A Poisson-based Prediction Model and Warning System for MRSA Daily Burden

ROCCO J PERLA, BRADFORD D ALLEN

OBJECTIVE: This study was designed to demonstrate that the number of methicillin-resistant *Staphylococcus aureus* (MRSA) isolates collected daily in a community hospital is Poisson distributed and that using a one-sided Poisson control table is a fast and easy way to recognize unusually high numbers of MRSA isolates collected daily that may signal possible outbreaks.

METHODS: A retrospective analysis of MRSA isolates collected daily over a three year period (2005-2007, N = 934) was performed. Observed MRSA isolate frequencies are compared to Poisson frequencies using chi-square goodness-of-fit tests. A regression equation on the mean number of MRSA isolates collected daily for the years 2005, 2006, and 2007 is used to predict the mean number of MRSA isolates for 2008. A warning system for MRSA isolates collected daily is presented and a one-tailed, mean + 2 sigma control table is provided.

SETTING: One-hundred-fifty bed community hospital in central Massachusetts.

RESULTS: Goodness-of-fit tests showed close agreement between actual MRSA isolates collected daily and Poisson frequencies for 2005 ($\chi^2_4 = 4.045$, $p = 0.39$), 2006 ($\chi^2_4 = 2.807$, $p = 0.59$), and 2007 ($\chi^2_4 = 1.494$, $p = 0.83$).

CONCLUSION: Theoretical and empirical support is provided for the Poisson probability model. The model can be used to identify unusually high occurrences of MRSA isolates collected daily. This study was limited to a single community healthcare system but the results may be generalized to other types of healthcare settings.

ABBREVIATIONS: ICPs = infection control practitioners; MDROs = multi-drug resistant organisms; MRSA = methicillin-resistant *Staphylococcus aureus*.

INDEX TERMS: infection control; microbiology; MRSA; Poisson distribution; statistical process control.

Clin Lab Sci 2009;22(1):22

***Rocco J Perla EdD** is **IHI/George W Merck Fellow**, Institute for Healthcare Improvement, Cambridge MA; and **section head**, Clinical Microbiology & Diagnostic Immunology, and **epidemiologist**, Department of Infection Control and Prevention, HealthAlliance Hospital, Leominster MA.*

***Bradford D Allen EdD** is chair, Department of Mathematics and Science, Lasell College, Newton, MA.*

***Address for correspondence:** Rocco J Perla EdD, IHI/George W Merck Fellow, Institute for Healthcare Improvement, 20 University Road, 7th Floor, Cambridge MA 02138. (617) 301-4800, (617) 301-4848 (fax). rperla@IHI.org.*

***ACKNOWLEDGEMENT:** The authors wish to acknowledge Francis R Landry, MSLS, HealthAlliance Hospital Library Manager, for his assistance with the literature review for this article.*

One of the problems hospital microbiologists, epidemiologists, and infection control practitioners (ICPs) routinely face when dealing with the more frequently encountered multi-drug resistant organisms (MDROs) reported to the infection control service is to determine the point at which comprehensive follow-up is needed or justified. Currently, the ubiquity of some resistant organisms in hospitals and surrounding communities, such as methicillin-resistant *Staphylococcus aureus* (MRSA), can result in data that resemble white noise, particularly to overburdened laboratories and ICPs. Understanding and modeling the daily burden of resistant organisms is critical for good resource management, prompt epidemiologic intervention, and benchmarking by institutions looking to eradicate, or at least significantly reduce, MDRO frequency.[1-3] Further, because some resistant organisms are now a community and outpatient problem, the burden of these organisms needs to be addressed at all levels of service and care.[2]

The purpose of this article is to demonstrate that the number of MRSA isolates collected daily in a community hospital and reported to the infection control service can be described with the Poisson distribution. Moreover, using the Poisson distribution is an effective way to understand MRSA daily occurrence data and to make valid decisions

as to when epidemiologic follow-up is reasonable. To date, the number of MRSA isolates (or other MDROs) collected daily in a hospital setting have not been modeled with the Poisson distribution.

MATERIALS AND METHODS

Data collection

The statistics presented here provide a retrospective analysis of the number of MRSA isolates collected daily over a three year period (2005-2007) at a 150-bed, non-major teaching community hospital in central Massachusetts. The range of services provided by the testing laboratory was consistent over the three year period and included inpatient and outpatient services. Laboratory records of MRSA isolates recovered daily during the years 2005 to 2007 were collected using the Vitek DataTrac Logbook Report program (bioMerieux, Durham NC) following the "first isolate rule" (i.e., one patient isolate per year) and following Clinical Laboratory and Standards Institute guidelines for the analysis of susceptibility data.[4] The MRSA isolate data were grouped into daily occurrence categories (0 to ≥5 occurrences per day) for each of the three years. The grouping procedure was repeated on two separate occasions to insure reliability. The data included all clinical isolates (inpatients and outpatients), with duplicate patient isolates, surveillance cultures, and screens omitted from the analysis. The actual specimen collection date (not the date reported to infection control) was used to organize the data because the date of collection provides a reasonable and standard estimate of frequency across time. When the warning system presented here is used in real time, MRSA counts by date of collection should be used rather than counts by date reported to infection control so false warnings are not generated by batched reports.

Statistical analysis

To conclude that MRSA isolates collected daily are Poisson distributed, it was assumed that the occurrence or non-occurrence of MRSA isolates in any small time interval is a Bernoulli trial (an experiment with two possible outcomes e.g., yes/no or infection/no infection). It was then observed that the mean number of MRSA isolates collected daily was independent of time interval length or period (i.e., seasonally, monthly, weekly, daily). To test whether the observed frequencies of MRSA isolates collected daily were Poisson distributed, the number of MRSA isolates collected daily during the years 2005 to 2007 were compared to Poisson frequencies using chi-square goodness-of-fit tests. Poisson frequencies were computed using $e_i = np_i$ where n is the total number of MRSA isolates and p_i is the Poisson probability of i occurrences where i is 0, 1, 2, 3, 4, ≥5. Goodness-of-fit

Table 1. Distribution characteristics for MRSA frequencies per day across study period

Year	Number of MRSA isolates collected daily						Total	λ*	s^2†	Skew‡	CV§
	0	1	2	3	4	≥5					
2005											
N	210	105	39	9	2	0	218	0.60	0.67	1.38	1.11
%	57.5	28.8	10.7	2.5	0.5	0	100				
2006											
N	153	119	68	21	3	1	335	0.92	0.95	0.94	1.03
%	42.0	32.6	18.6	5.7	0.8	0.3	100				
2007											
N	132	131	65	30	5	2	381	1.04	1.08	0.95	1.03
%	36.2	35.9	17.8	8.2	1.4	0.5	100				

*lambda (mean of the Poisson distribution)
†variance (≈λ)
‡Skew (≈1/√λ)
§coefficient of variation
Note: categories 4 and ≥ 5 were combined for χ^{-2} analysis.

tests were used for each of the three years. To follow common goodness-of-fit test guidelines, categories with ≥4 occurrences of MRSA isolates per day for the years 2006 and 2007 were combined to ensure that no more than 20% of the expected frequencies were less than 5. After all tests of the Poisson model were confirmed (see the Results section), the Poisson distribution was used to develop a MRSA daily burden warning system. The mean, variance, skew, and coefficient of variation (variance/mean) of MRSA isolates collected per day were found and compared to the theoretical Poisson distribution for each of the three years. It should be noted that for the Poisson distribution, the mean and variance are equal. That is, the mean = λ (lowercase Greek letter lambda), variance = λ, coefficient of variation = 1, and skew = $1/\sqrt{\lambda}$.

Table 2. Selected lambda (mean) values and associated warning limits

λ	λ+2 (sigma)*	X ≥ λ+2(sigma) (follow-up recommended)	Probability X ≥ λ+2 (sigma) (chance of a false alarm)
1.00	3.00	X ≥ 3	0.08
1.05	3.10	X ≥ 4	0.02
1.10	3.20	X ≥ 4	0.03
1.15	3.29	X ≥ 4	0.03
1.20	3.39	X ≥ 4	0.03
1.25	3.49	X ≥ 4	0.04
1.30	3.58	X ≥ 4	0.04
1.35	3.67	X ≥ 4	0.05
1.40	3.77	X ≥ 4	0.05
1.45	3.86	X ≥ 4	0.06
1.50	3.95	X ≥ 4	0.07
1.55	4.04	X ≥ 5	0.02
1.60	4.13	X ≥ 5	0.02
1.65	4.22	X ≥ 5	0.03
1.70	4.31	X ≥ 5	0.03
1.75	4.40	X ≥ 5	0.03
1.80	4.48	X ≥ 5	0.04
1.85	4.57	X ≥ 5	0.04
1.90	4.66	X ≥ 5	0.04
1.95	4.74	X ≥ 5	0.05
2.00	4.83	X ≥ 5	0.05

*sigma is equal to the square root of lambda ($\sqrt{\lambda}$)

Finally, a regression equation for the mean based on the years 2005, 2006, and 2007 was found and used to predict the mean number of MRSA isolates for 2008.

RESULTS

Distribution analysis

Goodness-of-fit tests showed close agreement between the number of MRSA isolates collected daily and Poisson frequencies for 2005 ($\chi^2_4 = 4.045$, $p = 0.39$), 2006 ($\chi^2_4 = 2.807$, $p = 0.59$), and 2007 ($\chi^2_4 = 1.494$, $p = 0.83$). Close agreements in variance and skew were found between the Poisson distribution and the actual numbers of MRSA isolates collected daily for each of the three years (Table 1).

Prediction and warning limits

The mean of a Poisson distribution is the only parameter necessary to completely describe the Poisson distribution. To predict the mean number of MRSA isolates per day in 2008, a regression equation of the form $\lambda_p = a+bt$ was found and used to describe the linear trend of the mean during 2005, 2006, and 2007. The coefficient of determination of the equation was found to be $R^2=0.93$. The regression equation was used to predict the mean value (and thus the distribution) for 2008. Using the regression equation, the mean number of MRSA isolates collected daily is predicted to be $\lambda_p = 1.29$ (95% CI, 1 – 1.58). (The value of $\lambda_p = 1.29$ is very close to the predicted mean value of our longer-term logistic model for the years 2003 - 2007. The logistic model is not presented here). The predicted value of lambda for 2008 ($\lambda_p = 1.29$) can be used to construct a table that serves as a warning system for unusually high numbers of MRSA isolates collected per day. Isolate occurrences that exceed the "follow-up recommended" values presented in Table 2 would trigger epidemiologic follow-up. Table 2 presents a range of lambda (mean) values, lambda plus 2-sigma values, the numbers of occurrences where follow-up is recommended (≥ lambda plus 2-sigma), and the probabilities of a false alarm. Although it is standard practice to use 3-sigma warning limits in statistical process control models,[5] a more cautious 2-sigma warning limit is used here that is more sensitive to potential infection control problems (but is more likely to have false alarms). For example, using a 3-sigma model as a warning limit would produce false alarms only once in 769 days (1/.0013) on average, but would probably miss all the real warning signals. Of course, the follow-up warning limits can be adjusted in different situations where it may be more prudent to reduce the risk of false alarms.

DISCUSSION

The need to address aspects of infectious disease epidemiology, such as the use of more systematic, rigorous and logical methods of information management, has been advocated by many, particularly in the context of statistical process control[6,7] and signal detection theory.[8] Today, the increased frequency and emergence of resistant microorganisms and other public health threats necessitate the use of more formal decision-making systems that augment the clinical interpretive dimension of laboratory and infection control practice. It has long been recognized that the microbiology laboratory is an "early warning center" for potential infection control problems and that microbiologists and ICPs need to work together to develop optimal surveillance strategies.[9] There are some aspects of clinical microbiology and infection control practice that are ideally suited for probabilistic modeling and surveillance, including frequency, distribution, and time series analysis of specified microorganisms. However, many microbiologists and ICPs may not be familiar or comfortable with probability modeling, thereby limiting the use of these important tools. This study demonstrates that the frequency of MRSA isolates collected daily in a community hospital is Poisson distributed and that a Poisson model can be used to determine when isolate frequencies per day indicate the need for a higher level of attention or action while understanding the chance of acting on a false alarm. Although the Poisson model presented here focuses on MRSA, the model is also useful for monitoring other MDROs such as extended-spectrum β-lactamase producing gram negatives, or vancomycin-resistant enterococcus.

The model presented here is appropriate, easy to use, and accessible to a large number of microbiologists and ICPs who may not have advanced training in statistics. Further, the model may be implemented using control chart features for modeling Poisson data that are available in many statistical packages such as Stata (Stat Corporation, College Station TX) and SPSS (SPSS Inc. Chicago IL). In addition, the probability model also provides a means of correcting misperceptions in clinical judgment and intuition. For example, prior to beginning this study, hospital infection control and microbiology staff members were asked about the expected frequency of zero MRSA isolate days during 2007. All three ICPs and seven microbiologists believed there would be few or no days with no MRSA isolates collected. Their expectations were incorrect however because finding no MRSA isolates daily was the most frequent category in each of the three years studied (2005-2007). In this case, the infection control and microbiology staff thought the MRSA daily occurrence was far greater than it actually was.

Reducing MRSA and other MDRO frequencies over time is of preeminent importance[10] and probability modeling needs to be on the front lines of this effort. If MRSA frequency is reduced in a defined area (e.g., community, hospital, outpatient service), then the change would be reflected in the distribution model and related statistics (mean, variance, skew and Poisson probabilities). In situations where reduction in frequencies is subtle, the warning system cutoff points can help quantify the effectiveness of different MDRO reduction efforts. Because infection control is becoming a more interdisciplinary and multidisciplinary activity, particularly in the area of antibiotic resistance surveillance,[11] it is likely that probability models will become important infection control tools that allow, in the spirit of W.E. Deming,[12] the sound use of probability and statistics as a basis for action.

REFERENCES

1. Bonten MJM, Austin DJ, Lipsitch M. Understanding the spread of antibiotic resistant pathogens in hospitals: mathematical models as tools for control. Clin Infect Dis 2001;33:1739-46.
2. Austin DJ, Anderson RM. Studies of antibiotic resistance within the patient, hospitals and the community using simple mathematical models. Phil Trans R Soc Lond B 1999;354:721-38.
3. Perencevich EN, Hartley DM. Of models and methods: our analytic armamentarium applied to methicillin-resistant *Staphylococcus aureus*. Infect Control Hosp Epidemiol 2005;26:594-97.
4. Clinical and Laboratory Standards Institute. Analysis and presentation of cumulative antimicrobial susceptibility test data: approved guidelines. 2nd ed. CLSI document M39-A2. Wayne (PA): Clinical and Laboratory Standards Institute; 2005.
5. Shewart WA. Economic control of quality of manufactured product. New York: D. Van Nostrand Company; 1931.
6. Benneyan JC. Statistical quality control methods in infection control and hospital epidemiology, part I: introduction and basic theory. Infect Control Hosp Epidemiol 1998;19:194-214.
7. Benneyan JC. Statistical quality control methods in infection control and hospital epidemiology, part II: chart use, statistical properties, and research issues. Infect Control Hosp Epidemiol 1998;19:265-83.
8. Wagner MW, Tsui F, Espino JU, and others. The emerging science of very early detection of disease outbreaks. J Public Health Management Practice 2001;7:51-9.
9. Weinstein RA, Mallison GF. The role of the microbiology laboratory in surveillance and control of nosocomial infections. Am J Clin Pathol 1978;69:130-6.
10. Siegel JD, Rhinehart E, Jackson M, Chiarello L. Management of multidrug-resistant organisms in health care settings, 2006. Am J Infect Control 2007;35:S165-193.
11. Larson E, Saiman L, Haas J, and others. Perspectives on antimicrobial resistance: Establishing an interdisciplinary research approach. Am J Infect Control 2005;33:410-8.
12. Deming WE. On probability as a basis for action. Am Stat 1975;29:146-52.