of stay outliers: implications for hospital payment systems. *Health Policy.* 2004; 68:159–168. [PubMed: 15063016]
33. Thungjaroenkul P, Cummings GG, Embleton A. The impact of nurse staffing on hospital costs and patient length of stay: a systematic review. *Nurs Econ.* 2007; 25:255–265. [PubMed: 18080621]
34. Hoonhout LH, de Bruijne MC, Wagner C, et al. Direct medical costs of adverse events in Dutch hospitals. *BMC Health Serv Res.* 2009; 9:27. [PubMed: 19203365]
35. Wen T, He S, Attenello F, et al. The impact of patient age and comorbidities on the occurrence of “never events” in cerebrovascular surgery: an analysis of the Nationwide Inpatient Sample. *J Neurosurg.* 2014; 121:580–586. [PubMed: 24972123]
36. Horn SD, Sharkey PD, Buckle JM, et al. The relationship between severity of illness and hospital length of stay and mortality. *Med Care.* 1991; 29:305–317. [PubMed: 1902275]
37. Agency for Healthcare Research and Quality. Patient Safety Indicators (PSI) log of revisions to PSI documentation and software. AHRQ website. [http://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V50/ChangeLog\\_PSI\\_v50.pdf](http://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V50/ChangeLog_PSI_v50.pdf). Published March 2015 Accessed November 16 2015
38. Simmering JE, Polgreen LA, Campbell DR, et al. Hospital transfer network structure as a risk factor for *Clostridium difficile* infection. *Infect Control Hosp Epidemiol.* 2015; 36:1031–1037. [PubMed: 26072907]

39. Kuntz JL, Polgreen PM. The importance of considering different healthcare settings when estimating the burden of *Clostridium difficile*. Clin Infect Dis. 2014; 60:831–836. [PubMed: 25477426]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 1**

Study Population in Study of CDI as a Proxy for Length of Stay

Variable	2009	2010	2011
Total patients in NIS	7,810,762	7,800,441	8,023,590
Total CDI cases in NIS	66,623	69,315	79,633
Total hospitals in NIS	1,050	1,051	1,049
Patients in final sample	4,126,716	4,585,926	4,841,294
Hospitals in final sample	942	949	945

Note. CDI, *Clostridium difficile* infection; NIS, National Inpatient Sample.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

## Summary Description of All Model Covariates in Study of CDI as a Proxy for Length of Stay

Variable	Description/specification
Hospital characteristics	
Bed size	Small, medium, or large
Control/ownership	Public, private (nonprofit), private (for profit)
Region	Northeast, Midwest, South, West
Teaching status/location	Rural, urban nonteaching, urban teaching
RN percent	Percentage of RNs among all licensed nurses
Nurse to patient ratio	Number of total licensed nurse full-time equivalents per 100 inpatient-days
Patient characteristics	
Age group	21 indicators for 5-year age ranges
Sex	Female or male
Race	White, black, Hispanic, Asian or Pacific Islander, Native American, or other
Admission type	Emergency, urgent, elective, newborn, trauma center, other
Disposition	6 indicators for routine, transfer to short-term hospital, other transfer, home healthcare, against medical advice, or unknown
OR procedure	Indicator of major OR procedure
Neonatal or maternal diagnosis/procedure	Indicators for maternal, neonatal, or both maternal and neonatal records
Admission month	12 monthly indicators
Admission weekend	Indicator of weekend admission
Died	Indicator of hospital death
Comorbidities	29 comorbid disease indicators <sup>a</sup>
Number of procedures	30 indicators for the total number of procedures coded on discharge record
Number of diagnoses	30 indicators for the total number of diagnoses coded on discharge record
Number of chronic conditions	30 indicators for the total number of unique chronic diagnoses reported on the discharge
APR DRGs <sup>b</sup>	316 APR DRG indicators
APR DRG risk mortality <sup>b</sup>	4 indicators for likelihood of dying: minor, moderate, major, or extreme
APR DRG severity <sup>b</sup>	4 indicators for severity of illness (loss of function): minor, moderate, major, or extreme
Primary expected payer	Medicare, Medicaid, private insurance, self-pay, no charge, or other
ZIP code median income	4 quartile indicators for estimated median household income in patient's ZIP code

Note. APR, All Patient Refined; CDI, *Clostridium difficile* infection; DRG, diagnosis-related groups; OR, operating room; RN, registered nurse.

<sup>a</sup>The 29 comorbidity indicators were assigned by the Agency for Healthcare Research and Quality Comorbidity Software, version 3.7.

<sup>b</sup>The variables for APR DRGs along with APR DRG severity and APR DRG risk mortality were created using software developed by 3M Health Information Systems, version 27.0, for year 2009 and version 28.0 for years 2010 and 2011.

Table 3

LOS and CDI Incidence Rates by Hospital CDI Decile (Weighted Average)

CDI decile	2009		2010		2011	
	LOS	CDI rate	LOS	CDI rate	LOS	CDI rate
1	3.511	.003	3.499	.028	3.487	.030
2	4.019	.181	3.912	.204	4.138	.263
3	4.118	.342	4.083	.378	4.045	.431
4	4.186	.470	4.323	.502	4.156	.567
5	4.362	.578	4.505	.623	4.439	.692
6	4.553	.711	4.681	.759	4.426	.846
7	4.502	.836	4.694	.924	4.438	1.018
8	4.634	1.029	4.562	1.106	4.491	1.231
9	4.605	1.360	4.624	1.455	4.898	1.583
10	6.280	2.559	6.722	2.803	6.002	2.880
p		.6709		.8243		.7551
Pvalue		<.001		<.001		<.001

Note. CDI, *Clostridium difficile* infection; LOS, length of stay. LOS is calculated as the average patient LOS across all patients without CDI for hospitals in a given decile. CDI rates are calculated as weighted averages across hospitals in a given decile, weighted by each hospital's total number of discharges. The correlation coefficient and *P* values correspond to the correlation between each hospital's average LOS and CDI incidence.

**Table 4**  
Hospital-Level Results From a Weighted Least Squares Regression With Weights Corresponding to the Number of Discharges per Hospital

Variable	2009			2010			2011		
	Coefficient (SE)	P value	Coefficient (SE)	P value	Coefficient (SE)	P value	Coefficient (SE)	P value	
CDI incidence	1.1909 (.0500)	<.001	1.6096 (.0593)	<.001	1.3011 (.0600)	<.001			
Hospital bed size									
Small	1 [Reference]		1 [Reference]		1 [Reference]		1 [Reference]		
Medium	.0782 (.1374)	.569	-.2108 (.1362)	.122	.2412 (.1437)	.094			
Large	.4407 (.1245)	<.001	.2704 (.1229)	.028	.5726 (.1299)	<.001			
Hospital control									
Government	1 [Reference]		1 [Reference]		1 [Reference]		1 [Reference]		
Private (nonprofit)	-.4635 (.1200)	<.001	-.5868 (.1208)	<.001	-.6954 (.1390)	<.001			
Private (for profit)	-.1394 (.1551)	.369	-.1578 (.1630)	.333	-.2744 (.1678)	.102			
Hospital region									
Northeast	1 [Reference]		1 [Reference]		1 [Reference]		1 [Reference]		
Midwest	-.4038 (.1195)	.001	-.5344 (.1224)	<.001	-.5696 (.1236)	<.001			
South	-.1828 (.1153)	.113	-.0216 (.1162)	.853	-.2477 (.1186)	.037			
West	-.3522 (.1342)	.009	-.3355 (.1335)	.012	-.6358 (.1376)	<.001			
Location and teaching status									
Rural	1 [Reference]		1 [Reference]		1 [Reference]		1 [Reference]		
Urban nonteaching	.5609 (.1375)	<.001	.2453 (.1371)	.074	.4476 (.1495)	.003			
Urban teaching	1.0930 (.1390)	<.001	1.0401 (.1375)	<.001	1.1040 (.1534)	<.001			
% of RNs among licensed nurses	-.0304 (.0063)	<.001	-.0307 (.0067)	<.001	-.0439 (.0071)	<.001			
Nurses per 1,000 inpatient-days	-.0333 (.0306)	.278	-.0073 (.0316)	.818	.0153 (.0308)	.620			
N	942		949		945				
R <sup>2</sup>	.4494		.5070		.4180				

Note. CDI, *Clostridium difficile* infection; RN, registered nurse. The dependent variable is average hospital length of stay for patients without CDI.

**Table 5**  
Patient-Level Results From GLM (Gamma Family, Log Link) and OLS Models

Variable	2009		2010		2011	
	Coefficient (SE)	P value	Coefficient (SE)	P value	Coefficient (SE)	P value
GLM-gamma (log link)						
CDI incidence	.0428 (.0005)	<.001	.0720 (.0005)	<.001	.0570 (.0005)	<.001
AIC	19,125,174		21,291,522		22,539,057	
OLS						
CDI incidence	.6420 (.0033)	<.001	1.0523 (.0039)	<.001	.8379 (.0036)	<.001
AIC	24,587,709		27,223,452		29,536,720	
N	4,126,716		4,585,926		4,841,294	

Note. AIC, Akaike information criterion; CDI, *Clostridium difficile* infection; GLM, generalized linear models; OLS, ordinary least squares; SE, standard error. The dependent variable is length of stay for patients without CDI.