2011

# Development and validation of a Clostridium difficile infection risk prediction model

Erik R. Dubberke
*Washington University School of Medicine in St. Louis*

Yan Yan
*Washington University School of Medicine in St. Louis*

Kimberly A. Reske
*Washington University School of Medicine in St. Louis*

Anne M. Butler
*Washington University School of Medicine in St. Louis*

Joshua Doherty
*Barnes-Jewish Hospital*

***See next page for additional authors***

Development and Validation of a Clostridium difficile Infection Risk Prediction Model
Author(s): Erik R. Dubberke, Yan Yan, Kimberly A. Reske, Anne M. Butler, Joshua Doherty, Victor Pham, Victoria J. Fraser
Reviewed work(s):
Source: *Infection Control and Hospital Epidemiology*, Vol. 32, No. 4 (April 2011), pp. 360-366
Published by: The University of Chicago Press on behalf of The Society for Healthcare Epidemiology of America
Stable URL: http://www.jstor.org/stable/10.1086/658944
Accessed: 06/03/2012 23:19

ORIGINAL ARTICLE

# Development and Validation of a *Clostridium difficile* Infection Risk Prediction Model

**Erik R. Dubberke, MD, MSPH;[1] Yan Yan, MD, PhD;[1] Kimberly A. Reske, MPH;[1] Anne M. Butler, MS;[1] Joshua Doherty, BS;[2] Victor Pham, BS;[2] Victoria J. Fraser, MD[1]**

OBJECTIVE. To develop and validate a risk prediction model that could identify patients at high risk for *Clostridium difficile* infection (CDI) before they develop disease.

DESIGN AND SETTING. Retrospective cohort study in a tertiary care medical center.

PATIENTS. Patients admitted to the hospital for at least 48 hours during the calendar year 2003.

METHODS. Data were collected electronically from the hospital's Medical Informatics database and analyzed with logistic regression to determine variables that best predicted patients' risk for development of CDI. Model discrimination and calibration were calculated. The model was bootstrapped 500 times to validate the predictive accuracy. A receiver operating characteristic curve was calculated to evaluate potential risk cutoffs.

RESULTS. A total of 35,350 admitted patients, including 329 with CDI, were studied. Variables in the risk prediction model were age, CDI pressure, times admitted to hospital in the previous 60 days, modified Acute Physiology Score, days of treatment with high-risk antibiotics, whether albumin level was low, admission to an intensive care unit, and receipt of laxatives, gastric acid suppressors, or antimotility drugs. The calibration and discrimination of the model were very good to excellent (C index, 0.88; Brier score, 0.009).

CONCLUSIONS. The CDI risk prediction model performed well. Further study is needed to determine whether it could be used in a clinical setting to prevent CDI-associated outcomes and reduce costs.

*Infect Control Hosp Epidemiol* 2011;32(4):360-366

The morbidity, mortality, and financial burden of *Clostridium difficile* infection (CDI) are significant.[1-9] Primary risk factors for CDI include receipt of antimicrobials, older age, high severity of illness, exposure to other patients with CDI, and receipt of gastric acid suppressants.[10-17] Despite the breadth and depth of knowledge about risk factors available in the medical literature, healthcare facilities still struggle with outbreaks of CDI and increasing rates of disease.[4-7,9,18-20] Most CDI prevention and control strategies, such as the use of contact precautions and environmental decontamination, emphasize prevention of secondary transmission (ie, the spread of *C. difficile* from an infected patient to uninfected patients).[15,21-24] Few data are available on strategies to identify patients at increased risk for CDI or interventions to prevent primary transmission of *C. difficile* and/or subsequent development of symptomatic CDI.

Risk prediction modeling has been used in the study of chronic and infectious diseases to identify important risk factors and to quantify or rank each factor's comparative importance in the development of disease.[25,26] Unlike risk factor studies, which identify characteristics that increase a patient's risk for disease, risk prediction modeling allows researchers to "score" patients according to the importance of these characteristics, thereby identifying patients at highest risk for disease. If applied to CDI prevention in medical facilities, risk prediction modeling could allow healthcare workers to identify patients at highest risk for CDI and to intervene before these patients become ill by modifying either the patients' individual risk factors or their environment. Risk prediction modeling of CDI is not well studied.

To our knowledge, only 2 studies have used this method to identify patients at high risk for CDI, and neither is broadly generalizable.[27,28] The study population used by Garey et al[27] was restricted to patients receiving antibiotics, and the models were developed with data collected from multiple sources, some of which were available only after hospital discharge. Tanner et al[28] used the Waterlow score, a measure not commonly used in the United States, to develop their index. The Waterlow score is based on patient data at the time of admission, but the risk of CDI may fluctuate over the course of a hospitalization, depending on patient characteristics, clinical events, exposure to *C. difficile*, and medications re-

Affiliations: 1. Washington University School of Medicine, St. Louis, Missouri; 2. Barnes-Jewish Hospital, St. Louis, Missouri.

Received June 1, 2010; accepted September 20, 2010; electronically published February 21, 2011.

 0899-823X/2011/3204-0007$15.00. DOI: 10.1086/658944

ceived.[29] These factors limit the usefulness of the Tanner et al[28] indices on a real-time or prospective basis. The purpose of our study was to develop and validate a CDI risk prediction model that could ultimately be employed in real time to prevent CDIs in a medical facility.

## METHODS

Data were collected electronically on a large, retrospective cohort of patients hospitalized at Barnes-Jewish Hospital (BJH) for at least 48 hours during the calendar year 2003.[14] Data collected included patients' demographics, medications received, and *C. difficile* toxin assay results; laboratory tests; *International Classification of Diseases, Ninth Revision* (*ICD-9*) discharge and procedure codes; and hospital ward(s) of stay. Patients with a history of CDI in the previous 60 days were excluded because initial analyses indicated that the effect of this characteristic was substantially higher than that of all the other variables considered in the analysis. Patients admitted to bone marrow transplantation (BMT) or leukemia wards at BJH were also excluded because further analyses suggested that the risk factors for developing CDI in BMT and leukemia patients were different from those in other hospitalized patients. Removal of these patients allowed for more detailed analyses of the remaining patients. The variable CDI pressure has been described previously.[13] Briefly, this variable is a modified version of colonization pressure: mean CDI pressure is measured as each patient's total exposure to infectious CDI patients divided by the patient's length of stay at risk for CDI. The cohort was analyzed previously for CDI risk factors by pooled logistic regression; detailed methods and results of this analysis have been published elsewhere.[1,2,13,14]

Data from this 2003 cohort were used to develop a CDI risk prediction model. Variables retained in the final risk factor model developed for the previous study were reviewed to determine whether the data were available in real time; if not, an appropriate surrogate variable was substituted. Data for receipt of mechanical ventilation, which was based on *ICD-9* procedure codes assigned at hospital discharge, was not available in real time. Admission to an intensive care unit (ICU) was selected as a surrogate because all BJH patients who receive mechanical ventilation are initially admitted to an ICU. No other variable substitutions were necessary.

As the first step in selecting variables for the model, we used data reduction techniques to combine variables on the basis of substantive knowledge and statistical methods to avoid overspecification of the model correlation among related variables in the model. Principal-component analyses, corresponding analyses, and cluster analyses were used to determine whether logical groupings of variables were statistically valid.[30] The only grouping retained in the final model was "high-risk antibiotics" (cephalosporins, clindamycin, and amoxicillin or ampicillin), designated as a group because risk for CDI among patients receiving these antibiotics has consistently been reported as high in the literature and because data analysis confirmed that they could be combined into one category. In this combined variable, receipt of at least one of these antibiotics in a single day was considered a single day of exposure. When the study was planned, a decision was made to investigate appropriate functional formats for the continuous variables in the model. The appropriate functional formats of continuous variables were determined by examining nonparametric regression (smoothing) plots with a restricted cubic spline function. If possible, these complicated forms were simplified using low-order polynomials or piecewise linear splines for more robust prediction. Logistic stepwise regression was used to determine a set of variables that best predicted the risk of CDI. Potential clinically important interactions were tested. To facilitate interpretation of results, odds ratios for the displayed polynomial and piecewise linear spline variables compare the odds of developing CDI for values between the 75th and the 25th percentiles of each variable.[26]

The predictive accuracy of the model was determined by discrimination (C statistic) and calibration (Brier score).[30] The C statistic, also known as the area under the curve (AUC), measures discrimination, the ability of the model to separate cases from noncases.[31] The closer the C statistic is to 1, the better the discrimination of the model. The Brier score measures calibration, the closeness of predicted probabilities from the risk model to the observed outcome (ie, a patient developing CDI).[31] The closer the Brier score is to 0, the better the calibration of the model. The final model was bootstrapped 500 times to validate predictive accuracy.[32] Briefly, the bootstrapping methods involved drawing a sample from the original data and estimating the bootstrap regression coefficients with the bootstrap sample. Next, the apparent predictive accuracy index was calculated by applying coefficients to the bootstrap sample, and the predictive accuracy index was calculated by applying the coefficients to the original data. The bootstrap optimism (difference between apparent and bootstrap accuracy indices) was calculated, and the process was repeated 500 times. Finally, the average optimism was subtracted from the apparent predictive accuracy index to obtain the bias-corrected predictive accuracy index. A receiver operating characteristic (ROC) curve was created to assess which CDI risk cutoffs might be used to initiate an intervention. Statistical analyses were performed with SPSS, SAS, and R. This study was approved by the Washington University Human Research Protection Office.

## RESULTS

The final data set included 35,350 admitted patients, of whom 329 were patients with CDI. Of the admitted patients, 1,275 (3.5%) were excluded because of a previous history of CDI or because they were BMT patients. Characteristics of the study population are presented in Table 1. All differences between patients with CDI and those without CDI were sig-

TABLE 1. Demographic Characteristics of the Study Population

| Characteristic | Patients with CDI | Patients without CDI |
|---|---|---|
| Female sex, proportion (%) of patients | 166/329 (51) | 20,438/35,021 (58) |
| White race, proportion (%) of patients | 232/329 (71) | 21,836/35,021 (62) |
| Age, years, median (range) | 66 (21–99) | 56 (12–106) |
| Modified APS on admission to hospital, median (range) | 6 (0–28) | 4 (0–31) |
| CDI pressure, median (range) | 1.21 (0–15.64) | 0.22 (0–20.39) |

NOTE:. All differences were significant at the $P \leq .05$ level. APS, Acute Physiology Score; CDI, *Clostridium difficile* infection.

nificant ($P \leq .05$). CDI patients were less likely to be female and more likely to be white than patients without CDI (51% vs 58% and 71% vs 62%, respectively). Patients with CDI were older than those without CDI (median ages of 66 and 56 years, respectively) and had a higher acuity of illness on admission to the hospital (median modified Acute Physiology Score [APS] of 6 for patients with CDI and 4 for those without CDI). Patients with CDI had higher mean CDI pressure scores than patients without CDI (median scores of 1.21 and 0.22, respectively).

Variables in the final logistic regression model included age, CDI pressure, previous hospital admissions within 60 days, modified APS,[14] days of treatment with high-risk antibiotics, admission to an ICU, receipt of laxatives, receipt of gastric acid suppressors, receipt of an antimotility drug, and whether albumin level was low (Table 2). After evaluating various functional forms of the continuous variables to determine which form best represented the data, age and the modified APS were modeled as polynomial functions. CDI pressure and days on high-risk antibiotics were modeled as spline functions. The number of admissions in the previous 60 days was modeled categorically with three levels (0, 1, and 2 or more admissions). All other variables were modeled as dichotomous (yes/no) categorical variables. The polynomial form of the modified APS was not statistically significant in the final model, but it was retained because it was significant prior to modification and because of the clinical importance of severity of illness as a risk factor for CDI. The CDI risk prediction model demonstrated excellent discrimination and calibration in the original data set and in the bootstrap samples (C statistic of 0.88 and Brier score of 0.009 for both; Figure 1). Our model slightly overestimates the probability of CDI for patients whose true probability is higher than 0.075. However, 91.98% of our sample have predicted probabilities lower than 0.075, and 99.8% of the sample have predicted probabilities lower than 0.21. Therefore, the overestimation affects only a very small proportion of patients.

Table 3 presents the predicted probabilities of developing CDI with the sensitivity and specificity of the CDI risk prediction model at different thresholds of CDI risk. For example, at the 0.0233 level, the predicted probability of CDI is 2.3%, the sensitivity of the model is 60%, and the specificity is 89%. As the predicted probability of developing CDI increases, the sensitivity of the model decreases and the specificity increases. Figure 2 presents the ROC curve, graphically depicting the ability of the model to discriminate between true cases and true noncases. On the basis of these data, potential thresholds for intervention could be when the probability of developing CDI reaches 0.023 (2.3% risk of CDI) or 0.030 (3.0% risk of CDI). At the 0.023 level, the number of CDI cases prevented over the course of 1 year at BJH would be up to 198 (60% of the total CDI cases included in this study), and the number of patients identified as at high risk for CDI who did not develop the disease would be 3,834. At the 0.030 level, the number of CDI cases prevented over 1 year would be up to 165 (50% of the CDI cases included in this study), and the number of patients identified as at high risk who did not develop CDI would be 2,721.

## CONCLUSION

Risk prediction modeling has not been used frequently in infectious-disease epidemiology, particularly with CDI. In this study we developed a CDI risk prediction model using readily available electronic data. The risk assessment provided by our model was strong, according to statistical measures that evaluate model discrimination and calibration. The discrimination of our model (C statistic, 0.88) was high, which indicates that the model successfully identified patients who would develop CDI. Calibration was low (Brier score, 0.009), suggesting that the risk prediction model can accurately estimate the future probability of the event in question (ie, a patient developing CDI). Having excellent discrimination and calibration are important characteristics of a CDI risk prediction index because the incidence of CDI is relatively low. In a low-incidence setting, the risk of a false positive result is increased (ie, identifying a patient who will not develop CDI as being at high risk for CDI). Interventions aimed at preventing CDI in patients at high risk for CDI should be cost-effective. If discrimination and calibration are not sufficient, the high number of false positives will decrease the cost-effectiveness of any potential intervention.

An intervention using the risk prediction model presented here and based on estimates of the CDI-attributable cost, length of stay, and mortality could be cost-effective. Potential interventions could include preemptive contact precautions, prophylactic oral vancomycin, or pharmacy-based detailing to reduce unnecessary medications that increase a patient's

TABLE 2. Variables in the Final Logistic Regression for the *Clostridium difficile* Infection (CDI) Risk Prediction Model

| Variable | Symbol | Odds ratio | 95% confidence interval |
|---|---|---|---|
| Age[a,b] | AGE | 2.5 | 1.9–3.3 |
| CDI pressure[b,c] | CP | 13.0 | 8.5–19.9 |
| 1 hospital admission in previous 60 days | ADMIT60D | 1.0 | 0.7–1.3 |
| 2 or more hospital admissions in previous 60 days | ADMIT60D | 2.7 | 2.0–3.7 |
| Modified APS on admission to hospital[a,b] | MOD_APS | 1.1 | 0.9–1.4 |
| Days received high-risk antibiotics[b,c,d] | HRABX | 1.9 | 1.6–2.3 |
| Admission to ICU | ICUPT | 1.6 | 1.2–2.0 |
| Received laxatives | LAX | 1.3 | 1.0–1.7 |
| Received gastric acid suppressor | GAS | 2.1 | 1.5–3.0 |
| Received antimotility drug | MOTIL | 2.1 | 1.6–2.6 |
| Low albumin (≤3.5 g/dL) on admission to hospital | ALBUMIN | 1.7 | 1.3–2.2 |

NOTE. Equation for the final model:

$$\begin{aligned}\text{Logit}(\Pr(\text{CDI}=1)) = {} & -11.231076 + 0.050270561 \times \text{AGE} - 0.00018855342 \times \text{AGE}^2 + 8.3591798 \times \text{CP} \\ & - 8.3063656 \times \max(\text{CP} - 0.3, 0) + 0.50361921 \times \min(\text{ADMIT60D}, 2) \\ & + 0.025235717 \times \text{MOD\_APS} + 0.00050038706 \times \text{MOD\_APS}^2 + 0.2117423 \times \text{HRABX} \\ & - 0.2163988 \times \max(\text{HRABX} - 5, 0) + 0.4614184 \times \text{ICUPT} + 0.26031057 \times \text{LAX} \\ & + 0.75635204 \times \text{GAS} + 0.72442624 \times \text{MOTIL} + 0.53639167 \times \text{ALBUMIN}.\end{aligned}$$

APS, Acute Physiology Score; ICU, intensive care unit.

[a] Modeled as a polynomial function.

[b] Third quantile compared to first quantile.[26]

[c] Modeled as a spline function.

[d] Defined as cephalosporins, clindamycin, amoxicillin, or ampicillin.

risk of CDI. Future interventions might include administering fidaxomicin, nontoxigenic *C. difficile*, antitoxin monoclonal antibodies, or *C. difficile* vaccination. Previous analyses at BJH indicate that the attributable cost of CDI is $3,240 per case, the attributable length of stay is 2.8 days, the attributable mortality is 6%, and the attributable readmission rate is 19.3%.[1,2] On the basis of these data, at the 0.023 risk level, the potential cost savings of preventing 198 CDI cases would be $641,520; therefore, the maximum cost required for a "cost-neutral" intervention would be $159 per patient (4,032 patients identified as high risk for CDI). An intervention at the 0.023 level potentially would prevent 12 deaths, 38 readmissions, 18 nursing home or long-term care facility admissions, and 554 patient-days. At the 0.030 risk level, the potential cost savings associated with preventing 165 CDI cases would be $534,600; therefore, the maximum cost required for a cost-neutral intervention would be $185 per patient (2,886 patients identified as high risk for CDI). An intervention at the 0.030 level potentially would prevent 10 deaths, 32 readmissions, 15 nursing homes or long-term care facility admissions, and 462 patient-days.

These calculations are rudimentary and assume that the interventions will be 100% effective at preventing CDI. If the intervention is not 100% effective, then the cost per patient of the intervention would be higher and the decreases in patient morbidity and mortality would be lower. Conversely, the intervention may have a greater impact on reducing the incidence of CDI than estimated, even if the intervention is not 100% effective at reducing the risk of CDI for individual patients. The variable with the strongest association with probability of developing CDI was CDI pressure (Table 2).[13] A reduction in CDI incidence will subsequently decrease the probability of other patients developing CDI, because a reduction in the number of patients with CDI will result in reduced transmission of *C. difficile*. The reduction in costs from prevention of CDI is likely greater than that estimated as well. Reducing length of hospital stay by preventing CDI will increase the availability of beds for new admissions, decrease the number of patients in contact precautions, and reduce the frequency of an isolated patient "occupying" 2 beds in a semiprivate room.

To the best of our knowledge, only 2 other risk prediction models for CDI have been published. Garey et al[27] developed 2 risk indices for use among patients receiving broad-spectrum antibiotics, one based on odds ratio estimates from logistic regression and the other a simple yes/no index (AUC/C statistic of 0.733 in the development cohort and 0.712 in the validation cohort). These AUC/C-statistic values are lower than that presented in our analysis (0.88). Several explanations for this are possible: our risk index contained more variables than the Garey et al[27] indices, including CDI pressure (possibly improving model discrimination); we used nonparametric smoothing and spline techniques to deal with nonlinearity in continuous variables; our model was based on probability estimates (not odds ratios or yes/no classification); and our risk index was on a continuous scale.

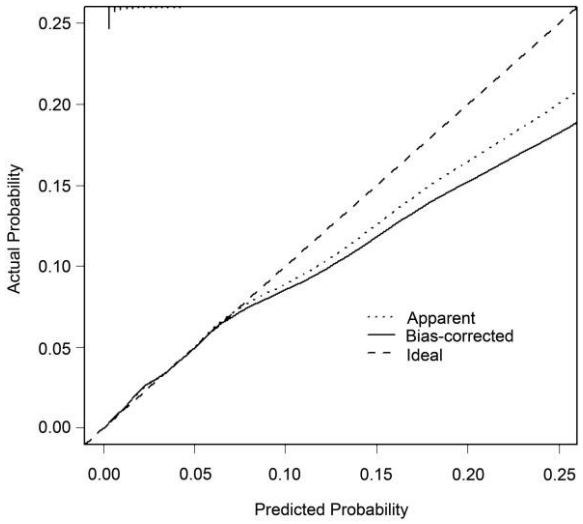


FIGURE 1. *Clostridium difficile* infection (CDI) risk prediction model calibration plot. The calibration plot graphically displays the agreement between the observed and predicted probabilities. The dashed line indicates perfect agreement. The lines below the dashed line indicate that the predicted probabilities are higher than the actual probabilities.

Tanner et al[28] developed a CDI risk prediction scale based on the Waterlow score, which was originally designed to identify patients at high risk for bedsores on admission to the hospital. A model with CDI risk factors plus the Waterlow score performed best, but its performance was not significantly different from that of the model with the Waterlow score alone (AUC/C statistic, 0.827). A key advantage of the Tanner et al[28] CDI risk prediction scale is ease of use for hospitals already calculating the Waterlow score; however, the Waterlow score may not be as widely used in the United States as the United Kingdom. Another limitation of this study is a potential lack of specificity among the variables included in the Waterlow score, such as mobility, below-waist fracture, and smoking. In addition, the investigators corrected Waterlow scores that were originally incorrectly calculated by healthcare workers in the validation cohort. This may have falsely increased the AUC in the validation cohort and decreases the generalizability of their model.

We excluded from the analyses 2 populations at high risk for CDI at our institution: patients with a history of CDI in the previous 60 days and patients who have undergone BMT or who have leukemia. These patient populations have unique characteristics that interfered with the ability of the model to predict patients at risk for CDI in the rest of the patient population. Patients with a history of CDI in the previous 60 days often are admitted to the hospital with CDI or develop CDI with minimal exposure. This adversely affected the model's ability to use these risk factors to identify other patients at risk for CDI. Hospitalized patients who have leukemia or who have undergone BMT typically have high severity of illness, prolonged hospital stays, and nearly universal exposure to medications that may predispose them to developing CDI,[33-36] decreasing the usefulness of these exposures for identifying patients in other populations at risk for CDI. Preliminary analyses identified these patients as almost universally at high risk for CDI; this was not true for any other patient population in our analyses, including general oncology patients. BMT patients and patients with a previous history of CDI are at high risk for CDI in most institutions, and there is no real practical benefit to including them in a risk prediction model.

There are several limitations to this study. First, the data used to develop the model were from 2003. There have not been any notable changes in CDI epidemiology or risk factors at our facility since 2003, indicating that the model should still be valid today. Second, the statistics used to develop the risk prediction model were complicated, and some of the variables used in the model are not readily available (eg, modified APS and CDI pressure), precluding the ability of a healthcare worker to determine a patient's CDI risk at the bedside. However, the CDI risk prediction index utilizes data available electronically in real time, allowing it to be automatically calculated each day a patient is in the hospital. This allows for the creation of a more robust risk prediction index and reduces barriers to determining a patient's CDI risk. Finally, this model may appear to require too many information technology resources or be too institution specific to be applied at other facilities. The information technology resources necessary for this project include an electronic clinical-data repository that is updated frequently, can be queried, and integrates different data sources or types (ie, demographic data and antimicrobial exposures). Developing this type of data repository also requires an information technology department with experience in application development. Al-

TABLE 3. Comparison of the *Clostridium difficile* Infection (CDI) Risk Prediction Model's Predicted Probability, Sensitivity, and Specificity at Different Levels of CDI Risk

| Predicted probability of developing CDI | Sensitivity | Specificity |
|---|---|---|
| 0.0065 | 0.9027 | 0.6835 |
| 0.0118 | 0.8024 | 0.7919 |
| 0.0173 | 0.7021 | 0.8507 |
| 0.0233 | 0.6018 | 0.8903 |
| 0.0302 | 0.5015 | 0.9222 |
| 0.0385 | 0.4012 | 0.9449 |
| 0.0557 | 0.3009 | 0.9705 |
| 0.0713 | 0.2006 | 0.9830 |
| 0.0997 | 0.1003 | 0.9920 |

NOTE. These probability values are deciles of the distribution of predicted probabilities among the patients.

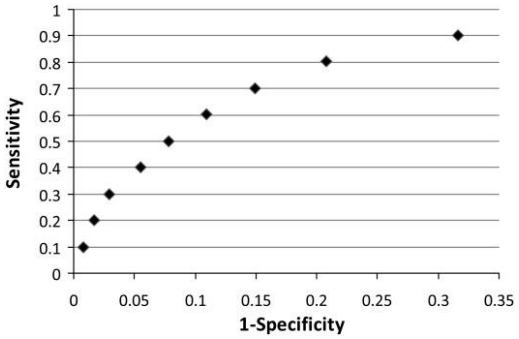


FIGURE 2. Receiver operating characteristic curve for the *Clostridium difficile* infection risk prediction model.

though not all healthcare facilities currently have the capability to combine data from multiple databases in such a fashion, the expansion of electronic medical records will increase the number of facilities where this approach is possible. We believe that clinical-data repositories will become more common in the future as clinicians, hospital administrators, and other healthcare professionals realize the benefits of readily available electronic data. In addition, any CDI risk prediction index will likely have to be individualized to the healthcare facility in question. *C. difficile* strain prevalence and practices that affect the risk of CDI, such as hand hygiene and contact precaution compliance and antimicrobial utilization, vary across healthcare facilities, and these differences will affect the predictive capability of the model. Once this model has been longitudinally validated, the methods for creating a CDI risk prediction index, rather than the CDI risk prediction index itself, would be applied at other healthcare facilities.

Validated methods to prevent CDI are needed. Current recommendations to prevent CDI, in general, are not supported by high-quality evidence, and CDI prevention efforts are not always successful.[23,37] One reason current recommendations may not always be effective is that prevention efforts are not initiated until after CDI has developed. Preventing CDI in patients at high risk for it may be an effective strategy. Our data suggest that risk prediction estimates generated from a model with readily available electronic data could identify hospital patients at highest risk for CDI. Theoretically, an intervention based on the results of this model may result in considerable financial savings, reduced length of stay, and prevention of hospital readmissions and long-term care facility admissions. The next step for this project is longitudinal validation of the model's predictive ability among hospitalized patients. If the predictive accuracy of the model proves as good in a longitudinal data set as in the development data set, then interventions could be designed to identify the highest-risk patients and to intervene before these patients develop CDI.

## ACKNOWLEDGMENTS

*Financial support.* This work was supported by grants from the Barnes-Jewish Hospital Foundation (00793-1106-01), the Centers for Disease Control and Prevention (UR8/CCU715087-06/1, 1U01C1000333-01), and the National Institutes of Health (1-R21-NR011362-01; 1K23AI065806-01A2).

*Potential conflicts of interest.* E.R.D. reports that he has performed research for Viropharma and Merck and has served as a consultant for Merck, Becton-Dickinson, Optimer, Meridian, and Steris. All other authors report no conflicts of interest relevant to this article.

Address correspondence to Erik R. Dubberke, MD, MSPH, Box 8051, 660 South Euclid, St. Louis, MO 63110 (edubberk@dom.wustl.edu).

Preliminary data were presented at the 18th Annual Scientific Meeting of the Society for Healthcare Epidemiology of America; Orlando, Florida; April 5–8, 2008 (Abstract 50).

## REFERENCES

1. Dubberke ER, Reske KA, Olsen MA, McDonald LC, Fraser VJ. Short- and long-term attributable costs of *Clostridium difficile*–associated disease in nonsurgical inpatients. *Clin Infect Dis* 2008; 46(4):497–504.
2. Dubberke ER, Butler AM, Reske KA, et al. Attributable outcomes of endemic *Clostridium difficile*–associated disease in nonsurgical patients. *Emerg Infect Dis* 2008;14(7):1031–1038.
3. Kyne L, Hamel MB, Polavaram R, Kelly CP. Health care costs and mortality associated with nosocomial diarrhea due to *Clostridium difficile. Clin Infect Dis* 2002;34(3):346–353.
4. Loo VG, Libman MD, Miller MA, et al. *Clostridium difficile:* a formidable foe. *CMAJ* 2004;171(1):47–48.
5. Loo VG, Poirier L, Miller MA, et al. A predominantly clonal multi-institutional outbreak of *Clostridium difficile*–associated diarrhea with high morbidity and mortality. *N Engl J Med* 2005; 353(23):2442–2449.
6. Muto CA, Pokrywka M, Shutt K, et al. A large outbreak of *Clostridium difficile*–associated disease with an unexpected proportion of deaths and colectomies at a teaching hospital following increased fluoroquinolone use. *Infect Control Hosp Epidemiol* 2005;26(3):273–280.
7. Pépin J, Valiquette L, Alary M-E, et al. *Clostridium difficile*–associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. *CMAJ* 2004;171(5):466–472.
8. Pépin J, Alary M-E, Valiquette L, et al. Increasing risk of relapse after treatment of *Clostridium difficile* colitis in Quebec, Canada. *Clin Infect Dis* 2005;40(11):1591–1597.
9. Pépin J, Valiquette L, Cossette B. Mortality attributable to nosocomial *Clostridium difficile*–associated disease during an epidemic caused by a hypervirulent strain in Quebec. *CMAJ* 2005; 173(9):1037–1042.
10. Dial S, Alrasadi K, Manoukian C, Huang A, Menzies D. Risk of *Clostridium difficile* diarrhea among hospital inpatients prescribed proton pump inhibitors: cohort and case-control studies. *CMAJ* 2004;171(1):33–38.
11. Dial S, Delaney JA, Barkun AN, Suissa S. Use of gastric acid–suppressive agents and the risk of community-acquired *Clostridium difficile*–associated disease. *JAMA* 2005;294(23):2989–2995.

12. Dial S, Delaney JA, Schneider V, Suissa S. Proton pump inhibitor use and risk of community-acquired *Clostridium difficile*–associated disease defined by prescription for oral vancomycin therapy. *CMAJ* 2006;175(7):745–748.
13. Dubberke ER, Reske KA, Olsen MA, et al. Evaluation of *Clostridium difficile*–associated disease pressure as a risk factor for *C difficile*–associated disease. *Arch Intern Med* 2007;167(10): 1092–1097.
14. Dubberke ER, Reske KA, Yan Y, Olsen MA, McDonald LC, Fraser VJ. *Clostridium difficile*–associated disease in a setting of endemicity: identification of novel risk factors. *Clin Infect Dis* 2007;45(12):1543–1549.
15. Gerding DN, Johnson S, Peterson LR, Mulligan ME, Silva J Jr. *Clostridium difficile*–associated diarrhea and colitis. *Infect Control Hosp Epidemiol* 1995;16(8):459–477.
16. Kyne L, Sougioultzis S, McFarland LV, Kelly CP. Underlying disease severity as a major risk factor for nosocomial *Clostridium difficile* diarrhea. *Infect Control Hosp Epidemiol* 2002;23(11):653–659.
17. Pépin J, Saheb N, Coulombe M-A, et al. Emergence of fluoroquinolones as the predominant risk factor for *Clostridium difficile*–associated diarrhea: a cohort study during an epidemic in Quebec. *Clin Infect Dis* 2005;41(9):1254–1260.
18. McDonald LC, Owings M, Jernigan DB. *Clostridium difficile* infection in patients discharged from US short-stay hospitals, 1996–2003. *Emerg Infect Dis* 2006;12(3):409–415.
19. Muto CA, Blank MK, Marsh JW, et al. Control of an outbreak of infection with the hypervirulent *Clostridium difficile* BI strain in a university hospital using a comprehensive "bundle" approach. *Clin Infect Dis* 2007;45(10):1266–1273.
20. Peterson CA, Calderon RL. Trends in enteric disease as a cause of death in the United States, 1989–1996. *Am J Epidemiol* 2003; 157(1):58–65.
21. Johnson S, Gerding DN, Olson MM, et al. Prospective, controlled study of vinyl glove use to interrupt *Clostridium difficile* nosocomial transmission. *Am J Med* 1990;88(2):137–140.
22. Kaatz GW, Gitlin SD, Schaberg DR, et al. Acquisition of *Clostridium difficile* from the hospital environment. *Am J Epidemiol* 1988;127(6):1289–1294.
23. Mayfield JL, Leet T, Miller J, Mundy LM. Environmental control to reduce transmission of *Clostridium difficile*. *Clin Infect Dis* 2000;31(4):995–1000.
24. McMullen KM, Zack J, Coopersmith CM, Kallef M, Dubberke E, Warren DK. Use of hypochlorite solution to decrease rates of *Clostridium difficile*–associated diarrhea. *Infect Control Hosp Epidemiol* 2007;28(2):205–207.
25. D'Agostino RB, Belanger AJ, Markson EW, Kelly-Hayes M, Wolf PA. Development of health risk appraisal functions in the presence of multiple indicators: the Framingham Study nursing home institutionalization model. *Stat Med* 1995;14(16):1757–1770.
26. Harrell FE Jr, Margolis PA, Gove S, et al. Development of a clinical prediction model for an ordinal outcome: the World Health Organization Multicentre Study of Clinical Signs and Etiological Agents of Pneumonia, Sepsis and Meningitis in Young Infants. *Stat Med* 1998;17(8):909–944.
27. Garey KW, Dao-Tran TK, Jiang ZD, Price MP, Gentry LO, DuPont HL. A clinical risk index for *Clostridium difficile* infection in hospitalised patients receiving broad-spectrum antibiotics. *J Hosp Infect* 2008;70(2):142–147.
28. Tanner J, Khan D, Anthony D, Paton J. Waterlow score to predict patients at risk of developing *Clostridium difficile*-associated disease. *J Hosp Infect* 2009;71(3):239–244.
29. Thom KA, Shardell MD, Osih RB, et al. Controlling for severity of illness in outcome studies involving infectious diseases: impact of measurement at different time points. *Infect Control Hosp Epidemiol* 2008;29(11):1048–1053.
30. D'Agostino RB, Griffith JL, Schmidt CH, Terrin N. Measures for evaluating model performance. In: *Proceedings of the Biometrics Section*. Alexandria, VA: American Statistical Association, 1997:253–258.
31. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem* 2008;54(1):17–23.
32. Efron B, Tibshirani R. *An introduction to the bootstrap*. New York: Chapman & Hall, 1993.
33. Bilgrami S, Feingold JM, Dorsky D, et al. Incidence and outcome of *Clostridium difficile* infection following autologous peripheral blood stem cell transplantation. *Bone Marrow Transplant* 1999; 23(10):1039–1042.
34. Chakrabarti S, Lees A, Jones SG, Milligan DW. *Clostridium difficile* infection in allogeneic stem cell transplant recipients is associated with severe graft-versus-host disease and non-relapse mortality. *Bone Marrow Transplant* 2000;26(8):871–876.
35. Tomblyn M, Gordon L, Singhal S, et al. Rarity of toxigenic *Clostridium difficile* infections after hematopoietic stem cell transplantation: implications for symptomatic management of diarrhea. *Bone Marrow Transplant* 2002;30(8):517–519.
36. Yuen K-Y, Woo PCY, Liang RHS, et al. Clinical significance of alimentary tract microbes in bone marrow transplant recipients. *Diagn Microbiol Infect Dis* 1998;30(2):75–81.
37. Dubberke ER, Gerding DN, Classen D, et al. Strategies to prevent *Clostridium difficile* infections in acute care hospitals. *Infect Control Hosp Epidemiol* 2008;29(suppl 1):S81–S92.