Available online at www.sciencedirect.com

SCIENCE DIRECT®

Diagnostic Microbiology and Infectious Disease 53 (2005) 61–64

DIAGNOSTIC MICROBIOLOGY AND INFECTIOUS DISEASE

www.elsevier.com/locate/diagmicrobio

Antimicrobial susceptibility studies

The effect of interlaboratory variability on antimicrobial susceptibility determination

David H. Annis[a,*], Bruce A. Craig[b]

[a]*Operations Research Department, Naval Postgraduate School, Monterey, CA 93943-5219, USA*
[b]*Department of Statistics, Purdue University, West Lafayette, IN 47907-2067, USA*

Received 14 September 2004; accepted 25 March 2005

Abstract

In the minimum inhibitory concentration (MIC) test literature, discussion concerning the effect of laboratory-to-laboratory variation is lacking. We present 2 sets of drug dilution test quality control data that illustrate considerable laboratory differences in measured MIC. In both isolates (*Escherichia coli*, ATCC 25922; *Staphylococcus aureus*, ATCC 29213) the laboratory-to-laboratory variability accounts for approximately half of the total variability. We illustrate the impact of this variability on the probability of correctly classifying the susceptibility level of an isolate and on the estimation of resistance prevalence. For example, we show that laboratory differences in the probability of correctly classifying the isolate (specifically near the lower breakpoint) can vary up to 80%.

Keywords: Interlaboratory variability; Antimicrobial susceptibility determination; *Escherichia coli*; *Staphylococcus aureus*

1. Introduction

To determine the susceptibility of an unknown pathogen to a specific drug, a laboratory often performs a broth dilution test, such as a 2-fold minimum inhibitory concentration (MIC) test. For this type of test, the observed MIC is classified into one of 3 categories (susceptible, intermediate, or resistant) based on lower and upper MIC breakpoints. It is common for the intermediate range (distance between the breakpoints) to be one dilution.

Repeated laboratory experiments using the same pathogen/drug combination commonly show a 3-fold dilution range in the observed MIC. This measurement or experimental variability is most likely because of variations in inoculum size, incubation time, temperature, and other environmental factors. Given the relatively narrow intermediate range, this variability can seriously impact the ability to correctly classify a pathogen if its true MIC were near the breakpoints (Craig, 2000).

In this article, we address another important source of variability, the interlaboratory variability, and its effect on the ability to correctly classify a pathogen and estimate the prevalence of pathogen resistance. Using MIC quality control data and the statistical model presented by Craig (2000), we demonstrate the large impact this error can have on classification when the isolate's MIC is near the breakpoint.

2. Materials and methods

In the standard MIC broth dilution test, 2-fold dilutions are performed, with the observed MIC being the lowest 2-fold dilution without visible growth. For the remainder of this article, consider the observed MIC in terms of binary logarithmic (log 2) units such that the test results are integer values (e.g., 0.5 μg/mL = −1 and 1.0 μg/mL = 0 and 2.0 μg/mL = 1).

2.1. The model

We used data from quality control experiments involving repeated measurements of MIC both within and across laboratories. An upward-rounded, hierarchical normal model is used to describe the distribution of the observed MIC at each laboratory (Craig, 2000). Under this formulation, the ith laboratory has a MIC, μ_i, which represents the exact concentration of the drug required to inhibit the particular isolate. To account for experimental variation, the jth

* Corresponding author. Tel.: +1-831-656-2590; fax: +1-831-656-2595.
E-mail address: annis@nps.edu (D.H. Annis).

0732-8893/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.diagmicrobio.2005.03.012

measurable MIC at the ith laboratory, $x_{i,j}$, is assumed to be normally distributed with mean μ_i and variance σ_e^2.

To allow for interlaboratory variability, we assume that the laboratory-specific MICs follow a normal distribution with mean θ and variance σ_L^2. Thus, θ represents the true underlying MIC for the isolate, and the quantities σ_L^2 and σ_e^2 represent the magnitudes of variation attributable to interlaboratory differences and experimental error, respectively. This orthogonal decomposition of variance is a common approach to account for several sources of variability. With additional information, this interlaboratory variability could be further broken down to account for other sources of variability such as day-to-day variation. As an alternative approach, one could assume these errors follow a specific correlation structure, but for our purposes, the particular choice of model does not matter.

The upward-rounding inherent in this procedure implies that the ith observed value at the ith laboratory is $y_{i,j} = \lceil x_{i,j} \rceil = \lceil \mu_i + \epsilon_{i,j} \rceil = \lceil \theta + \delta_i + \epsilon_{i,j} \rceil$, where $\delta_i \sim N(0, \sigma_L^2)$ is the the ith laboratory effect and $\epsilon_{i,j} \sim N(0, \sigma_e^2)$ is the experimental error. The $\lceil \cdot \rceil$ notation represents the "ceiling" function, which returns the smallest integer greater than or equal to its argument (e.g., $\lceil 1.28 \rceil = 2$). Therefore, the probability that an observed MIC, $y_{i,j}$, at laboratory i will be integer k is the probability that the underlying measurable MIC, $x_{i,j}$, is contained in the interval $(k-1, k]$. This can be expressed as

$$Pr(y_{i,j} = k \mu_i) = \Phi\left(\frac{k - \mu_i}{\sigma_e}\right) - \Phi\left(\frac{k - 1 - \mu_i}{\sigma_e}\right),$$

where μ_i is the mean MIC at laboratory i and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

2.2. The data

Table 1A and B presents quality control data from the National Committee on Clinical Laboratory Standards Subcommittee on Antimicrobial Susceptibility Testing (S. Cullen, personal communication) involving 50 measurements of the observed MIC (for an unspecified drug or drugs) taken on the same strain of *Escherichia coli* and *Staphylococcus aureus*, respectively, at 10 different laboratories. These quality control studies typically include at least 7 laboratories with replicates (i.e., individually prepared inoculum suspensions) tested using at least 3 different media lots over a 3- to 4-day period.

The expectation maximization (EM) algorithm of Dempster et al. (1977) can be used to compute laboratory-specific (censored) maximum likelihood estimates of the mean assuming the previously described model. Wolynetz (1979) details the implementation of the expectation maximization procedure for a censored normal population. These laboratory-specific estimates differ by as much as 0.8 and 1.1 dilutions for the *E. coli* and *S. aureus* data, respectively. Differences of this magnitude are striking, because the 1.1-unit dilution range is wider than the intermediate classification range of 1 unit.

3. Results

Under our proposed hierarchy, $x_{i,j} \sim N(\theta, \sigma_L^2 + \sigma_e^2)$ and $\mathrm{Cov}(x_{i,j}, x_{i,k}) = \sigma_e^2$; $j \neq k$ after integrating out the unknown laboratory means μ_i. The observed MICs are interval-censored observations of the realized $x_{i,j}$'s. Maximum likelihood estimates of σ_L^2 and σ_e^2 can be obtained numerically. However, because the data are heavily censored, numerical maximum likelihood estimates can be unstable (especially when one variance component is near zero). A Bayesian approach using diffuse (essentially noninformative) prior distributions remedies this concern. (R code for either numerical maximum likelihood or for Bayesian computations is available from the authors. The Bayesian solution requires WinBUGS, which is available at http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml.)

For the *E. coli* data in Table 1A, the posterior estimates of variance components are $\hat{\sigma}_L^2 = 0.101$ and $\hat{\sigma}_e^2 = 0.136$, indicating that the interlaboratory variation accounts for 43% of total variation. For the *S. aureus* data in Table 1B, $\hat{\sigma}_L^2 = 0.241$ and $\hat{\sigma}_e^2 = 0.198$, which suggest that laboratory-to-laboratory differences account for approximately 55% of the total variability.

Permutation tests were conducted to determine whether laboratory differences contribute a statistically significant amount of variation. Under the null hypothesis that there are

Table 1
Quality control data for *E. coli* (ATCC 25922) and *S. aureus* (ATCC 29213) show large variation in reported MIC from laboratory to laboratory

A. *E. coli* (ATCC 25922)
Observed MIC

Laboratory	−8	−7	−6	−5	Mean
I	8	36	6	–	−7.54
II	6	41	3	–	−7.57
III	7	32	11	–	−7.42
IV	–	48	2	–	−7.13
V	2	48	–	–	−7.68
VI	–	33	17	–	−7.10
VII	7	41	2	–	−7.62
VIII	–	15	35	–	−6.88
IX	–	33	16	1	−7.12
X	1	35	14	–	−7.22
All	31	362	106	1	−7.34

B. *S. aureus* (ATCC 29213)
Observed MIC

Laboratory	−8	−7	−6	−5	−4	−3	Mean
I	–	14	34	2	–	–	−6.75
II	–	–	−24	26	–	–	−5.99
III	–	19	29	2	–	–	−6.84
IV	–	–	37	8	4	1	−6.11
V	–	2	45	3	–	–	−6.47
VI	–	3	33	14	–	–	−6.28
VII	–	12	36	2	–	–	−6.71
VIII	–	2	8	40	–	–	−5.76
IX	–	–	50	–	–	–	−6.50
X	1	19	27	3	–	–	−6.86
All	1	71	323	100	4	1	−6.42