# Identification of patients at high risk for *Clostridium difficile* infection: development and validation of a risk prediction model in hospitalized patients treated with antibiotics

**C. H. van Werkhoven[1], J. van der Tempel[2], R. Jajou[5], S. F. T. Thijsen[4], R. J. A. Diepersloot[4], M. J. M. Bonten[1,3], D. F. Postma[1] and J. J. Oosterheert[2]**

1) *Julius Center for Health Sciences and Primary Care,* 2) *Department of Internal Medicine and Infectious Diseases,* 3) *Department of Medical Microbiology, University Medical Center,* 4) *Department of Medical Microbiology and Immunology, Diakonessenhuis, Utrecht and* 5) *Health Sciences, VU University Amsterdam, The Netherlands*

## Abstract

To develop and validate a prediction model for *Clostridium difficile* infection (CDI) in hospitalized patients treated with systemic antibiotics, we performed a case–cohort study in a tertiary (derivation) and secondary care hospital (validation). Cases had a positive *Clostridium* test and were treated with systemic antibiotics before suspicion of CDI. Controls were randomly selected from hospitalized patients treated with systemic antibiotics. Potential predictors were selected from the literature. Logistic regression was used to derive the model. Discrimination and calibration of the model were tested in internal and external validation. A total of 180 cases and 330 controls were included for derivation. Age >65 years, recent hospitalization, CDI history, malignancy, chronic renal failure, use of immunosuppressants, receipt of antibiotics before admission, nonsurgical admission, admission to the intensive care unit, gastric tube feeding, treatment with cephalosporins and presence of an underlying infection were independent predictors of CDI. The area under the receiver operating characteristic curve of the model in the derivation cohort was 0.84 (95% confidence interval 0.80–0.87), and was reduced to 0.81 after internal validation. In external validation, consisting of 97 cases and 417 controls, the model area under the curve was 0.81 (95% confidence interval 0.77–0.85) and model calibration was adequate (Brier score 0.004). A simplified risk score was derived. Using a cutoff of 7 points, the positive predictive value, sensitivity and specificity were 1.0%, 72% and 73%, respectively. In conclusion, a risk prediction model was developed and validated, with good discrimination and calibration, that can be used to target preventive interventions in patients with increased risk of CDI.



**Keywords:** Antibiotic-associated diarrhoea, antibiotics, *Clostridium difficile*, hospital infection control, primary prevention, risk prediction
**Original Submission:** 31 July 2014; **Revised Submission:** 7 April 2015; **Accepted:** 7 April 2015
Editor: I. Gyssens
**Article published online:** 16 April 2015

Results of the derivation part of the study were presented at the 2013 European Congress of Clinical Microbiology and Infectious Diseases 2013
**Corresponding author:** C. H. van Werkhoven, University Medical Centre Utrecht, Julius Centre for Health Sciences, PO Box 85500, 3508 GA Utrecht, The Netherlands
**E-mail:** c.h.vanwerkhoven@umcutrecht.nl

## Introduction

*Clostridium difficile* infection (CDI) is an important cause of antibiotic-associated diarrhoea and infectious colitis. CDI has shown increasing incidence and mortality rates worldwide since the 1990s [1,2] and is associated with increased healthcare costs [1,3,4]. Disruption of the intestinal bacterial flora by antibiotics is recognized as the most important pathogenic mechanism and risk factor of CDI [5–8].



http://dx.doi.org/10.1016/j.cmi.2015.04.005

Antibiotic stewardship programs, probiotics and infection control policies focussing on prevention of spread seem to be able to reduce CDI incidence [9–14]. However, applying these interventions to all hospitalized patients may not be cost-effective in all settings. For example, in hospitals with a low incidence of CDI, the number needed to treat (NNT) to prevent one case of CDI might be unfavourably high. In that case, if possible, preventive interventions could be targeted to those at highest risk of CDI, thereby reducing the NNT.

Despite good discriminative power in previously reported risk prediction models [15–17], none of these can be easily implemented in clinical practice because variables were included that are not generally measured, such as nutritional status, or that are unknown when prediction is warranted, such as total length of intensive care unit (ICU) stay. Moreover, two of these studies included all hospitalized patients, regardless of receipt of antibiotics. Because systemic antibiotic treatment is the most important pathogenesis for CDI, we developed and validated a model to identify patients at the highest risk of CDI who had received at least one dose of systemic antibiotics during their hospitalization.

## Methods

### Design and setting

For derivation of the model, we performed a retrospective case–cohort study at the University Medical Center Utrecht, a 1042-bed tertiary care hospital in the Netherlands, with about 28 000 clinical hospitalizations annually. We included patients who were admitted between January 2005 and December 2011. No outbreaks of CDI were experienced in this hospital during the study period. External validation of the derived model was performed in the Diakonessenhuis hospital in patients admitted between November 2010 and May 2013, using the same retrospective design and the same inclusion and exclusion criteria. This study period was chosen because data on antibiotic use before November 2010 were not electronically available. The Diakonessenhuis hospital is a 451-bed secondary care teaching hospital, with about 26 000 clinical hospitalizations annually. From January to June 2011 the Diakonessenhuis hospital experienced an outbreak of *C. difficile*. During this period, the incidence of CDI was increased threefold compared to the previous 3 years. Most CDI episodes occurred in one of the internal departments, the pulmonology department and the geriatric department, and three were found in the ICU. The outbreak was managed using isolation measures, intensified cleaning of paramedical devices and of affected rooms, and an antibiotic stewardship program, restricting use of cephalosporins, clindamycin and fluoroquinolones, for the involved departments. Both hospitals provide a broad range of medical and surgical services.

### Selection of cases and controls

Inclusion criteria were age over 16 years, hospitalization for at least 48 hours and receipt of systemic antibiotic therapy during hospitalization. Patients were excluded when only low-dose erythromycin was prescribed as prokinetic therapy because this is not associated with CDI [18], or when suspicion of CDI was present before initiation of in-hospital antibiotic treatment because in these patients prediction is no longer warranted. However, when patients were admitted for CDI within 28 days of a previous admission, the previous admission was considered for inclusion. In the derivation cohort, patients with multiple episodes of CDI were only included for the first eligible admission to prevent overestimation of effect estimates due to similarity of predictors within episodes of one patient.

Cases were selected based on the detection of *C. difficile* in clinical stool samples. As a result of differences in local testing practices, cases for the derivation cohort were selected on the basis of a positive *C. difficile* toxin assay (ImmunoCard, toxins A and B), whereas in the validation cohort CDI was defined as a positive *C. difficile* toxin assay (ImmunoCard; dubious test results were confirmed by the GeneXpert PCR test), a positive PCR result, or *C. difficile* cultured from stool samples. The clinical indication for testing was suspicion of infectious diarrhoea associated with antibiotic use in both hospitals. In the validation cohort, the toxin assay was the standard test, and upon receipt of negative results with sustained suspicion of CDI, PCR was performed. Screening procedures were not in place at either hospital. During the outbreak period in the validation cohort, testing was performed in case of diarrhoea, irrespective of association with antibiotics.

For the derivation cohort, we aimed for a case–control ratio of 1:2, and controls were randomly selected from the Utrecht Patient Oriented Database, an infrastructure of relational databases containing clinical and administrative data for all patients treated at the University Medical Center Utrecht, in accordance with current regulations concerning privacy and ethics as described elsewhere [19]. First, admission and pharmacy data were merged to only select admissions in which systemic antibiotics were provided. Next, controls were randomly selected from this data set. Patients with multiple admissions had a higher probability of being selected; however, selection of multiple admissions of a single patient was prohibited because most parameters are expected to be the same in the different episodes and this may cause overestimation of the odds ratios.

For the validation cohort, we aimed for a case–control ratio of 1:4, which was chosen to increase precision, given the lower number of cases in the validation cohort. Controls were

 

randomly selected from all hospitalized patients receiving antibiotic therapy during the study period, using the hospital admission and pharmacy registries. Selection of multiple admissions of the same patient was allowed in the validation cohort in order to allow model validation in a real-life situation.

### Data collection

Clinical data were manually extracted from the electronic patient records. For the purpose of prediction, only data that were available on or before the first day of in-hospital antibiotic therapy were recorded. In the derivation sample, the severity of comorbidities was classified using an adapted McCabe–Jackson comorbidity severity index [20], limited to three categories: no comorbidity, nonfatal comorbidities and fatal comorbidities. Exposure to CDI was pragmatically defined as positive if up to 2 weeks before the patient's antibiotics start date, a CDI-positive patient had been present at the same hospital ward. This definition was chosen because the prediction model has to be usable in clinical practice, taking into account that more misclassifications are expected for episodes that occurred more than 2 weeks ago. Antibiotic use before admission was defined as any systemic antibiotic use in the 14 days prior to the current admission. This time period was chosen because antibiotic use more than 2 weeks before admission is generally poorly documented, and because these data are routinely collected by the clinician, making it easier to apply the model in clinical practice.

Permission for the study was obtained from the ethics review board of both hospitals, and a waiver for informed consent was granted.

### Statistical analysis

Multiple logistic regression was performed, starting with all potential predictors. Stepwise backward elimination was used to determine the best predictive model, excluding variables with a p value of >0.05 using the likelihood ratio test. Internal validation of the model was performed using bootstrapping techniques, and bias-corrected estimates of the regression coefficients were derived by multiplying the original coefficients with the slope of the calibration curve. Discriminative capacity of the model in the derivation and validation cohort was assessed using the area under the receiver operating characteristic curve (AUC). Absolute predicted probabilities were corrected for the case–control ratio and sampling fraction, as described elsewhere [21]. Calibration of the model was tested by plotting the predicted probability vs. the observed outcome. A risk score was derived from the regression coefficients, multiplied by 2 and subsequently rounded. Sensitivity, specificity and positive and negative predictive values were estimated for different risk levels. Subgroup analysis was performed among patients during the outbreak period in the validation cohort. Analysis was performed in R software (version 2.15.2), including packages rms (version 3.6.3) and pROC (version 1.5.4) [22].

## Results

### Patients

Characteristics of the study populations are listed in Table 1. Compared to controls, CDI patients were generally older, had more comorbidities, had received more antibiotics before admission and more frequently received immunosuppressants and gastric acid suppressants.

### Model derivation

In the derivation study, 180 CDI patients and 319 controls were included. CDI developed during hospitalization in 146 cases and within 28 days of discharge in 34 cases. Reasons for exclusion of patients are summarized in Fig. 1. During the study period, the incidence of CDI was 5.2 (95% confidence interval (CI) 4.5–5.9) per 1000 admissions in patients receiving systemic antibiotic therapy.

Variables in the final model, including bias-corrected estimates, are presented in Table 2. Both the model and the risk score showed good discrimination, with AUC values of 0.84 (95% confidence interval (CI) 0.80–0.87) and 0.83 (95% CI 0.80–0.87), respectively (Fig. 2). The bias-corrected AUC, after internal validation, was 0.81.

### Model validation

In the validation study, 97 CDI patients and 419 controls were included. CDI developed during hospitalization in 77 cases and within 28 days of discharge in 20 cases. Of the patients, 95 were identified by a positive toxin assay and two were PCR positive only. Reasons for exclusion of patients are summarized in Fig. 1. During the study period, the incidence of CDI was 4.0 (95% CI 3.2–4.8) per 1000 admissions in patients receiving systemic antibiotic therapy. During and outside the outbreak period, the incidence was 8.8 (95% CI 6.2–11.6) and 2.8 (95% CI 2.1–3.6) per 1000 admissions, respectively.

The AUC of the model and risk score were 0.81 (95% CI 0.77–0.85) and 0.81 (95% CI 0.77–0.86), respectively (Fig. 2). The predicted risk was generally well calibrated with the observed outcome, with a Brier score of 0.004 (a Brier score of 0 represents perfect calibration). For those with the highest risk, the predicted risk was somewhat underestimated (Fig. 2). This group consisted of two controls, which represented less than 1% of the base population.

Positive predictive values, sensitivities and specificities of the risk score in the validation cohort are presented in Table 3.

**TABLE 1. Population characteristics**

| Characteristic | Derivation cohort | | Validation cohort | |
|---|---|---|---|---|
| | CDI cases ($n$ = 180) | Controls ($n$ = 319) | CDI cases ($n$ = 97) | Controls ($n$ = 419) |
| Male gender | 102 (56.7%) | 182 (57.1%) | 46 (47.4%) | 169 (40.3%) |
| Age (years) | 66 (54–76) | 61 (48–71) | 77 (68–85) | 71 (60–81) |
| Medical history | | | | |
| Previous (<90 days) hospitalization | 76 (42.2%) | 72 (22.6%) | 50 (51.5%) | 55 (13.1%) |
| Cardiovascular disease | 44 (24.4%) | 102 (32.0%) | 40 (41.2%) | 88 (21.0%) |
| Malignancy | 49 (27.2%) | 53 (16.6%) | 19 (19.6%) | 61 (14.6%) |
| Diabetes mellitus | 24 (13.3%) | 40 (12.5%) | 11 (11.3%) | 73 (17.4%) |
| Chronic renal failure | 31 (17.2%) | 15 (4.7%) | 2 (2.1%) | 3 (0.7%) |
| Inflammatory bowel disease | 3 (1.7%) | 7 (2.2%) | 3 (3.1%) | 6 (1.4%) |
| History of CDI | 5 (2.8%)[a] | 1 (0.3%) | 2 (2.1%) | 0 |
| Medication use | | | | |
| Immunosuppressants (past 90 days) | 60 (33.3%) | 44 (13.8%) | 11 (11.3%) | 18 (4.3%) |
| Gastric acids suppressants (past 14 days) | 82 (45.6%) | 122 (38.2%) | 58 (59.8%) | 133 (31.7%) |
| Antibiotics before hospitalization (14 days) | 61 (33.9%) | 40 (12.5%) | 37 (38.1%) | 67 (16.0%) |
| Admission | | | | |
| Surgical admission | 53 (29.4%) | 204 (63.9%) | 22 (22.7%) | 224 (53.5%) |
| Total length of stay | 26.5 (14–67.3) | 8 (5–16) | 22 (10–32) | 6 (4–11) |
| Length of stay until first *Clostridium difficile* test positive | 18.5 (9–34) | NA | 9 (3–18) | NA |
| ICU admission[b] | 28 (15.6%) | 23 (7.2%) | 18 (18.6%) | 19 (4.5%) |
| Tube feeding[b] | 20 (11.1%) | 7 (2.2%) | 7 (7.2%) | 4 (1.0%) |
| CDI exposure at time of prediction | 24 (13.3%) | 19 (6.0%) | NA | NA |
| Initial antibiotic treatment | | | | |
| β-Lactams | 39 (21.7%) | 85 (26.6%) | 26 (26.8%) | 117 (27.9%) |
| Cephalosporins | 102 (56.7) | 172 (53.9) | 60 (61.9%) | 238 (56.8%) |
| Macrolides | 8 (4.4%) | 26 (8.2%) | 13 (13.4%) | 26 (6.2%) |
| Quinolones | 24 (13.3%) | 35 (11.0%) | 7 (7.2%) | 23 (5.5%) |
| Carbapenems | 13 (7.2%) | 4 (1.3%) | 0 | 4 (1.0%) |
| Other antibiotics | 48 (26.7%) | 71 (22.3%) | 12 (12.4%) | 70 (16.7%) |
| Indication for antibiotics | | | | |
| Underlying infection | 132 (73.3%) | 120 (37.6%) | 86 (88.7%) | 274 (65.4%) |
| Respiratory tract infection | 53 (29.4%) | 39 (12.2%) | 39 (40.2%) | 87 (20.8%) |
| Urinary tract infection | 30 (16.7%) | 23 (7.2%) | 14 (14.4%) | 71 (16.9%) |
| Other infections | 49 (27.2%) | 58 (18.2%) | 36 (37.1%) | 120 (28.6%) |
| Prophylactic treatment | 48 (26.7%) | 199 (62.4%) | 11 (11.3%) | 145 (34.6%) |

Data are presented as $n$ (%) or median (interquartile range).
CDI, *Clostridium difficile* infection; ICU, intensive care unit; NA, not applicable.
[a]For CDI cases in the derivation cohort, these include CDI episodes that had occurred before 2005, in another hospital or in a noneligible previous admission.
[b]Until initiation of antibiotic treatment.

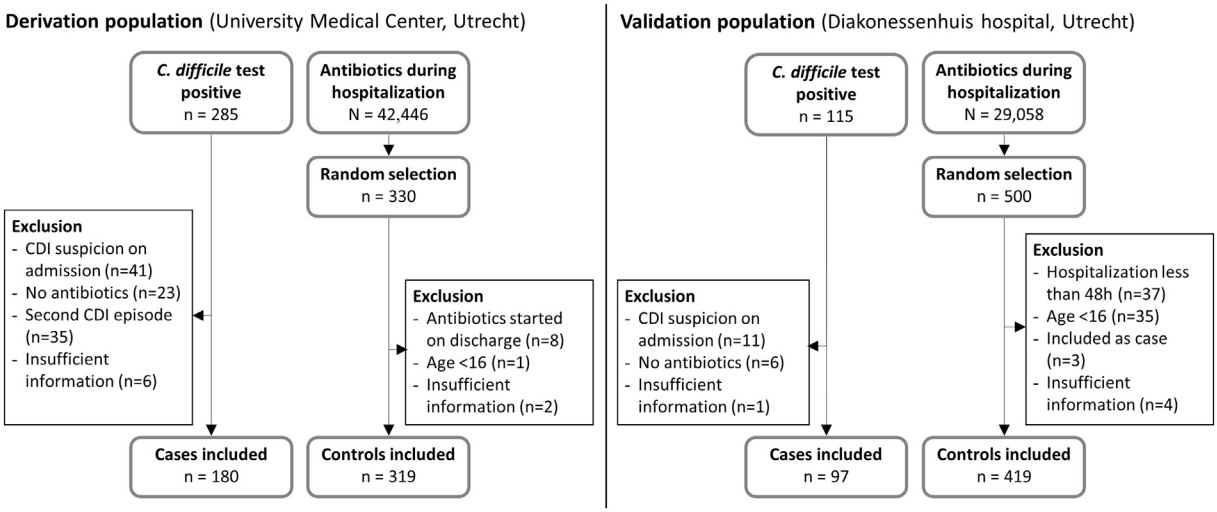


**FIG. 1.** Inclusion flowcharts of cases and controls for derivation and validation studies.

### Sensitivity analyses

For patients with the first in-hospital antibiotic dose within the outbreak period, the model had an AUC of 0.86 (95% CI 0.79–0.92), vs. 0.80 (95% CI 0.74–0.86) outside the outbreak period (p 0.20). Calibration of the model during the outbreak period showed an underestimation of the predicted probability in the high risk patients (Brier score 0.009).

Additionally, a sensitivity analysis was performed using a more liberal p value of 0.2 to select predictors because it has been debated whether higher p values lead to more generalizable

**TABLE 2. Parameters in the final model**

| | Model parameters | | Bias corrected[b] | | |
|---|---|---|---|---|---|
| **Predictor[a]** | **Beta** | **OR (95% CI)** | **Beta** | **OR (95% CI)** | **Risk score** |
| Age >65 years | 0.568 | 1.77 (1.11–2.80) | 0.456 | 1.58 (0.99–2.51) | 1 |
| Previous hospitalization[1] | 0.536 | 1.71 (1.04–2.81) | 0.430 | 1.54 (0.94–2.53) | 1 |
| History of CDI | 3.510 | 33.6 (2.82–397.3) | 2.819 | 16.8 (1.41–198.8) | 7 |
| Malignancy[2] | 0.574 | 1.78 (1.01–3.12) | 0.461 | 1.59 (0.90–2.79) | 1 |
| Chronic renal failure[3] | 1.104 | 3.02 (1.43–6.38) | 0.887 | 2.43 (1.15–5.13) | 2 |
| Use of immunosuppressants[4] | 0.764 | 2.15 (1.24–3.71) | 0.613 | 1.85 (1.07–3.19) | 2 |
| Antibiotics before admission[5] | 1.040 | 2.83 (1.62–4.96) | 0.835 | 2.31 (1.32–4.03) | 2 |
| Receipt of cephalosporins | 0.896 | 2.45 (1.49–4.02) | 0.719 | 2.05 (1.25–3.37) | 2 |
| Underlying infection[6] | 1.020 | 2.77 (1.65–4.66) | 0.819 | 2.27 (1.35–3.81) | 2 |
| Nonsurgical admission | 0.998 | 2.71 (1.59–4.64) | 0.801 | 2.23 (1.30–3.81) | 2 |
| ICU admission | 0.894 | 2.45 (1.20–5.00) | 0.718 | 2.05 (1.00–4.18) | 2 |
| Tube feeding | 1.403 | 4.07 (1.46–11.31) | 1.127 | 3.09 (1.11–8.59) | 3 |

OR, odds ratio; CI, confidence interval; CDI, *Clostridium difficile* infection; ICU, intensive care unit.
[a]Predictor values were recorded as available on the first day of in-hospital antibiotic administration. Definitions of predictors are as follows: 1) within the last 90 days before current admission; 2) any malignancy requiring chemotherapy or radiotherapy in the past 5 years; 3) with indication for dialysis or renal transplantation; 4) within the past 90 days; 5) with in the past 14 days before admission; and 6) receiving antibiotics for an infection, as opposed to receiving prophylactic antibiotic treatment. Variables not in the model include severity of comorbidities, CDI exposure, use of gastric acid suppressants within the past 14 days, classes of in-hospital antibiotic treatment (β-lactams, macrolides, quinolones, carbapenems and others).
[b]After internal validation, the regression coefficients (beta) were multiplied by 0.803 to arrive at the bias-corrected effect estimates.

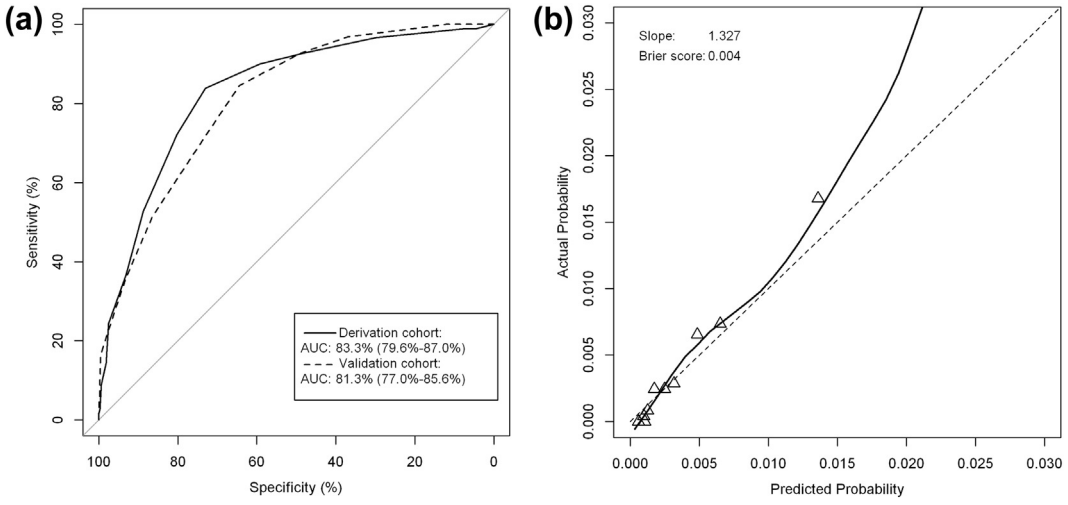


**FIG. 2.** Performance. (a) ROC curve of simplified risk score in derivation and validation cohort. (b) Calibration of predicted and observed probability of CDI in external validation, zoomed in to show 99.5% of the data. ROC, receiver operating characteristic; AUC, area under the ROC curve.

**TABLE 3. Model performance at different cutoff points using simplified risk score**

| | Total study period | | | | | Outbreak period | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Score** | **Positive (%)[a]** | **PPV (%)** | **NPV (%)** | **Sensitivity (%)** | **Specificity (%)** | **Positive (%)[a]** | **PPV (%)** | **NPV (%)** | **Sensitivity (%)** | **Specificity (%)** |
| 3 | 88 | 0.5 | 100 | 100 | 12 | 90 | 1.0 | 100 | 100 | 10 |
| 4 | 63 | 0.6 | 100 | 97 | 37 | 62 | 1.4 | 100 | 100 | 39 |
| 5 | 51 | 0.7 | 99.9 | 93 | 49 | 53 | 1.6 | 99.9 | 95 | 47 |
| 6 | 36 | 0.9 | 99.9 | 85 | 64 | 33 | 2.4 | 99.9 | 90 | 67 |
| 7 | 27 | 1.0 | 99.8 | 72 | 73 | 22 | 2.9 | 99.7 | 73 | 79 |
| 8 | 14 | 1.5 | 99.8 | 52 | 86 | 9 | 4.8 | 99.5 | 49 | 91 |
| 9 | 7 | 2.0 | 99.7 | 36 | 93 | 3 | 9.6 | 99.4 | 34 | 97 |
| 10 | 4 | 2.7 | 99.7 | 27 | 96 | 0 | 100 | 99.3 | 24 | 100 |

Each row that represents NPV, PPV, sensitivity and specificity when a score equalled or exceeded the cutoff point is considered high risk.
PPV, positive predictive value; NPV, negative predictive value.
[a]Percentage of patients with risk score above the cutoff point.

models [23]. In this analysis, receipt of carbapenems remained in the model after stepwise elimination. However, overall model performance did not improve, both in the derivation and in the validation cohort.

## Discussion

We developed and validated a model to predict the risk of CDI in hospitalized patients treated with systemic antibiotics. Both the model and the simplified risk score showed good discrimination, illustrated by an AUC of 0.81 in internal and external validation. Discrimination of the simplified risk score was similar to the full model in all settings. The parameters in the model are plausible, given the epidemiology of *C. difficile*. This model can be used to target risk-reducing interventions to patients at high risk of CDI. For example, if all patients in our validation cohort would undergo the restrictive antibiotic stewardship program, with an efficacy of 54% to prevent CDI [10], the NNT to prevent one CDI case would be 463 (the inverse of CDI incidence divided by the efficacy). By limiting the intervention to patients with a risk score of 7 or higher, assuming the same efficacy of the intervention, the NNT would decrease to 185 (the inverse of the positive predictive value divided by the reduction percentage). With this approach, up to 39% of the CDI cases could still be prevented by the intervention (sensitivity times efficacy), while the intervention would be restricted to only 27% of the patients (number of patients scoring positive; Table 3). Depending on the costs and disadvantages of the intervention, this could be a more acceptable approach than treating the whole population.

The incidence of CDI in our study population was low compared to some other countries, with incidences up to 10 per 1000 admissions in all hospitalized patients [15], and the previously published incidence in the derivation hospital was relatively low compared to other Dutch hospitals (on average, 14.7 per 10 000 admissions) [24] (http://www.rivm.nl/Documenten_en_publicaties/Algemeen_Actueel/Uitgaven/Infectieziekten/CDiffNL/Seventh_Annual_Report_of_the_National_Reference_Laboratory_for_Clostridium_difficile_May_2012_to_May_2013_and_results_of_the_sentinel_surveillance). During the study periods, incidence in the derivation cohort was approximately 11 per 10 000 admissions (independent of antibiotic use), with minor fluctuations; in the validation cohort, this was approximately 15 per 10 000 admissions, varying from 11 outside the outbreak period to 32 during the outbreak period (data not shown). As a result of these low incidences, the positive predictive values, even in the group with high risk scores, remained low. In patients receiving systemic antibiotics, incidence was higher in the derivation study compared to the validation study, presumably because the derivation study took place in a tertiary care hospital.

The outbreak of *C. difficile* in the Diakonessenhuis hospital may have influenced the results of the validation. However, discrimination of the model was similar during the outbreak period compared to the nonoutbreak setting. This suggests that although the model has been developed in a low-incidence setting, it is generalizable to a setting with higher incidence. It should be noted that the incidence during the outbreak period was still relatively low. Therefore, generalizability of the model to high incidence settings should be assessed before implementation of the model in clinical practice.

CDI exposure did not improve model performance in our study. This may be due to the pragmatic and (consequently) imprecise definition of CDI exposure, or to the low incidence of CDI in the derivation study. A previously proposed sensitive method to measure CDI exposure uses admission data and CDI test results in a computerized model [15]. However, this requires adaptation of the model to the infrastructure of every individual hospital and limits applicability of the model to hospitals where these data are electronically available. Moreover, differences in testing and isolation strategies and of patient migration patterns will limit the generalizability of CDI exposure effects. It has also been suggested that infection pressure partly represents the physician's alertness instead of real infection pressure [25].

Colonization with *C. difficile* at admission has been suggested as a risk factor for CDI [26]. Currently no routine screening is performed; therefore, as a result of the retrospective design of our study, we were not able to assess this predictor. Colonization can only be a useful predictor if routine screening is implemented. However, because *C. difficile* colonization and infection share the most important risk factors [26], we expect the added value to be limited.

To our knowledge, three models for the prediction of CDI risk have been previously published. The first study examined a risk score in patients receiving broad-spectrum antibiotics who were admitted for at least 48 hours. Discrimination was moderate, with an AUC of 0.712 in the validation cohort [16]. One limitation of this study is that the total length of ICU stay, up to the date of the stool test for cases and up to the discharge date for controls, was one of the predictors in the model. Although this is a known risk factor of CDI, prediction is warranted at a time when length of stay is not yet known. In contrast, in our study, we only collected data that were known at the start of in-hospital antibiotic treatment. The second study investigated performance of the Waterlow score in predicting CDI occurrence. This score is routinely used in UK hospitals to predict the risk of acquiring decubitus ulcers, and it also appeared to

 

predict CDI, with an AUC of 0.837 [17]. Because the evaluation was performed in a cohort with only 20 cases of CDI, results should be interpreted with caution. The third study used electronically available data and included a measure for CDI pressure. The model has an AUC of 0.88 [15]. As mentioned by the authors, electronic information sources are prone to be too institution specific, thus limiting generalizability. Furthermore, the last two risk scores have not been evaluated in a population restricted to patients receiving systemic antibiotics. Because this is known to be the most important factor in the pathogenesis of CDI, validation of CDI risk prediction models in a population of patients receiving systemic antibiotic treatment is desirable. Recently, a fourth prognostic model has been published [27]. However, development of diarrhoea during the admission was one of the predictors, and the model should therefore be considered diagnostic.

Several limitations of our study need to be addressed. We assumed that patients did not have CDI when they were not tested. Although it is common practice to test for CDI in patients with antibiotic-associated hospital-acquired diarrhoea, testing may have been omitted in patients with mild or nonspecific symptoms. Additionally, the sensitivity of the toxin test is limited, although the interpretation of a positive nucleic acid amplification test and negative toxin test is also debated [28]. As a result, we may have missed CDI cases, especially the less severe, which may have caused over- or underestimation of the model performance. Also, because we based the definition of CDI mainly on a positive toxin test, some cases may have been falsely positive. However, in our hospitals, the *Clostridium* toxin test is only performed when infectious diarrhoea is suspected, and we therefore assume the number of false-positive results will be limited. Differences in diagnostic tests used in the hospitals will have a limited effect, given that only two additional cases were identified with PCR in the validation cohort. Yet as the diagnostic algorithm used in the validation hospital is more compatible with the European guideline on diagnosing CDI [29], differences will be in favour of the validation study, which showed similar model performance compared to the derivation study. This suggests that the differences in diagnostic algorithms, as they were present, did not affect the results of the model to a large degree. Nevertheless, prospective evaluation of the risk score using a combination of sensitive *C. difficile* detection methods and objective clinical criteria of CDI may be necessary.

We have excluded a substantial part of CDI cases in our derivation cohort. Most of these cases were excluded because the participants were already suspected of having CDI at admission and tested positive within 48 hours. Prediction is irrelevant if the outcome is already present, and inclusion of these patients would have biased the results. This was most frequent in the derivation cohort, probably because of the tertiary care setting. In the derivation cohort, we also excluded subsequent CDI episodes within the same patient. Including multiple episodes that occur in the same patient might result in more homogeneous data within the cases, which in turn could lead to overestimation of the predictor effects. Last of all, a number of CDI patients in the derivation cohort were excluded because they did not receive antibiotics during the admission and therefore were not part of our study domain. Most of these patients received chemotherapy (data not shown), which is a known factor predisposing to CDI infection. Validation of the model in this patient group may be warranted in future investigations.

Because this is a retrospective study, information bias may have occurred. Having a history of CDI is prone to bias because physicians may be more inclined to report this in patients with a current episode of CDI. On the other hand, it is not surprising that history of CDI is a strong predictor because these are susceptible patients, and CDI is associated with high recurrence rates [30]. For other variables in the model, information bias is expected to be limited. A known aetiologic factor that was not significant in our derivation study is use of gastric acid suppressants. The difference in gastric acid use between cases and controls was relatively low in our derivation cohort, which may explain this finding.

In conclusion, we developed and validated a model to predict the occurrence of CDI in hospitalized patients who receive systemic antibiotic treatment. The model showed good discrimination, which was maintained in external validation. If preventive interventions have a negative cost–benefit balance in the entire population, the risk prediction score could be a valuable tool to select a population at increased risk of CDI. Future research could compare different CDI control strategies—for example, discontinuation of gastric acid suppressants, use of probiotics or discouragement of broad-spectrum antibiotics, including risk stratification and restriction of the intervention to high-risk populations.

## Transparency declaration

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

## References

[1] 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;34:346–53.

[2] Redelings MD, Sorvillo F, Mascola L. Increase in *Clostridium difficile*–related mortality rates, United States, 1999–2004. Emerg Infect Dis 2007;13:1417–9.
[3] O'Brien JA, Lahue BJ, Caro JJ, Davidson DM. The emerging infectious challenge of *Clostridium difficile*–associated disease in Massachusetts hospitals: clinical and economic consequences. Infect Control Hosp Epidemiol 2007;28:1219–27.
[4] 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;46:497–504.
[5] 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;45:1543–9.
[6] Howitt JR, Grace JW, Schaefer MG, Dolder C, Cannella C, Schaefer RS. *Clostridium difficile*–positive stools: a retrospective identification of risk factors. Am J Infect Control 2008;36:488–91.
[7] Kuntz JL, Chrischilles EA, Pendergast JF, Herwaldt LA, Polgreen PM. Incidence of and risk factors for community-associated *Clostridium difficile* infection: a nested case–control study. BMC Infect Dis 2011;11:194.
[8] Owens Jr RC, Donskey CJ, Gaynes RP, Loo VG, Muto CA. Antimicrobial-associated risk factors for *Clostridium difficile* infection. Clin Infect Dis 2008;46(Suppl. 1):S19–31.
[9] Aldeyab MA, Kearney MP, Scott MG, et al. An evaluation of the impact of antibiotic stewardship on reducing the use of high-risk antibiotics and its effect on the incidence of *Clostridium difficile* infection in hospital settings. J Antimicrob Chemother 2012;67:2988–96.
[10] Feazel LM, Malhotra A, Perencevich EN, Kaboli P, Diekema DJ, Schweizer ML. Effect of antibiotic stewardship programmes on *Clostridium difficile* incidence: a systematic review and meta-analysis. J Antimicrob Chemother 2014;69:1748–54.
[11] Goldenberg JZ, Ma SS, Saxton JD, et al. Probiotics for the prevention of *Clostridium difficile*–associated diarrhea in adults and children. Cochrane Database Syst Rev 2013;5:CD006095.
[12] Rebmann T, Carrico RM, Association for Professionals in Infection Control and Epidemiology. Preventing *Clostridium difficile* infections: an executive summary of the Association for Professionals in Infection Control and Epidemiology's elimination guide. Am J Infect Control 2011;39:239–42.
[13] Talpaert MJ, Gopal RG, Cooper BS, Wade P. Impact of guidelines and enhanced antibiotic stewardship on reducing broad-spectrum antibiotic usage and its effect on incidence of *Clostridium difficile* infection. J Antimicrob Chemother 2011;66:2168–74.
[14] Vonberg RP, Kuijper EJ, Wilcox MH, et al. Infection control measures to limit the spread of. Clostridium difficile Clin Microbiol Infect 2008;14(Suppl. 5):2–20.
[15] Dubberke ER, Yan Y, Reske KA, et al. Development and validation of a *Clostridium difficile* infection risk prediction model. Infect Control Hosp Epidemiol 2011;32:360–6.
[16] 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;70:142–7.
[17] 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;71:239–44.
[18] Nguyen NQ, Ching K, Fraser RJ, Chapman MJ, Holloway RH. Risk of *Clostridium difficile* diarrhoea in critically ill patients treated with erythromycin-based prokinetic therapy for feed intolerance. Intensive Care Med 2008;34:169–73.
[19] ten Berg MJ, Huisman A, van den Bemt PM, Schobben AF, Egberts AC, van Solinge WW. Linking laboratory and medication data: new opportunities for pharmacoepidemiological research. Clin Chem Lab Med 2007;45:13–9.
[20] McCabe WR, Jackson GG. Gram-negative bacteremia: etiology and ecology. Arch Intern Med 1962;110:847–55.
[21] Huang Y, Pepe MS. Assessing risk prediction models in case–control studies using semiparametric and nonparametric methods. Stat Med 2010;29:1391–410.
[22] Core Team R. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012.
[23] Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. BMJ 2009;338:b604.
[24] Bruns AH, Oosterheert JJ, Kuijper EJ, et al. Impact of different empirical antibiotic treatment regimens for community-acquired pneumonia on the emergence of Clostridium difficile. J Antimicrob Chemother 2010;65:2464–71.
[25] Goorhuis A, van Dissel JT, Kuijper EJ. Novel risk factors for *Clostridium difficile*–associated disease in a setting of endemicity? Clin Infect Dis 2008;47:429–30.
[26] Hung YP, Lee JC, Lin HJ, et al. Clinical impact of *Clostridium difficile* colonization. J Microbiol Immunol Infect 2014. http://dx.doi.org/10.1016/j.jmii.2014.04.011.
[27] Chandra S, Thapa R, Marur S, Jani N. Validation of a clinical prediction scale for hospital-onset *Clostridium difficile* infection. J Clin Gastroenterol 2014;48:419–22.
[28] Humphries RM, Uslan DZ, Rubin Z. Performance of *Clostridium difficile* toxin enzyme immunoassay and nucleic acid amplification tests stratified by patient disease severity. J Clin Microbiol 2013;51:869–73.
[29] Crobach MJ, Dekkers OM, Wilcox MH, Kuijper EJ. European Society of Clinical Microbiology and Infectious Diseases (ESCMID): data review and recommendations for diagnosing *Clostridium difficile*–infection (CDI). Clin Microbiol Infect 2009;15:1053–66.
[30] Vardakas KZ, Polyzos KA, Patouni K, Rafailidis PI, Samonis G, Falagas ME. Treatment failure and recurrence of *Clostridium difficile* infection following treatment with vancomycin or metronidazole: a systematic review of the evidence. Int J Antimicrob Agents 2012;40:1–8.