An Agent-Based Model for Evaluating Surveillance Methods for Catheter-Related Bloodstream Infection

Michael A. Rubin, MD, PhD[1,2], Jeanmarie Mayer, MD[1], Tom Greene, PhD[1], Brian C. Sauer, PhD[1,2], Bala Hota, MD, MPH[3], William Trick, MD[4], John A. Jernigan, MD[5], and Matthew H. Samore, MD[1,2], for the CDC Prevention Epicenter Program
[1]University of Utah, Salt Lake City, UT; [2]Salt Lake City VAMC IDEAS Center, Salt Lake City, UT; [3]Rush University, Chicago, IL; [1]Stroger Hospital of Cook County, Chicago, IL; [5]Centers for Disease Control and Prevention, Atlanta, GA

Abstract

Surveillance for catheter-related bloodstream infections (CRBSI) is hindered by the fact that clinical case criteria are complex and subjective. Simplified objective criteria, based only on microbiologic data, may be a less valid, but potentially more reliable system for estimating and comparing institutional infection rates. We developed an agent-based simulation model to examine the impact of these two different criteria on the measurement of CRBSI in a simulated 12-bed hospital intensive care unit (ICU). We found that, on average, the clinical criteria was more accurate at estimating the true CRBSI rate than the simple criteria (3.36+/-1.11 vs. 5.41+/-1.36 infections/1000 catheter-days, compared with a true rate of 3.54+/-1.60). However, ecologic correlation (i.e., the accurate ranking of CRBSI rates across institutions) was higher for simple criteria than clinical criteria. Thus, simplified objective criteria are potentially superior to clinical criteria in identifying the true differences in CRBSI rates between institutions.

Introduction

Public disclosure of hospital-acquired infection (HAI) rates is gaining momentum across the United States, as public officials seek to provide health consumers with the ability to make better-informed decisions about their care[1-4]. The generation and interpretation of HAI rates, however, is complex. Traditional HAI surveillance, led by infection control practitioners (ICPs), is a laborious process involving highly subjective case criteria which are not true "gold standards." The application of these criteria can also be highly biased by such factors as ICP training and experience, individual hospital practices, and specific hospital characteristics. This raises serious concerns about the reliability of these criteria and the rates calculated by different individuals applying the criteria to their data. As such, the accurate interpretation of these data is implausible, and their use by the public for comparing the performance of different facilities may be inappropriate and more confusing than helpful[5].

Alternative, simplified methods of surveillance that rely entirely on objective criteria (such as microbiology data extracted from an electronic medical record[6]) present an appealing option because of their ability to be automated and because their reliability is potentially much greater. These methods may not be as accurate at estimating true infection rates, however, because they draw upon a smaller pool of data and, as a result, are less selective. Still, a system that is objective and much more reliably applied at different institutions could be more useful when the ultimate goal is the comparison of rates across those institutions.

Because there is no true "gold standard" for diagnosing CRBSI, it is difficult to evaluate and compare these two surveillance approaches using real-world data. Instead, we chose to create an agent-based simulation model to simulate the occurrence and surveillance of CRBSI, since this approach allowed us to definitively know which patients had true infections. We then explored the tradeoff between reliability and validity when using the two surveillance methods to (a) estimate an institution's true CRBSI rate, and (b) compare the ranking of estimated rates across multiple institutions.

Methods

Agent-based simulation. We used a type of computer simulation called agent-based modeling, a technique in which a system is modeled as a collection of autonomous decision-making entities called agents. Each agent can be assigned internal states and behaviors; interactions between agents can then produce complex, emergent system-level dynamics and behavior patterns. The advantages of this technique for this study are its flexibility and its ability to create completely synthetic data sets that resemble real-word data sets and are suitable for statistical analysis.

<u>Overview.</u> The agent-based model was constructed in Anylogic 5.5 (XJ Technologies, St. Petersburg, Russia) using Java-based graphical editing tools to create replicated agents and their associated

parameters, variables, and statecharts. The base model simulated a cohort of patients admitted to a 12-bed intensive care unit (ICU) over a three-year time period. The broad overview of the simulation is as follows (see **Figure 1**): During the simulation, patients flowed in and out of the ICU, had central venous catheters inserted and removed, and were at risk of developing infections. Catheterized patients were at risk for developing CRBSI and other types of infection, while non-catheterized patients were at risk only for other types of infection. The infection status of a patient influenced the discharge rate and the development of signs and symptoms of infection, such as fever; in turn, the symptom status of a patient influenced the ordering of blood cultures and cultures from other body sites. The result of these cultures was impacted by the patient's underlying infection status. At the time of death or discharge, surveillance was performed on the simulated data collected during each patient's ICU stay, using a process that involved the interpretation of symptom/sign data and culture results.

The model thus simulates four main processes: a patient flow process, an infection event process, a clinical event process, and a surveillance classification process (**Figure 1**). Each of these four processes is described in more detail below.

Patient Flow. Patient flow was dictated primarily by the discharge hazard rate, which was modeled as a cubic spline function of time since admission, using local ICU length-of-stay data, to produce an average stay of approximately 2.8 days. The hazard function was designed such that the presence of an infection decreased a patient's daily discharge hazard by 25%. Death was combined into the discharged state.

Clinical Events. During the simulated hospital stay, clinical events could occur which would alter critical internal states of the patient. First, central venous catheter insertion and removal events were implemented as constant (time-independent) hazard rates, to produce mean times to catheterization and catheter removal of 2 days and 10 days, respectively. Next, we modeled the development of signs and symptoms of infection. NHSN criteria for CRBSI surveillance specify that at least one of fever, chills, or hypotension should be present in the context of blood cultures positive for common skin commensals (CSCs) in order to meet criteria for a CRBSI[7]. We therefore chose to model all three signs/symptoms above as a single binary variable, where a value of true indicated the presence of at least one of the three criteria. The rate of symptom development was also approximated using a constant hazard rate to produce a mean time to symptom development of

approximately 1 day.

The ordering of cultures from blood and other body sites was tested on a daily basis for each patient, and was modeled as a Bernoulli probability based on multiple factors such as number of prior cultures positive for a true pathogen, number of prior cultures that were negative, and whether or not the patient had similar cultures obtained on the previous day of hospitalization. Culture results were also modeled as simple Bernoulli probabilities based primarily on the patient's underlying infection state (infected, not infected) and infection type (primary CRBSI or other infection). Repeatedly positive cultures were always assumed to be of the same organism.

Infection Event. The rates of infection for both CRBSI and infections at other sites were modeled as time-dependent hazard rates governed by individual gamma distributions. Patients were allowed to have only one episode of infection per ICU admission.

Parameter	Value
ICU Rooms per hospital	12
Min days from admission for positive culture to be considered nosocomial	2
Max days between two positive CSC cultures to be considered infection	2
Mean days to insertion of CVC	2
Mean days to removal of CVC	10
Mean days to symptom development	1
Mean days to CRBSI infection	14
Mean days to infections at other sites	10.5
Proportion of true CRBSI due to CSC	0.50
Days around a positive blood culture to check for other positive cultures	-3 to +7
ICP distance (d') (mean ± std dev)	2 ± 0.1
ICP specificity (mean)	0.833
ICP specificity (variance)	0.004

Table 1. Base values used for key experimental model parameters. (CSC = common skin commensals; CVC = central venous catheter)

Model parameters were derived from the literature or from local ICU data when available; when unavailable, plausible parameter values were estimated and tested with robustness analysis. **Table 1** provides the values assigned to many of the key parameters in the model. Individual-level patient data on each ICU stay were collected for the performance of surveillance upon discharge (or death), as below.

Surveillance. Surveillance for CRBSI was performed upon each patient discharge, and was modeled using two methods: subjective clinical (SC) surveillance by simulated ICPs applying adapted NHSN clinical case criteria; and objective simplified (OS) surveillance

similar to that which could be performed electronically using primarily microbiology data, as has been published previously[6]. During surveillance, the patient under consideration was assigned an SC score and an OS score, reflective of the type of surveillance method used. Each score has a threshold for interpretation, with scores above the threshold classified as a CRBSI. Thus, the fraction of true CRBSI and non-CRBSI patients classified as having a CRBSI defines the sensitivity and false positivity of the surveillance criteria.

Surveillance criteria. The subjective clinical (SC) criteria used by simulated ICPs was based on the widely-used NHSN criteria for the surveillance of catheter-related bloodstream infections published by CDC[7]. These criteria contain a number of subjective components, as shown in Table 3. The ability of ICPs to correctly interpret these subjective components was accomplished by modeling ICP reliability, as discussed in the following section. The objective simplified (OS) criteria used for automated surveillance analyzed simulated microbiologic data predominantly, following an ordered algorithmic approach adapted from work done previously by Trick et. al.[6]

ICP reliability. ICPs were modeled as separate agents, each with its own sensitivity and specificity for accurately detecting clinical criteria (e.g., presence or absence of clinical symptoms/signs). As we wanted to model the typical tradeoff between sensitivity and specificity, we utilized the following equation derived from Signal Detection Theory[8]:

$$d' = z(\text{sensitivity}) - z(1 - \text{specificity})$$

where d' (distance) is the distance between noise and (noise + signal), and z represents the z-score of the specified value. Distance thus represents the ability of an ICP to accurately separate true positives and false positives. For each individual ICP, we first randomly assigned values to specificity (from a beta distribution) and distance (from a normal distribution), and then calculated sensitivity as a function of specificity using the above equation.

Simulation conditions and analysis. The model was used to simulate 100 identically configured hospitals over a 3-year period. Each hospital had its own 12-bed ICU and its own true CRBSI rate. We then compared the ability of the two surveillance methods to (a) estimate each facility's true CRBSI rate, and (b) accurately maintain the rank order of all facilities. The values used for many of the various important model parameters are shown in **Table 1**.

To compare the rate ratios of SC criteria (relative to Truth) and OS criteria (relative to Truth), we fit mixed-effects Poisson regression. The model included SC and OS methods as indicator variables, where both are relative to Truth. The exponentiated regression coefficient for the two indicator variables estimated the relative rate for each method, relative to truth. We next compared the two regression coefficients to determine if the ratio of the estimated infection rates to the true infection rates was closer to unity for SC criteria than for OS criteria. Correspondence between the rankings provided by the two methods was assessed using Kendall's tau rank correlation method[9], which is an alternative measure of the proportion of pairs of hospitals with concordant ranks. The proportion of concordant pairs is equivalent to (tau + 1) / 2. A bootstrap simulation, using the bias-corrected method, was done to provide a statistical comparison of the two concordance measures. This simulation provided a 95% confidence interval around the difference between the concordances.

All statistical analyses were performed using the Stata 9.2 statistical software (StataCorp LP, College Station, TX).

Results

The results of our simulation experiment are provided in **Table 2**. When the rate estimates produced by each surveillance method were aggregated among the 100 simulated hospitals, we found that the mean rate estimate produced using the SC criteria (3.36 ± 1.11 CRBSI/1,000 catheter-days) was more reflective of the mean of the true underlying CRBSI rates (3.54 ± 1.60) than was the mean rate estimate produced using the OS criteria (5.41 ± 1.36). This was also reflected in the rate ratios for each group (**Table 2**); while the SC criteria underestimated the true CRBSI rate by 16% (rate ratio 0.84), the OS criteria overestimated the true rate by 38% (rate ratio 1.38; *P*<0.001).

We then used Kendall's tau rank correlation to assess the extent to which each surveillance method was able to maintain the appropriate rank order of institutions based on their true CRBSI rates. In fact, the Kendall's tau for the OS criteria (0.484) was higher than that for the SC criteria (0.405), indicating that the percentage of correlated measurements produced by the OS criteria (74%) was higher than that produced by the SC criteria (70%) [concordance difference = 4%; 95% CI (0.00% to 10.00%)]. Thus, the rate estimates generated using the OS criteria better reflected the appropriate rank order of institutions than those generated using the SC criteria, despite the fact that the rates estimated by the OS criteria exhibited a greater overall bias relative to the true rates.

Rate Estimates	CRBSI/1,000 catheter-days (mean ± SD)
True CRBSI Rate	3.54 ± 1.60
SC Estimated Rate	3.36 ± 1.11
OS Estimated Rate	5.41 ± 1.36

Rate Ratios	Rate Ratio	95% CI
True CRBSI Rate	referent	---
SC Criteria	0.84	0.80 - 0.88
OS Criteria	1.38*	1.32 - 1.44

Performance	Sensitivity	Specificity
SC Criteria	62.5% ± 7.2%	94.7% ± 2.1%
OS Criteria	82.4% ± 6.2%	86.4% ± 2.7%

Rank Correlation	Kendall's tau	% Concordant
SC Criteria	0.405	70%
OS Criteria	0.484	74%

Table 2. Results of the simulation experiment using the base parameters. (SC = subjective clinical; OS = objective simplified; SD = standard deviation) *$P<0.001$

Discussion

With the growing movement across the U.S. to require hospitals to report their hospital-acquired infection (HAI) rates[1-4], increasing scrutiny is being placed on the methodology used to estimate these rates. Because the traditional, ICP-based method is unreliable and subject to multiple biases, its use in comparing the performance of different facilities may be inappropriate[5]. Newer, objective methods of surveillance may be a less accurate but more reliable method of HAI surveillance, but their acceptance has thus far been modest.

Through the use of simulation, we were able to establish an experimental environment that provided a means of assessing the tradeoff between accuracy and reliability in the surveillance of one HAI, nosocomial CRBSI. Our findings demonstrated that, in a setting in which objective, simplified CRBSI surveillance criteria yielded a larger bias in the estimated CRBSI rate than subjective clinical criteria, the simplified criteria nonetheless provided more accurate rankings of the CRBSI rates between institutions. Thus, the more reliable objective criteria are most likely better suited for measuring and publicly reporting institutional CRBSI rates if the ultimate goal is to allow public comparisons between facilities.

How could one set of surveillance criteria be more accurate at estimating the true CRBSI rate of a facility, but less able to estimate the true differences in rates between facilities? Although this may seem contradictory, we believe this is a reflection of the effect of reliability in this setting. Currently, little is known about the inter-rater reliability of ICPs in the performance of infection surveillance. A recent study[6] assessing automated detection of bloodstream infections did evaluate the kappa statistic between ICP and investigator reviews of prospective bloodstream infection data and found a value of 0.37, which suggests a high degree of variability exists. This subject is currently a focus of our research and will be used to inform and validate future simulation work.

The SC criteria, being more valid, are also more selective (as reflected in their low sensitivity and high specificity). This results in overall rate estimations that are relatvely close to the true rates. However, because these criteria are applied inconsistently from one facility to the next—a reflection of reliability—the degree to which an individual rate estimation varies from the truth is less predictable and more likely to result in changes to the rank orders of facilities. In contrast, the simplified criteria, which are are less selective and thus lead to higher rate estimations, are nevertheless applied with great consistency across all facilities due to the absence of any reliability factor. Thus, the degree to which these estimations vary from the truth is also consistent across facilities, leading to a better preservation of the appropriate rank order.

These findings could potentially have a significant impact on how HAI surveillance is performed, and how infection rates are estimated, at hospitals nationwide. There has been considerable reluctance to the adoption of simplified, objective surveillance criteria to this point, most likely because of the perceived inaccuracy of this method. Still, as more and more hospitals across the country are required to report their HAI rates for the purpose of public reporting and comparison, findings such as these may convince decision makers that the reliability of the chosen surveillance criteria could be a more important consideration than overall validity.

The main limitations of this work relate to the fact that, as a simulation experiment, our model design relies on a large series of assumptions for processes and parameters that are, in many cases, unmeasured or unmeasurable. While we have attempted in most cases to provide accurate assumptions through the use of local hospital data, in many cases we were required to use plausible estimates for parameters for which there are little available data. Subsequent work will focus on further refining the model through iteration and calibration, as well as further collecting and incorporating missing or estimated parameters from real-world data. We are also now working to

determine the full range of settings in which the objective simplified criteria outperform the subjective clinical criteria.

Conclusion

With the current nationwide push toward mandatory public reporting of HAI rates, greater attention must be placed on the methods used for infection surveillance and the inconsistent manner with which case criteria are applied at different institutions. Our findings suggest that a more simplified set of surveillance criteria may improve the reliability and comparability of HAI rate estimates across institutions, despite the fact that such criteria might produce rate estimates that are generally less accurate at the inividual facility level. Still, for the purposes of mandatory public reporting of HAI rates, the sacrifice for improved reliability might be well worth the perception of less accuracy.

References

1. Edmond MB, Bearman GM. *Mandatory public reporting in the USA: an example to follow?* J Hosp Infect. 2007;65 Suppl 2:182-8.
2. McKibben L, Fowler G, Horan T, et al. *Ensuring rational public reporting systems for health care-associated infections: systematic literature review and evaluation recommendations.* Am J Infect Control. 2006;34:142-9.
3. McKibben L, Horan T, Tokars JI, et al. *Guidance on public reporting of healthcare-associated infections: recommendations of the Healthcare Infection Control Practices Advisory Committee.* Am J Infect Control. 2005;33:217-26.
4. Thompson JW, Pinidiya SD, Ryan KW, et al. *Health plan quality-of-care information is undermined by voluntary reporting.* Am J Prev Med. 2003;24:62-70.
5. Emori TG, Edwards JR, Culver DH, et al. *Accuracy of reporting nosocomial infections in intensive-care-unit patients to the National Nosocomial Infections Surveillance System: a pilot study.* Infect Control Hosp Epidemiol. 1998;19:308-16.
6. Trick WE, Zagorski BM, Tokars JI, et al. *Computer algorithms to detect bloodstream infections.* Emerg Infect Dis. 2004;10:1612-20.
7. Garner JS, Jarvis WR, Emori TG, et al. *CDC definitions for nosocomial infections, 1988.* Am J Infect Control. 1988;16:128-40.
8. Wickens TD. *Elementary Signal Detection Theory.* New York: Oxford University Press (2002).
9. Kendall M, Gibbons JD. *Rank correlation methods.* 5th ed. New York: Oxford University Press (1990).

① **PATIENT FLOW**

admission

② **CLINICAL EVENTS**

not catheterized

λ_2 λ_3

catheterized

asymptomatic

λ_4

symptomatic

blood culture ordering

other culture ordering

③ **INFECTION EVENT**

not infected

$\lambda_5(t)$ $\lambda_6(t)$

infected
CRBSI   Other

$\lambda_1(t)$

discharge

④ **SURVEILLANCE CLASSIFICATION**
• Clinical criteria (manual)
• Simple criteria (automated)

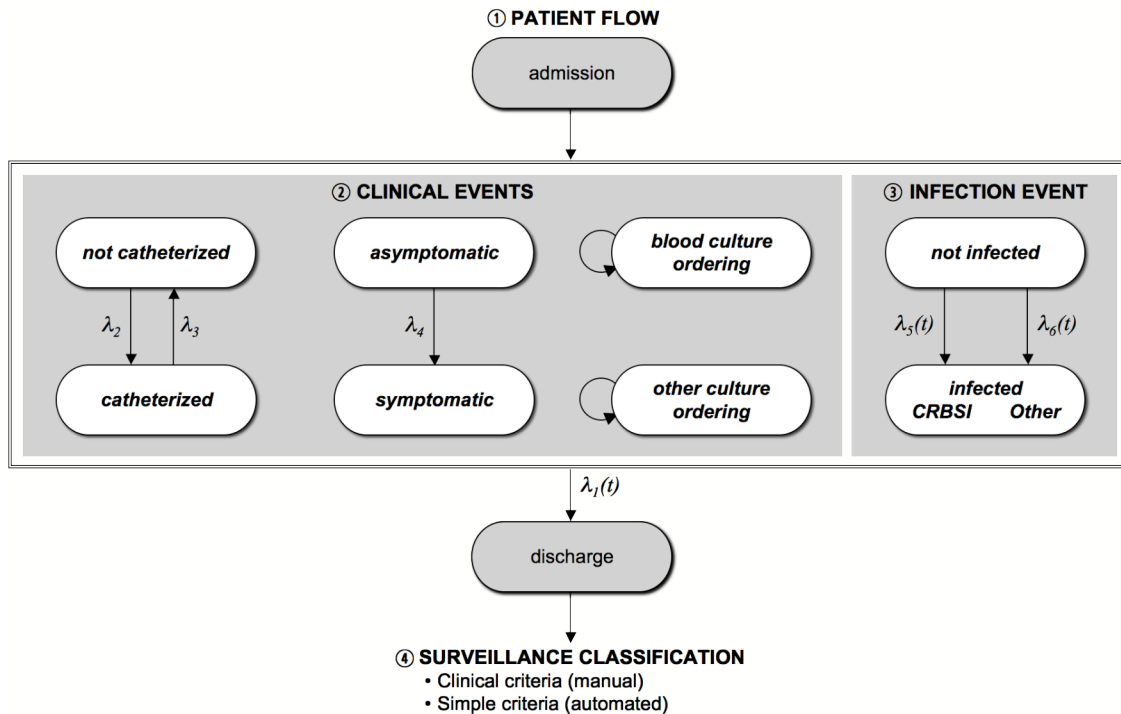**Figure 1.** Basic structure of the simulation, demonstrating the four main processes modeled in the simulation.