Infection Control & Hospital Epidemiology (2022), **43**, 298–305
doi:10.1017/ice.2021.114

SHEA

Original Article

Predicting healthcare-associated infections, length of stay, and mortality with the nursing intensity of care index

Bevin Cohen PhD, MS, MPH, RN[1], Elioth Sanabria MS[2], Jianfang Liu PhD, MAS[3], Philip Zachariah MD, MS[4], Jingjing Shang PhD, RN[3], Jiyoun Song PhD, MS, AGACNP-BC, RN[3], David Calfee MD, MS[5], David Yao PhD[2] and Elaine Larson PhD, RN, CIC, FAAN[3]

[1]Center for Nursing Research and Innovation, The Mount Sinai Hospital, New York, New York, [2]Columbia University Fu Foundation School of Engineering and Applied Sciences, New York, New York, [3]Columbia University School of Nursing, New York, New York, [4]Columbia University Vagelos College of Physicians and Surgeons, New York, New York and [5]Weill Cornell Medical College, New York, New York

Abstract

Objectives: The objectives of this study were (1) to develop and validate a simulation model to estimate daily probabilities of healthcare-associated infections (HAIs), length of stay (LOS), and mortality using time varying patient- and unit-level factors including staffing adequacy and (2) to examine whether HAI incidence varies with staffing adequacy.

Setting: The study was conducted at 2 tertiary- and quaternary-care hospitals, a pediatric acute care hospital, and a community hospital within a single New York City healthcare network.

Patients: All patients discharged from 2012 through 2016 (N = 562,435).

Methods: We developed a non-Markovian simulation to estimate daily conditional probabilities of bloodstream, urinary tract, surgical site, and *Clostridioides difficile* infection, pneumonia, length of stay, and mortality. Staffing adequacy was modeled based on total nurse staffing (care supply) and the Nursing Intensity of Care Index (care demand). We compared model performance with logistic regression, and we generated case studies to illustrate daily changes in infection risk. We also described infection incidence by unit-level staffing and patient care demand on the day of infection.

Results: Most model estimates fell within 95% confidence intervals of actual outcomes. The predictive power of the simulation model exceeded that of logistic regression (area under the curve [AUC], 0.852 and 0.816, respectively). HAI incidence was greatest when staffing was lowest and nursing care intensity was highest.

Conclusions: This model has potential clinical utility for identifying modifiable conditions in real time, such as low staffing coupled with high care demand.

(Received 5 November 2020; accepted 4 March 2021; electronically published 16 April 2021)

Aimed at facilitating real-time surveillance and early intervention, prediction models for healthcare-associated infections (HAIs) have proliferated as healthcare data have become increasingly accessible and robust.[1–4] Though model performance has improved with new data sources and engineering methods, actionable models that provide individualized infection probabilities that are updated continuously based on patients' actual hospital event sequences and care trajectories are lacking.[5]

In addition, nurse staffing adequacy, a key predictor of infection risk,[6,7] has been largely omitted from prediction models due to challenges with defining and capturing this factor in electronic records.[8–12] Staffing adequacy represents the balance between staffing supply (the number and composition of staff on a unit) and patient demand (the type and quantity of care required by patients on a unit).[13,14] Although staffing supply can be measured and extracted from electronic records, existing methods of determining patient demand using healthcare data have significant limitations.[15,16] To more accurately capture patient demand, we previously developed and validated the Nursing Intensity of Care Index to quantify patients' daily nursing care needs based on their actual use of services.[17]

This study had 2 main objectives. The first objective was to develop and validate a simulation model to estimate patients' daily probabilities of seven outcomes (bloodstream infection, urinary tract infection, surgical site infection, *Clostridioides difficile* infection, pneumonia, length of stay, mortality) based on time-varying patient- and unit-level factors, including nurse staffing adequacy. The second objective was to examine whether the incidence of HAIs considered to be preventable through evidence-based nursing care (central line-associated bloodstream infection, catheter-associated urinary tract infection, *C. difficile* infection, pneumonia)[18–22] vary according to nurse staffing adequacy.

Author for correspondence: Bevin Cohen, E-mail: bevin.cohen@mountsinai.org
Cite this article: Cohen B, *et al.* (2022). Predicting healthcare-associated infections, length of stay, and mortality with the nursing intensity of care index. *Infection Control & Hospital Epidemiology*, 43: 298–305, https://doi.org/10.1017/ice.2021.114

Methods

Sample and setting

This federally funded study was performed using data from 2 tertiary- and quaternary-care hospitals, a pediatric acute care hospital, and a community hospital within a single New York City healthcare network. All inpatients discharged from 2012 through 2016 were included (N=562,435). The participating medical centers' institutional review boards reviewed and approved the study.

The Nursing Intensity of Care Index

The Nursing Intensity of Care Index was developed at the study institution using previously described methods.[17] In brief, a team of clinician researchers reviewed lists of all procedures occurring at the study institutions and identified those thought to increase workload for nurses on inpatient units. Eleven staff nurses in a range of pediatric and adult units reviewed the curated list, indicated whether the procedures increased their workload by at least 15 minutes per shift, and suggested additional procedures that had not been included. Each procedure in the final list was ascribed a weight based on previously published data describing the time burden of each activity.[23] The elements of the final index, including weights and data sources, are presented in Table 1.

Data sources and definitions

Our team of clinician researchers identified fixed and time-varying patient- and unit-level factors that contribute to infection risk and are obtainable from electronic hospital records. Data on these factors were obtained from administrative records, human resources staffing records, medication administration records, perioperative records, provider order entries, structured nursing documentation in the electronic medical record, *International Classification of Diseases Ninth or Tenth Revision, Clinical Modification* (ICD-9/10-CM) procedure and diagnosis codes, and records of unit-level infectious disease outbreak periods reported to New York State through the Nosocomial Outbreak Reporting Application.[24] The variables included in the simulation model are detailed in Table 2.

HAIs were identified using previously validated case-detection algorithms based on Centers for Disease Control and Prevention National Healthcare Safety Network definitions, and time-stamped records from the institutions' clinical microbiology laboratories and patients' ICD-9/10-CM codes.[25] Bloodstream infection was defined as a positive blood culture with any organism in the absence of a positive culture with the same organism from another body site within the previous 14 days. Urinary tract infection was defined as a positive urine culture with any organism ($\geq 10^5$ CFU/mL or 10^3–10^5 CFU/mL plus pyuria). Surgical site infection was defined as a positive wound culture with any organism within 30 days following a National Healthcare Safety Network designated procedure.[26] *C. difficile* infection was defined by positive stool culture. Pneumonia was defined as a positive respiratory culture with any organism plus any ICD-9/10-CM pneumonia code. Any infection that occurred >2 calendar days after admission was considered healthcare-associated and was included in the study. Lengths of stay and mortality outcomes were obtained from hospital administrative records.

Simulation modeling

We used a non-Markovian simulation to estimate daily conditional probabilities of bloodstream infection, urinary tract infection, surgical site infection, *C. difficile* infection, pneumonia, length of stay, and in-hospital mortality. In brief, each patient characteristic, unit characteristic, and hospital event in Table 2 was assigned a unique integer. A patient's current state (X_t) is a value determined by the presence and order of onset of these integers, representing all that the patient has experienced to date during the admission. The model $(X_{\{t+1\}}, h_{\{t+1\}}) = f\{(X_t, h_t), \theta, u_{\{t+1\}}\}$ outputs the patient's next state (X_{t+1}) and an updated memory of previous states (h_{t+1}) as a function of the patient's current state (X_t), the previous memory (h_t), model parameters θ, and uniform random variable u_{t+1}. A detailed description of the model is included in Appendix 1 (online), where the model is parameterized using recurrent neural networks.

Model validation

To validate the model, we took an independent random sample of 100,000 patient admissions and simulated a path to discharge (or death) for each. We then compared the average mortality, infection, and length of stay implied by the simulation model with the observed metrics to see how closely the output of the simulation resembles real events. To measure the predictive performance of the model, we devised a comparison with logistic regression using the patient and unit variables available at admission to predict *C. difficile* infection, noting that this is meant to establish the soundness of the model, as our model formulation can accommodate more complex scenarios than traditional supervised learning predictive models. *C. difficile* was chosen for the validation because the factors that affect risk are generalized and not heavily dependent on specific events such as surgery or indwelling device placement. We compared model performance using area under the receiver operating characteristic curve (AUC of ROC) as a metric. To illustrate the potential clinical utility of the model, we present case studies illustrating daily changes in outcome probability based on patient and unit characteristics and events.

Nurse staffing adequacy

To evaluate whether incidence of central-line–associated bloodstream infection, catheter-associated urinary tract infection, *C. difficile* infection, or pneumonia vary according to nurse staffing adequacy, we calculated infection incidence by unit-level nurse staffing (total registered nurse, licensed practical nurse, and nursing assistant hours per patient day by tertile) and unit-level patient care intensity (Nursing Intensity of Care Index by tertile) on the day of infection.

Results

Model performance

Table 3 presents performance metrics for the simulation model. Most metrics implied by the model fall within 95% confidence intervals of the actual estimates ($\pm 1.96\sigma_n$), meaning that the simulation model captures the structure and order of patient transitions from admission to discharge. The predictive power of our simulation model was slightly higher than that of the logistic regression model (AUC 0.852 and 0.816, respectively) (Fig. 1). The case studies presented in Figure 2 illustrate how our simulation

https://doi.org/10.1017/ice.2021.114 Published online by Cambridge University Press

Table 1. The Nursing Intensity of Care Index Variables and Scoring[a]

Level	Variable	Scoring	Data Source
Patient	Charlson Comorbidity Index (for adults aged ≥18 y) or Pediatric Chronic Complex Condition Index (for children aged <18 y)	0–33	ICD-9/10-CM POA
	Medications, blood products, and feedings		
	Blood product administration, whole blood transfusion	2	Provider order
	Enteral infusion, tube feeding, gastric lavage	2	Provider order
	Injection, oral, topical medications[b]	0–3	MAR
	Intravenous medications[b]	0–3	MAR
	Nebulizer, airway inhalation medications[b]	0–3	MAR
	Procedures		
	Hemodialysis performed on unit	1	Provider order
	Preparation for surgical procedure	1	Perioperative records
	Thoracentesis on unit with unit nurse assist	1	ICD-9/10-CM and perioperative records
	Devices, catheters, and continuous therapies		
	Alcohol or drug detoxification	3	ICD-9/10-CM
	Central venous catheter maintenance	1	EMR structured documentation
	Continuous renal replacement therapy	3	Provider order
	Extracorporeal membrane oxygenation	3	ICD-9/10-CM
	Implantable cardiac assist device	3	ICD-9/10-CM
	Isolation precautions	3	Provider order
	Mechanical ventilation	2	ICD-9/10-CM
	Ostomy	2	Provider order
	Peritoneal dialysis	3	MAR
	Peripherally inserted central catheter	1	Provider order
	Restraints	1	Provider order
	Urinary catheter insertion, indwelling	1	EMR structured documentation
	Urinary catheter maintenance, indwelling	1	EMR structured documentation
	Urinary catheter, intermittent	2	Provider order
Unit	Admissions, discharges, transfers	Total count	Administrative records
	Outbreak period	2	NORA reports
	Intensive care	2	Administrative records

Note. HR, human resources; ICD-9/10-CM, *International Classification of Diseases 9th or 10th Revision, Clinical Modification*; MAR, medication administration record; NORA, Nosocomial Outbreak Reporting Application; POA, present on admission.
[a]The Nursing Intensity of Care Index is calculated daily for each patient in a unit and averaged to create a daily score for each unit.
[b]Scoring of medications was determined by tertile of total number administered: 0=none, 1=first tertile, 2=second tertile, 3=third tertile.

model can provide insights beyond traditional supervised learning methods by accounting for the timing of each additional risk bearing event as well as the history and sequence of previous risk bearing events to calculate a more individualized real-time assessment of the probability of infection.

Nurse staffing adequacy and infection

Table 4 presents a matrix showing the percent of patients with central-line–associated bloodstream infection, catheter-associated urinary tract infection, *C. difficile* infection, and pneumonia under a range of nurse staffing and nursing care intensity conditions. For all 4 types of infection, incidence was greatest when staffing was lowest and nursing care intensity was highest.

Discussion

Prediction models for identifying preventable hospital-acquired conditions, such as HAIs, have great promise for improving patient outcomes.[1] Still, the utility of prediction models in patient care remains limited due to low predictive power, lack of timely, actionable interventions, or inability to capture ongoing changes in patient risk.[27,28] In this study we addressed 2 major limitations of HAI prediction models. First, we used a novel approach to modeling the likelihood of infection, death, and length of stay by estimating daily probabilities that are conditional not only on the events that the patient experienced to date during their hospital stay but also the sequence of those events. Second, we used a novel approach to measuring nurse

Table 2. Patient and Unit Characteristics Included in the Simulation Model

Timing	Variable	Data Source
Integers assigned once at admission	Age (<1, 1–4, 5–12, 13–21, 22–54, 55–64, 65–74, 75–84, ≥85 y)	Administrative records
	Sex (yes, no, unknown)	Administrative records
	Malignancy (yes, no)	ICD-9/10-CM
	Diabetes mellitus (yes, no)	ICD-9/10-CM
	Charlson Comorbidity Index (0–6, 7–13, ≥14)	ICD-9/10-CM POA
	Initial unit of admission (intensive care, stepdown, medical, surgical, medical/surgical, other, unknown)	Administrative records
Integers assigned on each day the event occurred (all yes/no)	**Medications, blood products, and feedings**	
	Antibiotic agents	MAR
	Blood product administration, whole blood transfusion	Provider order
	Chemotherapeutic, immunosuppressive, anti-inflammatory agents	MAR
	Enteral infusion, tube feeding, gastric lavage	Provider order
	Injection, oral, topical medications	MAR
	Intravenous medications	MAR
	Nebulizer, airway inhalation medications	MAR
	Procedures	
	Biopsy	ICD-9/10-CM
	Cardiac catheterization	ICD-9/10-CM
	Catheter angiography	ICD-9/10-CM
	Coronary angioplasty	ICD-9/10-CM
	Endotracheal intubation	ICD-9/10-CM
	Feeding tube insertion	ICD-9/10-CM
	General anesthesia	Perioperative records
	Hemodialysis performed on unit	Provider order
	Intra abdominal vascular shunt insertion	ICD-9/10-CM
	Operating room procedure (>30 minutes)	Perioperative records
	Operating room procedure (any)	Perioperative records
	Thoracentesis performed on unit	ICD-9/10-CM and perioperative records
	Transplant (major organ)	ICD-9/10-CM
	Vascular stenting	ICD-9/10-CM
	Devices and continuous therapies	
	Alcohol or drug detoxification	ICD-9/10-CM
	Central venous catheter	EMR structured documentation
	Continuous renal replacement therapy	Provider order
	Dialysis	ICD-9/10-CM
	Extracorporeal membrane oxygenation	ICD-9/10-CM
	Implantable cardiac assist device	ICD-9/10-CM
	Mechanical ventilation	ICD-9/10-CM
	Ostomy	Provider order
	Peritoneal dialysis	MAR
	Peripherally inserted central catheter	Provider order
	Restraints	Provider order
	Urinary catheter, indwelling	EMR structured documentation
	Urinary catheter, intermittent	Provider order

(*Continued*)

https://doi.org/10.1017/ice.2021.114 Published online by Cambridge University Press

Table 2. (*Continued*)

Timing	Variable	Data Source
	Hospitalization characteristics	
	Intensive care	Administrative records
	Isolation precautions	Provider order
	Overnight stay (ie, not discharged that day)	Administrative records
	Clinical outcomes	
	Death	Administrative records
	Bloodstream infection	Infection identification algorithm
	Clostridioides difficile infection	Infection identification algorithm
	Pneumonia	Infection identification algorithm
	Surgical site infection	Infection identification algorithm
	Urinary tract infection	Infection identification algorithm
Integers assigned daily by unit	Nursing Intensity of Care Index score (continuous)	See Table 1
	Proportion of registered nurse hours to total nursing hours (categorical by tertile)	HR records
	Total registered nurse, licensed practical nurse, and nursing assistant hours per patient day (categorical by tertile)	HR records

Note. HR, human resources; ICD-9/10-CM, *International Classification of Diseases 9th or 10th Revision, Clinical Modification*; MAR, medication administration record; POA, present on admission.

Table 3. Key Metrics Implied by the Simulation Model Compared to Reality After Simulating 100,000 Patient Paths

Outcome	Implied by the Simulation Model (95% Confidence Interval)	Observed
Length of stay, d	6.62 (6.52–6.72)	6.94
Death, %	1.8 (1.49–2.11)	1.76
Urinary tract infection, %	1.64 (1.51–1.76)	1.84
Bloodstream infection, %	0.75 (0.58–0.92)	0.72
Pneumonia, %	0.44 (0.38–0.50)	0.52
Clostridioides difficile infection, %	0.92 (0.73–1.11)	0.96
Surgical site infection, %	0.24 (0.20–0.28)	0.25

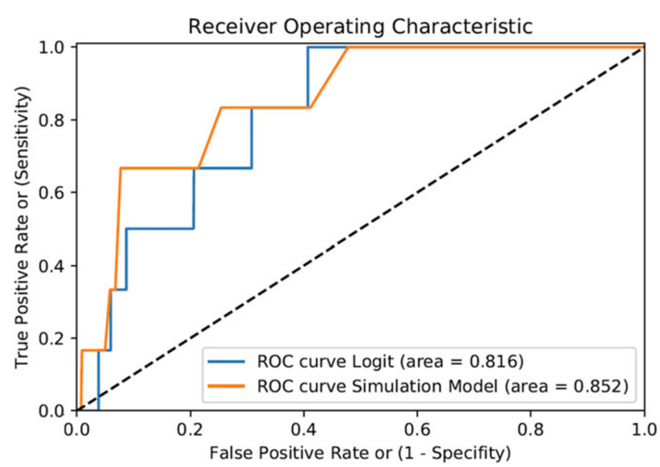

Figure 1. Baseline comparison between simulation model and logistic regression using only variables available at admission to predict *Clostridioides difficile*.

staffing by considering both staffing supply and patient care demand using the Nursing Intensity of Care Index.[17]

To our knowledge, this is the first study to incorporate staffing adequacy into a prediction model for HAIs. Countless studies have examined the association between nurse staffing and preventable complications, including HAIs.[29,30] The evidence strongly suggests that staffing plays some role in infection risk, though findings are not consistent across studies. A review by Shang et al[15] reported that inconsistent findings are likely explained by differences in the type and quality of staffing supply data sources. Our findings suggest that mixed results may also be due to whether and how studies account for patient care demand.

Despite challenges obtaining accurate staffing supply data, definitions of staffing levels and skill mix are relatively consistent among high-quality studies, with nurse-to-patient and registered nurse (RN)-to-support staff ratios among the most common metrics.[15] Patient care demand, on the other hand, is more difficult to define and measure, and meaningful methods for capturing patient care demand are not incorporated in most studies.[15,16,31] Severity of illness indices designed for mortality risk adjustment are sometimes used as proxies for how much nursing time patients require, but comorbidity burden does not necessarily correlate with inpatient nursing need.[31] Case-mix indices based on demographics, procedures, and

Table 4. Infection Incidence by Level of Nurse Staffing and Care Intensity

		Nurse Staffing		
Infection Type	Nursing Intensity of Care Index	Low	Moderate	High
Catheter-associated urinary tract infection	Low	1.16 (0.00069)	1.13 (0.00056)	1.28 (0.000004)
	Moderate	1.24 (0.000005)	1.29 (0.00045)	1.44 (0.00057)
	High	**1.56 (0.00042)**	1.49 (0.00052)	1.27 (0.00081)
Central-line––associated bloodstream infection	Low	0.44 (0.00043)	0.45 (0.00035)	0.46 (0.00024)
	Moderate	0.40 (0.00029)	0.47 (0.00027)	0.49 (0.00034)
	High	**0.57 (0.00026)**	0.46 (0.00029)	0.41 (0.00046)
Clostridioides difficile infection	Low	0.60 (0.000005)	0.61 (0.00041)	0.55 (0.00026)
	Moderate	0.53 (0.00033)	0.49 (0.00028)	**0.62 (0.00038)**
	High	**0.62 (0.00027)**	0.58 (0.00032)	0.49 (0.00051)
Pneumonia	Low	0.48 (0.00045)	0.54 (0.00038)	0.53 (0.00026)
	Moderate	0.53 (0.00033)	0.57 (0.000003)	0.64 (0.00038)
	High	**0.65 (0.00027)**	0.56 (0.00032)	0.32 (0.00041)

Data are infection incidence per 100 patients (standard error). The highest infection incidence within the matrix for each infection type is shown in bold.

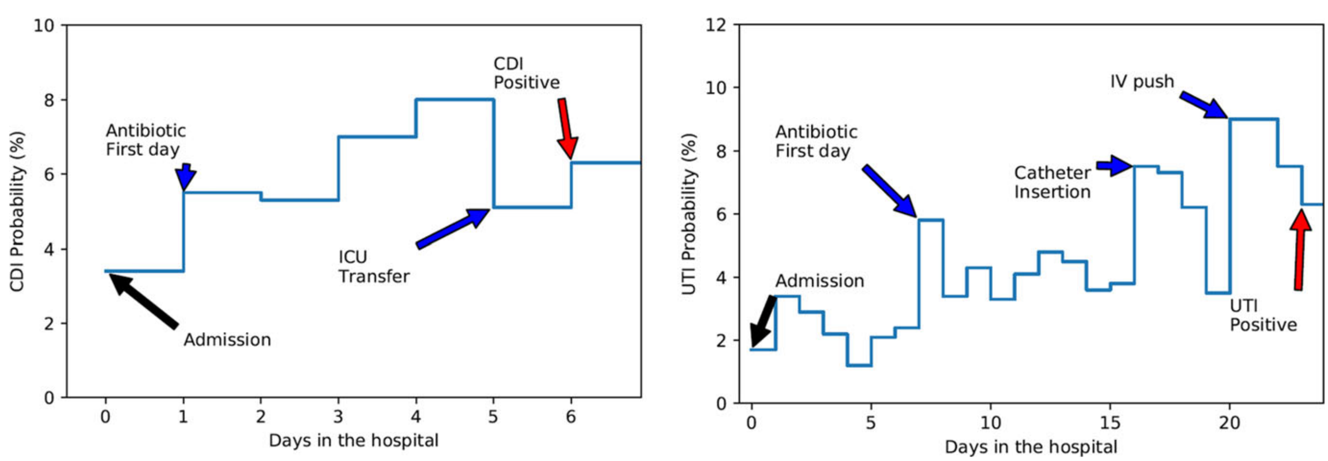

Figure 2. In **Case Study A**, a 13-year-old female with a malignancy was admitted with a low Pediatric Chronic Complex Condition score (0-6) to a unit with higher staffing (upper tertile) and moderate Nursing Intensity of Care unit score (middle tertile). She developed *Clostridioides difficile* infection (CDI) on day six after taking antibiotics for six consecutive days. Her daily risk of *C. difficile* infection estimated from the simulation model is plotted and annotated with landmark risk factors, illustrating how the probability of infection progressively increases with each consecutive day of antibiotics and following transfer to the intensive care unit (ICU). This case study highlights the usefulness of the simulation model to jointly assess the risk of an outcome that is caused by multiple co-dependent time-varying factors occurring simultaneously, which is difficult to achieve with traditional supervised machine learning methods. In **Case Study B**, a 62-year-old female diabetic patient with a malignancy was admitted with a high Charlson Comorbidity score (7-13) to a unit with lower staffing (lower tertile) and higher Nursing Intensity of Care unit score (upper tertile). She developed a urinary tract infection (UTI) on day 23 after seven days with a urinary catheter. Her daily risk of urinary tract infection is plotted, which illustrates how the probability of infection progressively increases after an initial spike following antibiotic administration beginning on day seven, urinary catheter insertion on day 16, and the addition of new IV push medications on day 20.

comorbidities are used to estimate nursing needs at the unit level, yet such measures are insensitive to changes in care demand throughout a patient's stay.[32] Patient classification and acuity systems incorporate a more holistic perspective of nursing care needs; however, these are generally designed to categorize patients into broad categories of low, medium, and high need.[33]

To improve measurement of nursing care demand, we developed and validated the Nursing Intensity of Care Index to capture patients' nursing needs on a daily basis to account for the fact that needs change over the course hospitalization and depend on more than medical diagnoses and demographic characteristics.[17,31] Using this index, we demonstrated that infection incidence was greatest when units had the lowest nurse staffing and highest patient care intensity. There is no clear dose–response relationship between staffing and nursing care intensity aside from the most extreme categories, which may suggest a threshold at which the staffing versus intensity ratio becomes unsafe. This warrants further analysis. Although it was not possible to precisely account for incubation period in this analysis, meaning that infection incubation could have begun before the time of staffing measurement, these findings suggest that the Nursing Intensity of Care Index may be useful tool to aid hospitals in safe staffing allocation.

Our study to improve prediction modeling for HAIs has some limitations. First, the data used to train the model are limited to a single health system. Although we aimed to create a model that

https://doi.org/10.1017/ice.2021.114 Published online by Cambridge University Press

can be generalized and operationalized elsewhere by including only electronic health data that is widely available across institutions in the United States, we acknowledge that local methods of data capture, billing code practices, order sets, and EMR structure may vary in ways that affect the predictive value of some variables. In addition, hospitals serving different patient populations or offering a different balance of services may need to incorporate other variables that impact infection risk, nursing care intensity, or both. Finally, the usefulness of prediction models for preventable hospital complications is determined in large part by whether they are integrated into practice such that real-time intervention to change a patient's infection probability is possible. Although our model identified modifiable factors that impact risk of infection, we did not deploy the model in clinical practice to assess its impact on infection rates in a real-world setting; however, the real-time availability of the data elements used for prediction would make this possible.

In summary, this study built on previously developed methods for predicting infection risk using a modeling approach that considers the order in which risk factors occur and incorporates a holistic consideration of staffing adequacy. The model had high predictive value for determining a patient's risk of infection on each day of hospitalization, and our results suggest that patients have the greatest risk of infection when unit-level staffing is low and patient care demand is high.

Acknowledgments.

Financial support. This work was supported by a grant from the Agency for Healthcare Research and Quality (grant no. R01 HS024915).

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

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/ice.2021.114

References

1. Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clin Infect Dis* 2018;66:149–153.
2. Roth JA, Battegay M, Juchler F, Vogt JE, Widmer AF. Introduction to machine learning in digital healthcare epidemiology. *Infect Control Hosp Epidemiol* 2018;39:1457–1462.
3. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? a primer for the epidemiologist. *Am J Epidemiol* 2019;188:2222–2239.
4. Beeler C, Dbeibo L, Kelley K, *et al.* Assessing patient risk of central line-associated bacteremia via machine learning. *Am J Infect Control* 2018;46:986–991.
5. Thorsen-Meyer H-C, Nielsen AB, Nielsen AP, *et al.* Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. *Lancet Digit Heal* 2020;2:e179–e191.
6. Shang J, Needleman J, Liu J, Larson E, Stone PW. Nurse staffing and healthcare-associated infection, unit-level analysis. *J Nurs Adm* 2019;49:260–265.
7. Mitchell BG, Gardner A, Stone PW, Hall L, Pogorzelska-Maziarz M. Hospital staffing and health care–associated infections: a systematic review of the literature. *Jt Comm J Qual Patient Saf* 2018;44:613–622.
8. Wiens J, Guttag J, Horvitz E. Patient risk stratification with time-varying parameters: a multitask learning approach. *J Mach Learn Res* 2016;17:1–23.
9. Revuelta-Zamorano P, Sánchez A, Rojo-Álvarez JL, Álvarez-Rodríguez J, Ramos-López J, Soguero-Ruiz C. Prediction of healthcare associated infections in an intensive care unit using machine learning and big data tools. *IFMBE Proc* 2016;57:834–839.
10. Harper EM. Staffing based on evidence: Can health information technology make it possible? *Nurs Econ* 2012;30:262–267, 281.
11. Noaman AY, Ragab AHM, Al-Abdullah N, Jamjoom A, Nadeem F, Ali AG. Predicting and reducing "hospital-acquired infections" using a knowledge-based e-surveillance system. *Expert Syst* 2020;37:e12402.
12. Sánchez-Hernández, Ballesteros-Herráez, Kraiem, Sánchez-Barba, Moreno-García. Predictive modeling of ICU healthcare-associated infections from imbalanced data. using ensembles and a clustering-based undersampling approach. *Appl Sci* 2019;9:5287.
13. Kalisch BJ, Friese CR, Choi SH, Rochman M. Hospital nurse staffing: choice of measure matters. *Med Care* 2011;49:775–779.
14. Jackson M, Chiarello LA, Gaynes RP, Gerberding JL. Nurse staffing and healthcare-associated infections: proceedings from a working group meeting. *J Nurs Adm.* 2002 Jun;32(6):314–322. doi: 10.1097/00005110-200206000-00007.
15. Shang J, Stone P, Larson E. Studies on nurse staffing and health care-associated infection: methodologic challenges and potential solutions. *Am J Infect Control* 2015;43:581–588.
16. Morris R, MacNeela P, Scott A, Treacy P, Hyde A. Reconsidering the conceptualization of nursing workload: literature review. *J Adv Nurs* 2007;57:463–471.
17. Larson EL, Cohen B, Liu J, Zachariah P, Yao D, Shang J. Assessing intensity of nursing care needs using electronically available data. *Comput Informatics Nurs* 2017;35:617–623.
18. Khanafer N, Voirin N, Barbut F, Kuijper E, Vanhems P. Hospital management of *Clostridium difficile* infection: a review of the literature. *J Hosp Infect* 2015;90:91–101.
19. Marschall J, Mermel LA, Fakih M, *et al.* Strategies to prevent central-line–associated bloodstream infections in acute-care hospitals: 2014 update. *Infect Control Hosp Epidemiol* 2014;35:753–771.
20. Gould CV, Umscheid CA, Agarwal RK, Kuntz G, Pegues DA. Guideline for prevention of catheter-associated urinary tract infections. Centers for Disease Control and Prevention website. https://www.cdc.gov/infectioncontrol/pdf/guidelines/cauti-guidelines-H.pdf. Published 2019. Accessed October 30, 2020.
21. Boltey E, Yakusheva O, Costa DK. 5 Nursing strategies to prevent ventilator-associated pneumonia. *Am Nurse Today* 2017;12:42–43.
22. Pássaro L, Harbarth S, Landelle C. Prevention of hospital-acquired pneumonia in non-ventilated adult patients: a narrative review. *Antimicrob Resist Infect Control* 2016;5:43.
23. De Cordova PB, Lucero RJ, Hyun S, Quinlan P, Price K, Stone PW. Using the nursing interventions classification as a potential measure of nurse workload. *J Nurs Care Qual* 2010;25:39–45.
24. Nosocomial outbreak reporting. New York State Department of Health website. https://www.health.ny.gov/professionals/diseases/reporting/communicable/infection/98-07-16_nosocomial_outbreak_reporting.htm. Published 2014. Accessed October 30, 2020.
25. Apte M, Neidell M, Furuya EY, Caplan D, Glied S, Larson E. Using electronically available inpatient hospital data for research. *Clin Transl Sci* 2011;4:338–345.
26. Surgical site infection (SSI) event. Centers for Disease Control and Prevention website. https://www.cdc.gov/nhsn/pdfs/pscmanual/9pscssicurrent.pdf. Published 2020. Accessed October 30, 2020.
27. Scardoni A, Balzarini F, Signorelli C, Cabitza F, Odone A. Artificial intelligence-based tools to control healthcare associated infections: a systematic review of the literature. *J Infect Public Health* 2020;13:1061–1077.
28. Luz CF, Vollmer M, Decruyenaere J, Nijsten MW, Glasner C, Sinha B. Machine learning in infection management using routine electronic health records: tools, techniques, and reporting of future technologies. *Clin Microbiol Infect* 2020;26:1291–1299.
29. Aiken LH, Sloane DM, Bruyneel L, *et al.* Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. *Lancet* 2014;383:1824–1830.
30. Schreuders LW, Bremner AP, Geelhoed E, Finn J. The relationship between nurse staffing and inpatient complications. *J Adv Nurs* 2015;71:800–812.

31. van Oostveen CJ, Ubbink DT, Huis in het Veld JG, Bakker PJ, Vermeulen H. Factors and models associated with the amount of hospital care services as demanded by hospitalized patients: a systematic review. *PLoS One* 2014;9:e98102.
32. Garcia AL. Variability in acuity in acute care. *J Nurs Adm* 2017;47: 476–483.
33. Defining staffing: workforce management patient classification and acuity systems the request for proposal process. American Nurses Association website. https://www.nursingworld.org/˜497e37/globalassets/practiceandpolicy/work-environment/nurse-staffing/website-staffing-and-acuity-systems-pdf-final_2017.pdf. Published 2017. Accessed October 30, 2020.
34. Resnick SI. *Adventures in Stochastic Processes*. Boston: Birkhäuser; 2002.
35. Cho K, Van Merriënboer B, Gulcehre C, *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation. Paper presented at: 2014 Conference on Empirical Methods in Natural Language Processing. Proceedings of the Conference Association for Computational Linguistics 2014:1724–1734.

https://doi.org/10.1017/ice.2021.114 Published online by Cambridge University Press