C H A P T E R 5

EXTERNAL VALIDATION OF THREE PREDICTION TOOLS FOR PATIENTS AT RISK OF A COMPLICATED COURSE OF CLOSTRIDIUM DIFFICILE INFECTION: DISAPPOINTING IN AN OUTBREAK SETTING

Yvette H. van Beurden, Marjolein P.M. Hensgens, Olaf M. Dekkers, Saskia Le Cessie, Chris J.J. Mulder, Christina M.J.E. Vandenbroucke-Grauls

Infection Control & Hospital Epidemiology 2017 Aug;38(8):897-905.

ABSTRACT

Objective

Estimating the risk of a complicated course of *Clostridium difficile* infection (CDI) might help doctors guide treatment. We aimed to validate three published prediction models: Hensgens (2014), Na (2015), and Welfare (2011).

Methods

The validation cohort comprised 148 patients diagnosed with CDI between May 2013 and March 2014. During this period, 70 endemic cases of CDI occurred as well as 78 cases of CDI related to an outbreak of *C. difficile* ribotype 027. Model calibration and discrimination were assessed for the three prediction rules.

Results

A complicated course (ie, death, colectomy, or ICU admission due to CDI) was observed in 31 patients (21%), and 23 patients (16%) died within 30 days of CDI diagnosis. The performance of all three prediction models was poor when applied to the total validation cohort with an estimated area under the curve (AUC) of 0.68 for the Hensgens model, 0.54 for the Na model, and 0.61 for the Welfare model. For those patients diagnosed with CDI due to non-outbreak strains, the prediction model developed by Hensgens performed the best, with an AUC of 0.78.

Conclusion

All three prediction models performed poorly when validated using our total cohort, which included CDI cases from an outbreak as well as endemic cases. The prediction model of Hensgens performed relatively well for patients diagnosed with CDI due to non-outbreak strains, and this model may be useful in endemic settings.

INTRODUCTION

Clostridium difficile infection (CDI) is a frequent cause of healthcare-associated diarrhea. It is associated with an infectionrelated mortality of 5%, and has an all-cause mortality of 15% - 20%.[1,2] The epidemiology of CDI has changed since the emergence of the B1/NAP1/027 strain in the early 2000s. This ribotype has been responsible for an increase in incidence and severity of CDI.[3] Current guidelines for the treatment of CDI recommend metronidazole for mild-to-moderate infection and vancomycin for severe infection.[4] A first recurrence is usually treated with vancomycin, and subsequent recurrences are treated with a tapered regimen of vancomycin, fidaxomicin or fecal microbiota transplantation (FMT).[4,5] Fidaxomicin and FMT both appear to lead to fewer recurrences.[6,7]

Early identification of patients at risk of a complicated course or death could help clinicians inform patients and might help doctors guide antibiotic treatment. Several scoring systems to predict a complicated course of CDI have been developed (Table 1).[8-19] However, none has gained widespread clinical acceptance due to lack of external validation, the retrospective study design, or the limited numbers of patients on which they are based.[20] To our knowledge, to date, only one study has been externally validated.[21] In four studies, internal validations were performed.[8,9,17,19] Between May 2013 and July 2014, the frequency of CDI cases increased above our usual endemic levels due to an outbreak of *C. difficile* ribotype 027.[22] The aims of the current study were to test published prediction models for complicated CDI for clinical use and to provide external validation for existing prediction models.

METHODS

Source of data

We searched PubMed and Embase for studies on prediction tools for a severe or complicated course of CDI up to February 2016 (Appendix A). We selected studies that (1) predicted at least one relevant outcome (ie, severity, complications, mortality) and (2) developed a prediction model or risk score. Only completed studies were included. Prediction tools that used a selected patient group (eg, only ICU patients), included nonquantitative parameters (eg, Horn's index, mental status), or parameters that were not available at the day of diagnosis in our cohort (eg, radiological findings or albumin concentration) were excluded.

Patient validation cohort and data collection

All consecutive hospitalized patients (≥18 years) diagnosed with CDI between May 2013 and March 2014 at the VU University medical center (VUmc), a 750-bed tertiary care center in Amsterdam, The Netherlands, were eligible. Patients diagnosed at the outpatient clinic, and patients who were lost to follow up were excluded. CDI was defined as the presence of diarrhea (ie, ≥3 stools per 24 hours) in combination with a positive *C. difficile* toxin test (VIDAS *C. difficile* A&B test, bioMèrieux, Marcy-l'Étoile, France). All isolates of *C. difficile* are routinely typed in our microbiology laboratory by amplified fragmentlength polymorphism, which permits differentiation between the different ribotypes. Medical charts were reviewed for demographic and clinical characteristics. For every participant, we recorded data for the predicting variables used in the different prediction models (Table 1): age, department of CDI diagnosis, presence of hypotension (systolic blood pressure below 100 mmHg and/or diastolic blood pressure below 60 mmHg), diarrhea as the reason for hospital admission, recent abdominal surgery, peak white blood count, peak creatinine level (ie, between 5 days before and 2 days after the diagnostic stool sample was obtained, as suggested by the original prediction model), and presence of renal disease or cancer (following

the definitions used in the Charlson comorbidity index[23]). In addition, data related to gender, Charlson comorbidity index, and *C. difficile* ribotype (outbreak strain ribotype 027 vs other ribotypes) were recorded. The study was approved by the VUmc Medical Ethics Committee.

Outcome

A course of CDI was considered complicated if any of the following criteria were met within 30 days after the diagnosis of CDI: (1) death as a direct or indirect consequence of CDI, (2) admission to the ICU for treatment of CDI or its complications, (3) surgery (colectomy) for toxic megacolon, perforation or refractory colitis.[24,25] This definition of a complicated course of CDI was identical to the definition used in the prediction rules by Hensgens *et al.*[8] and by Na *et al*.[9] In addition, we collected data on all cause 30-day mortality for validation of the prediction score by Welfare *et al*.[17] Whether the course of CDI was considered complicated was assessed by the study physician after chart review.

Missing Data

Most predictors had <1% missing data. However, 7% - 10% of values were missing for blood pressure on the day of diagnosis, white blood count, and serum creatinine level. When both diastolic and systolic blood pressure measurements were missing, we presumed that the patients had no hypotension on the day of diagnosis. To account for missing data on white blood count (N = 14) and serum creatinine level (N = 10), values were imputed using multiple imputation (10 imputed datasets).[26] The predictors in the selected studies, the outcome variable, and 5 additional variables (sex, Charlson comorbidity index[23], CDI due to ribotype 027, temperature ≥38.5°C, ward of CDI diagnosis) were included in the imputation procedure.

Table 1. Prediction tools for a complicated course of *Clostridium difficile* infection

Studies and predictors	Rubin 1995[13]	Velazquez Gomez 2008[12]	Vali-quete 2009[11]	Drew 2009[15]	Zilber-berg 2009[18]	Bhangu 2010[16]	Arora 2011[10]	Lungulescu 2011[14]	Welfare 2011[17]*	Butt 2013[19]	Hensgens 2014[8]*	Na 2015[9]*	Predictor used in different studies, %
White blood cell count/CRP	X	X	X	X		X		X		X		X	67
Age	X				X	X		X	X		X	X	58
Serum albumin	X	X	X	X		X		X		X			58
Abdominal pain	X	X											17
Admission or transfer ICU		X									X		17
Altered mental status	X	X											17
Hypotension		X									X		17
Immunosuppressive medication	X		X										17
Renal insufficiency/disease	X								X				17
Serum creatinine level								X				X	17
Serum urea				X		X							17
Antibiotic use	X												8
Anti-peristaltic/narcotic use	X												8
Ascites or colitis		X											8
Clinically severe disease[#]						X							8
COPD	X												8
CT findings			X										8
Diarrhea reason for admission											X		8
Fever		X											8
Hematocrit	X												8
History of malignancy									X				8
Horn's index							X						8
Recent abdominal surgery											X		8
Tachycardia		X											8
Absence of chronic respiratory disease					X								8
APACHE II score					X								8
Respiratory rate										X			8

*Validated prediction models in this manuscript; #Clinically severe disease: sepsis, peritonitis, and/or ≥10x diarrhea in 24h. Abbreviations: CRP: C-reactive protein; ICU: intensive care unit; COPD: chronic obstructive pulmonary disease; CT: computed tomography; APACHE: Acute Physiology and Chronic Health Evaluation

Statistical Analysis

Using the same multivariable models as used in the original studies, demographic characteristics and risk factors used in the prediction models were compared between patients with and without the outcome to give insight in the association between these predictors and the outcome in our validation cohort. These procedures were conducted using standard logistic regression. Risk scores were calculated for each patient by adding the allocated points for each variable according to the respective prediction model. To quantify how close predictions are to the actual outcome (calibration), we plotted the observed number of complicated cases against the predicted number of complicated CDI courses in the simplified risk categories provided by the original studies. The ability of the prediction models to discriminate between those with and without a complicated CDI course was assessed by calculating the area under the receiver operating characteristic (ROC) curve (AUC), which ranges from no discrimination (0.5) to perfect discrimination (1.0).[27] Because all prediction models were developed in an endemic setting, in a second step, we reexamined calibration and discrimination within the group of patients who had been diagnosed with CDI due to endemic strains (ie, strains that had never caused outbreaks in our hospital). SPSS version 22 software (IBM, Armonk, NY) was used for statistical analysis.

RESULTS

Selection of Prediction Models

We identified 12 papers presenting a prediction model for a complicated course of CDI (Table 1)[8-18], of which eight studies were excluded because they included a selected patient group[18], used nonquantitative parameters (eg, altered mental status, Horn's index)[10-12], or used variables that were not available in our cohort on the day of

CDI diagnosis (eg, serum albumin, radiological findings).[12-16,19] Age, white blood count, and albumin level were used in the prediction model in >50% of the studies. All other 24 predictors were used in only 10% or 20% of models. In this study, we sought to validate three prediction models from three different studies. Study and patient characteristics of the 3 derivation studies, and of our validation cohort are shown in Table 2.

Prediction model by Hensgens *et al*.[8]

The prediction model developed by Hensgens *et al.* calculates the probability to develop a complicated course. Based on total points, Hensgens *et al.* defined four risk categories: no risk (<0 points), low risk (0 – 1 point), medium risk (2 – 3 points), and high risk (≥4 points).

Prediction model by Na *et al.*[9]

The prediction model developed by Na *et al.* calculates the probability to develop a severe clinical course of CDI. Based on total points, Na *et al.* defined two risk categories: low risk (0 – 1 points), and high risk (2 – 3 points).

Prediction model by Welfare *et al*.[17]

The prediction model developed by Welfare *et al.* estimates all-cause mortality risk at 30 days. Based on total points, Welfare *et al.* defined three risk categories: low risk (0 – 3 points), medium risk (4 – 7 points), and high risk (8 points).

Table 2. Study and patient characteristics of the derivation and validation sets

	Derivation set Hensgens *et al.*[8]	Derivation set Na *et al.*[9] No. (%)	Derivation set Welfare *et al.*[17] No. (%)	Validation set No. (% / ±SD)
Study setting				
Inclusion period	2006 – 2009	2004 –2006	2002-2009	2013 – 2014
No. of patients	395	263	2571	148
Setting	Endemic	Endemic	Endemic	Outbreak + endemic
No. of patients diagnosed in University hospital	266 (67%)	263 (100%)	0 (0%)	148 (100%)
Location	9 Dutch hospitals[a]	BIDMC	Northumbria NHS[b]	VUmc
Patient characteristics				
Age	Median 65 (IQR: 52-77)	Mean 67 (± 17)	Mean 80 (± 11)	Mean 65 (±18)
Male sex	220 (56%)	131 (50%)	933 (36%)	94 (64%)
CDI due to ribotype 027	8%	NR	NR	53%
Charlson score, mean (±SD)	NA	3.3 (± NA)	NR	2.57 (±2.3)
0	59 (15%)	NR	NR	25 (17%)
1 – 2	150 (38%)	NR	NR	65 (44%)
3 – 4	120 (31%)	NR	NR	38 (26%)
≥ 5	64 (16%)	NR	NR	20 (14%)
CDI treatment				
Metronidazole	74%	NR	NR	22 (15%)
Vancomycin	3%	NR	NR	1 (1%)
Metro + vanco	11%	NR	NR	100 (68%)
Fidaxomicin	0%	NR	NR	1 (1%)
No treatment	12%	NR	NR	24 (16%)
Complicated course	46 (12%)	51 (19%)	NR	31 (21%)
30-day mortality	65 (17%)	NR	834 (33%)	23 (16%)
Predictors Hensgens *et al.*				
Age				
≤ 49 years	85 (22%)	NR	NR	22 (15%)
50 – 84 years	275 (70%)	NR	NR	113 (76%)
≥ 85 years	35 (9%)	NR	NR	13 (9%)
Dept. of diagnosis				
Other	293 (74%)	NR	NR	92 (62%)
Surgery	83 (21%)	NR	NR	45 (30%)
Intensive Care Unit	19 (5%)	NR	NR	11 (7%)
Recent abdominal surgery	110 (28%)	NR	NR	25 (17%)
Hypotension	117 (30%)	NR	NR	20 (14%)
Diarrhea reason admission	104 (27%)	NR	NR	13 (19%)
Predictors Na *et al.*				
Age, mean	NR	66.5 (±17.4)	80 (± 11)	64.7 (±17.9)
Peak[c] white blood cells (x10^9/L), mean (±SD)	NR	15.5 (±11.4)	NR	17.8
Peak[c] creatinine (umol/L), mean (±SD)	NR	159.1 (±159.1)	NR	135.2
Predictors Welfare *et al.*				
Age				
< 60	NR	NR	176 (6%)	48 (32%)
60-79	NR	NR	955 (35%)	68 (46%)
≥80	NR	NR	1630 (59%)	32 (22%)
Renal disease	NR	NR	629 (23%)	10 (7%)
Cancer	NR	NR	416 (15%)	55 (37%)

[a]5 university + 4 community hospitals; [b]Three general hospitals and seven community hospitals; [c]During 7 day time period from 5 days before to 2 days after the diagnostic stool sample was obtained. Abbreviations: SD: standard deviation; BIDMC: Beth Israel Deaconess Medical Center, Boston, USA; NHS: National Health Services, United Kingdom; VUmc: VU University medical center, Amsterdam, The Netherlands; CDI: *Clostridium difficile* infection; IQR: interquartile range; NR: not reported.

Patient Characteristics of the Validation Cohort

Patient characteristics are shown in Table 2. In total, 148 CDI patients were included in the validation set, of which 78 patients diagnosed with CDI due to the outbreak strain ribotype 027, and 70 patients with CDI due to other ribotypes (Table 3). In general, no differences were observed between the derivation study cohorts of Hengens *et al.* and of Na *et al.* and the validation cohort in age, gender, or Charlson comorbidity index. The derivation cohort of Welfare *et al.* differed from the validation cohort in mean age and gender.

Table 3. Patient characteristics of the validation cohort: patients diagnosed with CDI due to the outbreak strain ribotype 027 versus other strains

Patient characteristics	Validation set Outbreak strain (N = 78)	Validation set Other strains (N = 70)	Crude odds ratios (95% CI)
Age, median (range)	70 (18-93)	65 (18-94)	1.0 (1.0-1.0)
Male sex	50 (64%)	44 (63%)	0.9 (0.5-1.9)
CDI due to ribotype 027	100%	0%	NA
Charlson score, mean (±SD)	2.5 (±2.0)	2.7 (± 2.6)	1.0 (0.8-1.1)
CDI treatment			
Metronidazole	12 (16%)	10 (14%)	1.1 (0.4-2.8)
Vancomycin	0 (%)	1 (1%)	NA
Metronidazole + vancomycin	56 (73%)	44 (63%)	1.6 (0.8-3.2)
Fidaxomicin	1 (1%)	0 (0%)	NA
No treatment	9 (12%)	15 (21%)	2.1 (0.9-5.1)
Complicated course	22 (28%)	9 (13%)	2.7 (1.1-6.3)
30-day mortality	13 (17%)	10 (14%)	1.2 (0.5-2.9)

Abbreviations: CI: confidence interval; CDI: *Clostridium difficile* infection ; NA: not applicable

In the total validation cohort, 84% of patients received antibiotic treatment for CDI. Most frequently, a combination of metronidazole and vancomycin was used (68%). This differs from the derivation cohort of Hensgens *et al*, where metronidazole monotherapy was used most frequently (74%). No information regarding antibiotic treatment was available for the cohorts of Na *et al*. or of Welfare *et al*.

A complicated course due to CDI was observed in 31 patients (21%), and 23 patients (16%) died within 30 days after CDI diagnosis; death was CDI related in 13 of these

cases. Because of severe pseudomembranous colitis, one patient with CDI underwent a colectomy. Overall, 20 patients were admitted to the ICU after CDI diagnosis; 13 of these cases were related to CDI.

Model Performance

The association between predictors and a complicated course of CDI[8,9] or 30-day mortality[17] in the validation cohort was analyzed by multivariable logistic regression using the same multivariable models as the original studies (Table 4).

Prediction model by Hensgens *et al*.[8]

In the validation cohort, the median score using the prediction tool developed by Hensgens *et al*. was 1 (range, −3 to 6). Age and department of diagnosis (ICU) were significantly associated with a complicated course of CDI in our cohort. We compared the observed and predicted outcomes for the different risk groups defined by the prediction rule (Figure 1 and Appendix B). In the validation cohort, a higher score corresponded with a higher chance of a complicated course. This finding is similar to the results of the pilot external validation performed in the original manuscript. With an estimated AUC of 0.68 (95% CI 0.57-0.79), discrimination between patients with and without a complicated course of CDI was poor (Figure 2.1).[27] Analysis restricted to patients with CDI due to non-outbreak strains (N = 70) showed fair discrimination, with an AUC of 0.78 (95% CI 0.61-0.95) (Figure 2.2).[27]

Prediction model by Na *et al*.[9]

The median score, using the prediction tool developed by Na *et al*. in the validation cohort, was 1 (range, 1 – 3). No statistically significant association was identified in a multivariable model between the predictors by Na *et al*. and a complicated course of CDI in our cohort. We compared the observed and predicted outcomes for the different risk groups defined by the prediction rule (Figure 1 and Appendix C). Discrimination of the prediction model failed (AUC 0.53; 95% CI 0.42-0.65; Figure

2.1).[27] Analysis restricted to patients with CDI due to non-outbreak strains, also showed very poor discrimination (AUC 0.54; 95% CI 0.39-0.70; Figure 2.2).

Table 4. Multivariate analysis of predictors identified by the different prediction tools compared to multivariate OR of the original prediction studies

Patient characteristics	Scores original study	Complicated course / all cause 30-day mortality[a]		Adjusted OR original studies	Adjusted OR validation set
		Yes	No	(95% CI)	(95% CI)[b]
Predictors Hensgens *et al.*[c]					
Age					
≤ 49 years	0	1 (3%)	21 (18%)	1 (reference)	Reference
50-84 years	1	25 (81%)	88 (75%)	1.8 (0.68-4.97)	6.3 (0.7-55.1)
≥ 85 years	3	5 (16%)	8 (7%)	5.0 (1.4-17.6)*	14.9 (1.3-175.0)*
Department of diagnosis					
Other departments	0	13 (42%)	73 (62%)	1 (reference)	Reference
Surgery	0	10 (32%)	41 (35%)	1.0 (0.3-3.2)	1.2 (0.4-3.9)
Intensive Care Unit	3	8 (26%)	3 (3%)	7.0 (2.0-24.4)*	19.3 (3.1-119.7)*
Recent abdominal surgery	-3	3 (10%)	22 (19%)	0.2 (0.1-0.7)*	0.4 (0.1-1.9)
Hypotension	2	7 (23%)	13 (11%)	3.3 (1.5-6.9)*	0.6 (0.1-2.9)
Diarrhea reason for admission	2	6 (19%)	16 (14%)	3.3 (1.6-6.8)*	1.4 (0.4-5.0)
Predictors Na *et al.*[d]					
Age ≥ 65 years	1	20 (65%)	61 (52%)	2.4 (1.1-5.4)*	1.7 (0.8-4.0)
Peak WBC ≥ 20x10^9/L	1	11 (36%)	43 (37%)	4.2 (2.1-8.6)*	1.0 (0.4-2.4)
Peak creatinine ≥ 177 umol/L	1	4 (13%)	20 (17%)	8.1 (2.5-26.3)*	0.6 (0.2-2.2)
Predictors Welfare *et al.*[e]					
Age					
< 60	0	4 (8%)	44 (92%)	Reference	Reference
60-79	3	8 (12%)	60 (88%)	2.6 (1.6-4.3)*	1.5 (0.4-5.1)
≥ 80	4	11 (34%)	21 (66%)	4.2 (2.6-6.9)*	5.9 (1.7-20.8)*
Renal disease	2	2 (20%)	8 (80%)	2.0 (1.6-2.4)*	1.2 (0.5-2.9)
Cancer	2	8 (15%)	47 (86%)	2.0 (1.7-2.6)*	1.1 (0.6-1.8)

[a]Complicated course for the models of Hensgens *et al.* and Na *et al,* 30-day mortality for model of Welfare *et al.*; [b]Multivariable models using the same variables as used in the original studies; [c]Age, department, hypotension measured on the day of diagnosis, recent abdominal surgery gathered from 3 months prior to CDI diagnosis; [d]Age on the day of CDI diagnosis, peak WBC and peak creatinine measured between 5 days before to 2 days after the diagnostic CDI stool sample was obtained; [e]Data measurements were taken on the day of CDI diagnosis; *Statistically significant. Abbreviations: CDI: *Clostridium difficile* infection; OR: odds ratios; CI: confidence interval; WBC: white blood cells

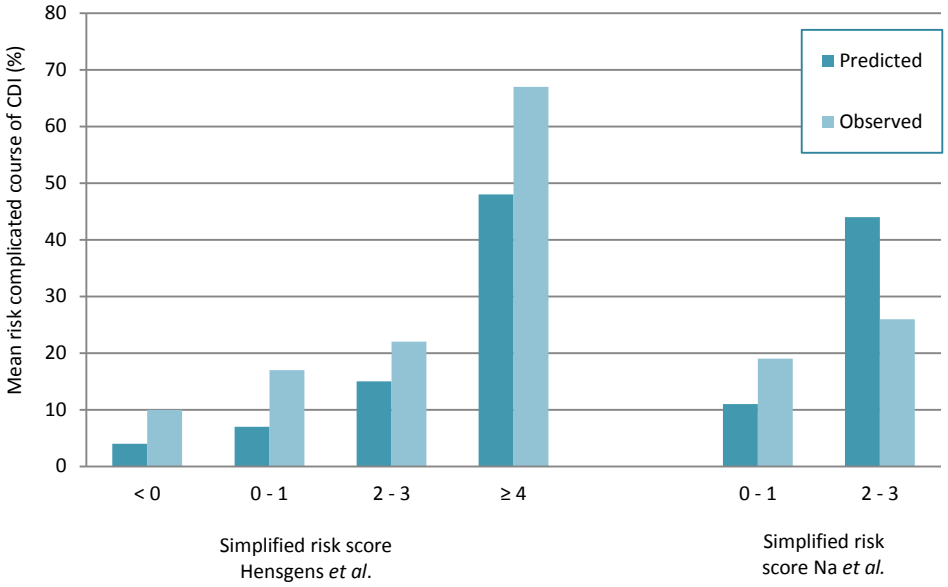


Figure 1. Validation of the prediction tools developed by Hensgens *et al.*[8] and by Na *et al.*[9] Observed versus predicted complicated course per simplified risk group following their prediction scores.

Prediction model by Welfare *et al.*[17]

In the validation cohort, the median score using the prediction model of Welfare *et al.* was 3 (range, 0–6). Only age ≥80 years was significantly associated with 30-day mortality. We were not able to compare the observed outcome in our cohort with the predicted outcome defined by the prediction rule due to lack of data in the original study by Welfare *et al.* (Appendix D). Therefore, the prediction model developed by Welfare *et al.* was not included in Figure 1. Discriminatory power of the prediction tool was poor (AUC 0.61; 95% CI 0.50-0.73; Figure 2.1).[27] Analysis restricted to patients with CDI due to non-outbreak strains showed similarly poor discrimination (AUC 0.56; 95% CI 0.37-0.76; Figure 2.2).

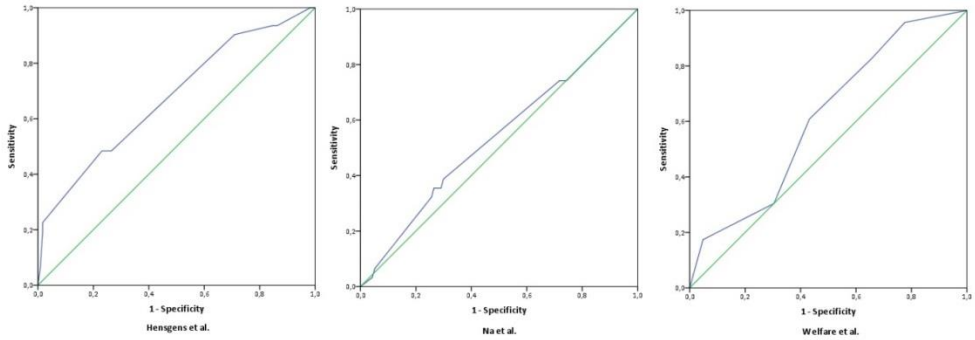


Figure 2.1. Pooled ROC curve performance prediction models, using the original scores.

Hensgens *et al.* prediction model: AUC 0.68 (95% CI 0.57-0.79)

Na *et al.* prediction model: AUC 0.54 (95% CI 0.42-0.65)

Welfare *et al.* prediction model: AUC 0.61 (95% CI 0.50-0.73)

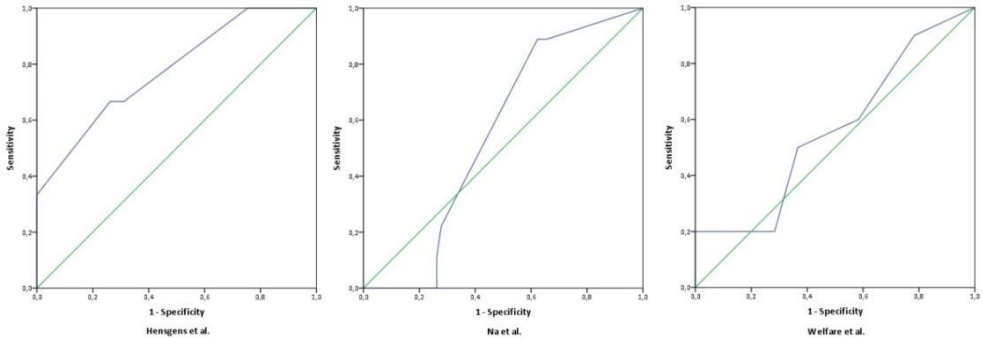


Figure 2.2. Pooled ROC curve prediction models (only non-027), using the original scores.

Hensgens *et al.* prediction model: AUC 0.78 (95% CI 0.61-0.95)

Na *et al.* prediction model: AUC 0.55 (95% CI 0.40-0.70)

Welfare *et al.* prediction model: AUC 0.56 (95% CI 0.37-0.76)

DISCUSSION

In this study, we validated three prediction models, all designed to identify patients at risk for a complicated course of CDI or death within 30 days after diagnosis, using a validation cohort in which approximately half of patients were part of a CDI outbreak

caused by ribotype 027. When applied to our total cohort (outbreak strain and non-outbreak strains), all three prediction models performed poorly. However, because the prediction tools were derived in endemic settings and because infection with *C. difficile* ribotype 027 has been associated with more severe outcomes[28-31], we also tested the prediction models on our cohort excluding patients infected with the outbreak strain *C. difficile* ribotype 027. In this restricted analysis, the prediction model developed by Hensgens *et al*. performed much better, with an AUC of nearly 0.8. This refinement provided strength to our study, showing that a prediction rule can only be applied in a cohort that is comparable with the derivation cohort.

The recent guideline of the Infectious Diseases Society of America (IDSA) recommends that age, peak leucocyte count, and peak serum creatinine level could be considered as indicators of a complicated course of CDI when treatment is started.[4] Na *et al.* used these three variables in their prediction model. However, our study findings did not support their prediction tool, which seems counterintuitive. This difference may be due to the dichotomized scoring system or to the chosen cutoff points of the different variables. In addition, it has been suggested that measurements of leucocyte count and serum creatinine level at different times around the day of CDI diagnosis lead to different severity classifications in many cases.[32] Na *et al*. used a time interval of seven days for the measurement of peak leucocyte count and serum creatinine level, which could have contributed to the poor performance of their prediction rule. The prediction model developed by Hensgens *et al*. was constructed using adequate methodology: the predictors were selected using backward selection, and bootstrapping techniques were used to adjust for overfitting.[33] Additionally, this scoring system was based on clinical parameters that are available after the completion of medical history and physical examination, which enables the attending physician to use it for treatment guidance.[8] Welfare *et al.* also based their prediction tool on clinical parameters that are readily available at a patient's bedside. They

performed the largest study on comorbidities as predictors of 30-day mortality in patients with CDI, and they developed the age, renal disease, and cancer (ARC) score to predict 30-day mortality. Welfare *et al.* reported a mortality rate of 33%, which is double the mortality rate in our validation cohort. The mean age of the cohort of Welfare *et al.* was 15 years higher than that of our cohort, and our cohort did not contain patients in the highest risk category. Additionally, in our cohort, data on clinical diagnoses were obtained from medical charts, which differs from the derivation study in which administration codes were used. These differences are a limitation of our study, and they could have influenced the performance of their prediction tool in our cohort. Unfortunately, the number of patients in the older age category in our validation cohort was too small to allow a validation in this subset. A second limitation of our study was the difference between the validation and derivation cohorts: the derivation cohorts consisted of patients diagnosed with CDI in an endemic setting, whereas our validation cohort consisted of all patients diagnosed with CDI (due to the outbreak strain, and due to other strains) during an outbreak of *C. difficile* ribotype 027. The number of patients who developed a complicated course of CDI was significantly higher among patients diagnosed with CDI due to ribotype 027 than among non-027 CDI patients. Third, we used existing patient data files; therefore, some data were missing. We found 12 studies on prediction rules for a complicated course of CDI. Due to missing data for several variables per prediction tool, we were not able to validate all 12 different prediction tools. However, as shown in Table 1, except for number of leucocytes, serum albumin (which is not measured routinely in our hospital), and age, there was much heterogeneity in the variables used in the various scores. In addition, inclusion of nonquantitative parameters and parameters that are not routinely available at time of diagnosis limits the use of prediction tools in clinical practice. Importantly, a recent systematic review showed that most prediction tools presented several methodological limitations and weak validities, which lead to suboptimal quality and debatable utility in clinical

settings.[20,34] Fourth, the relatively small sample size affected the precision of the estimates. Missing data were imputed using multiple imputation; when data on blood pressure measurements were missing, we made the assumption that patients had no hypotension on the day of diagnosis. Although we feel that these corrections for missing data were accurate, this factor could have influenced the results. Finally, different antibiotic regimens may have different risks for the development of a complicated course of CDI.[35,36] Data on CDI treatment were missing for both the derivation cohort of Na *et al.* and that of Welfare *et al*; CDI treatment regimens used in the study by Hensgens *et al.* differed from the regimens used in the validation cohort. These differences made it difficult to fully assess similarities between the validation cohort and the derivation cohorts.

In conclusion, our study shows that a prediction rule can only be used in a cohort comparable with the derivation cohort. The performance of all three validated prediction models was disappointing in a combined epidemic/endemic setting, and can therefore not be used during a CDI outbreak to predict a complicated course of CDI. The prediction model by Hensgens *et al.* performed relatively well, however, to identify patients with CDI at risk for a complicated course when the validation cohort was restricted to patients with CDI due to non-outbreak strains. This prediction model can therefore be used in an endemic setting to identify patients at risk for CDI complications, aiding clinicians in deciding which patients to monitor closely for CDI-related complications. Future clinical trials should be directed at determining whether early treatment and choice of treatment directed by the severity score improves patient outcomes.

REFERENCES

1. Lessa FC, Mu Y, Bamberg WM, et al. Burden of Clostridium difficile infection in the United States. *N Engl J Med* 2015; 372:825-834.
2. Leffler DA, Lamont JT. Clostridium difficile infection. *N Engl J Med* 2015; 372:1539-1548.
3. O'Connor JR, Johnson S, Gerding DN. Clostridium difficile infection caused by the epidemic BI/NAP1/027 strain. *Gastroenterology* 2009; 136:1913-1924.
4. Cohen SH, Gerding DN, Johnson S, et al. Clinical practice guidelines for Clostridium difficile infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). *Infect Control Hosp Epidemiol* 2010; 31:431-455.
5. Keller JJ, Kuijper EJ. Treatment of recurrent and severe Clostridium difficile infection. *Annu Rev Med* 2015; 66:373-86.
6. van Nood E, Vrieze A, Nieuwdorp M, et al. Duodenal infusion of donor feces for recurrent Clostridium difficile. *N Engl J Med* 2013; 368:407-415.
7. Louie TJ, Miller MA, Mullane KM, et al. Fidaxomicin versus vancomycin for Clostridium difficile infection. *N Engl J Med* 2011; 364:422-431.
8. Hensgens MPM, Dekkers OM, Goorhuis A, LeCessie S, Kuijper EJ. Predicting a complicated course of Clostridium difficile infection at the bedside. *Clin Microbiol Infect* 2014; 20:O301-O308.
9. Na X, Martin AJ, Sethi S, et al. A Multi-Center Prospective Derivation and Validation of a Clinical Prediction Tool for Severe Clostridium difficile Infection. *PLoS One* 2015; 10:e0123405.
10. Arora V, Kachroo S, Ghantoji SS, Dupont HL, Garey KW. High Horn's index score predicts poor outcomes in patients with Clostridium difficile infection. *J Hosp Infect* 2011; 79:23-6.
11. Valiquette L, Pepin J, Do XV, et al. Prediction of complicated Clostridium difficile infection by pleural effusion and increased wall thickness on computed tomography. *Clin Infect Dis* 2009; 49:554-60.
12. Velazquez-Gomez I, Rocha-Rodriguez R, Toro DH, et al. A Severity Score Index for *Clostridium difficile* Infection. *Infectious Diseases in Clinical Practice* 2008; 16:376-378.
13. Rubin MS, Bodenstein LE, Kent KC. Severe Clostridium difficile colitis. *Dis Colon Rectum* 1995; 38:350-4.
14. Lungulescu OA, Cao W, Gatskevich E, et al. CSI: a severity index for Clostridium difficile infection at the time of admission. *J Hosp Infect* 2011; 79:151-4.
15. Drew RJ, Boyle B. RUWA scoring system: a novel predictive tool for the identification of patients at high risk for complications from Clostridium difficile infection. *J Hosp Infect* 2009; 71:93-4; author reply 94-5.

16. Bhangu S, Bhangu A, Nightingale P, Michael A. Mortality and risk stratification in patients with Clostridium difficile-associated diarrhoea. *Colorectal Dis* 2010; 12:241-6.
17. Welfare MR, Lalayiannis LC, Martin KE, et al. Co-morbidities as predictors of mortality in Clostridium difficile infection and derivation of the ARC predictive score. *J Hosp Infect* 2011; 79:359-63.
18. Zilberberg MD, Shorr AF, Micek ST, Doherty JA, Kollef MH. Clostridium difficile-associated disease and mortality among the elderly critically ill. *Crit Care Med* 2009; 37:2583-9.
19. Butt E, Foster JA, Keedwell E, et al. Derivation and validation of a simple, accurate and robust prediction rule for risk of mortality in patients with Clostridium difficile infection. *BMC Infect Dis* 2013; 13:316.
20. Abou Chakra CN, Pepin J, Valiquette L. Prediction tools for unfavourable outcomes in Clostridium difficile infection: a systematic review. *PLoS One* 2012; 7:e30258.
21. Toro DH, Amaral-Mojica KM, Rocha-Rodriguez R, Gutierrez-Nunez J. An innovative severity score index for *Clostridium difficile* infection: A prospective study. *Infectious Diseases in Clinical Practice* 2011; 16:336-339.
22. van Beurden YH, Dekkers OM, Bomers MK, et al. An Outbreak of Clostridium difficile Ribotype 027 Associated with Length of Stay in the Intensive Care Unit and Use of Selective Decontamination of the Digestive Tract: A Case Control Study. *PLoS One* 2016; 11:e0160778.
23. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; 40:373-383.
24. Kuijper EJ, Coignard B, Tull P. Emergence of Clostridium difficile-associated disease in North America and Europe. *Clin Microbiol Infect* 2006; 12 Suppl 6:2-18.
25. McDonald LC, Coignard B, Dubberke E, et al. Recommendations for surveillance of Clostridium difficile-associated disease. *Infect Control Hosp Epidemiol* 2007; 28:140-145.
26. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; 338:b2393.
27. Tazhibi M, Bashardoost N, Ahmadi M. Kernel smoothing for ROC curve and estimation for thyroid stimulating hormone. *Int J Public Health Res* 2011:239–242.
28. 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; 353:2442-2449.
29. See I, Mu Y, Cohen J, et al. NAP1 strain type predicts outcomes from Clostridium difficile infection. *Clin Infect Dis* 2014; 58:1394-400.

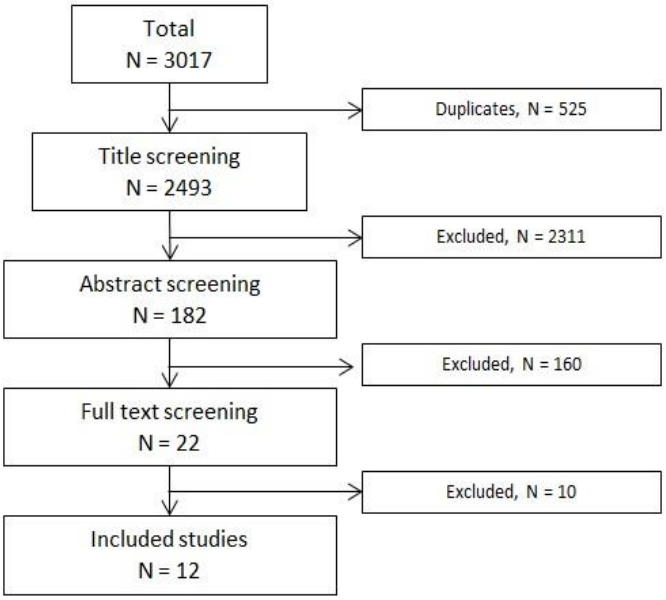
30. Cloud J, Noddin L, Pressman A, Hu M, Kelly C. Clostridium difficile strain NAP-1 is not associated with severe disease in a nonepidemic setting. *Clin Gastroenterol Hepatol* 2009; 7:868-873.
31. Pepin J, Valiquette L, Gagnon S, Routhier S, Brazeau I. Outcomes of Clostridium difficile-associated disease treated with metronidazole or vancomycin before and after the emergence of NAP1/027. *Am J Gastroenterol* 2007; 102:2781-8.
32. Bauer MP, Hensgens MP, Miller MA, et al. Renal failure and leukocytosis are predictors of a complicated course of Clostridium difficile infection if measured on day of diagnosis. *Clin Infect Dis* 2012; 55 Suppl 2:S149-53.
33. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. *Springer* 2009:1-497.
34. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999; 130:515-24.
35. Stevens VW, Nelson RE, Schwab-Daugherty EM, et al. Comparative Effectiveness of Vancomycin and Metronidazole for the Prevention of Recurrence and Death in Patients With Clostridium difficile Infection. *JAMA Intern Med* 2017.
36. Nelson RL, Kelsey P, Leeman H, et al. Antibiotic treatment for Clostridium difficile-associated diarrhea in adults. *Cochrane Database Syst Rev* 2011:CD004610.

APPENDICES

Appendix A

Databases used: Pubmed and Embase

Keywords: *Clostridium difficile* infection (CDI), *Clostridium difficile* associated diarrhea (CDAD), *Clostridium difficile* associated disease, prediction model, clinical prediction rule, sensitivity, specificity, risk score, risk scale, risk index, prognosis, pseudomembranous (entero)colitis, complicated course, mortality, severe. The search was limited to humans, English language, and availability of full text.



Appendix B

Predicted and observed complicated course by prediction rule Hensgens *et al.*[8] for patients with CDI

Prediction score	Patients (N)	Observed complicated courses of CDI (N)	Observed complicated course of CDI (%)	Average predicted complicated course of CDI (%)[a]
Complete score				
-3	2	0	0%	3%
-2	16	2	13%	3%
-1	2	0	0%	10%
0	17	1	6%	5%
1	65	13	20%	7%
2	4	1	25%	18%
3	33	8	24%	15%
4	1	0	0%	28%
5	5	4	80%	49%
6	3	2	67%	51%
≥ 7	0	NA	NA	NA
Simplified score[b]				
< 0 (no risk)	20	2	10%	4%
0 – 1 (low risk)	82	14	17%	7%
2 – 3 (medium risk)	37	8	22%	15%
≥ 4 (high risk)	9	6	67%	48%
Total	148	31	21%	12%

[a]For each patient the predicted probability on a complicated course was calculated using the formula of Hensgens *et al.*: P = 1 / (1 + exp-(-3.15 + 0.52 x age 50-84 + 1.38 x age ≥ 85 ± 0.02 x department of surgery + 1.68 x department of ICU ± 1.26 x recent abdominal surgery + 1.01 x hypotension + 1.01 x diarrhea reason for admission)). [b]Simplified score derived from original publication.

Appendix C

Predicted and observed complicated course by prediction model of Na *et al.*[9] for patients with CDI

Prediction score	Patients (N)	Observed complicated courses of CDI (N)	Observed complicated course of CDI (%)	Predicted complicated course of CDI (%)[a]
Complete score				
0	40	8	20%	7%
1	65	12	19%	12%
2	35	9	27%	41%
3	8	2	22%	58%
Simplified score[b]				
0 or 1	105	20	19%	11%
2 or 3	43	11	26%	44%
Total	148	31	21%	19%

[a]Based on the actual observed severe outcomes in the derivation cohort of Na *et al.* (see table 1). [b]Simplified score derived from original publication.

Appendix D

Predicted and observed 30-day mortality by prediction rule of Welfare *et al.*[17] for patients with CDI

Prediction score	Patients (N)	Observed 30-day mortality (N)	Observed 30 day mortality (%)	Average predicted 30 day mortality (%)[a]
Complete score				
0	29	1	3.4%	9%
2	18	3	16.7%	NA
3	33	5	15.2%	21%
4	23	7	30.4%	31%
5	35	3	8.6%	NA
6	10	4	40%	NA
7	0	NA	NA	48.0%
8	0	NA	NA	66%
Simplified score[b]				
0 – 3 (low risk)	80	9	11.3%	9% - 21%
4 – 7 (medium risk)	68	14	20.6%	31% - 48%
8 (high risk)	0	NA	NA	66%
Total	148	23	16%	33%

[a]Based on the actual observed severe outcomes in the derivation cohort of Welfare *et al.* (see table 1).

[b]Simplified score derived from original publication. Abbreviations: NA: not applicable.