

Forecast probabilities - modelling literature review

Nicolas Blöchliger, Institute of Medical Microbiology, University of Zurich

June 21, 2016

Contents

1 Error rate-bound method [1]	1
2 Craig [2]	2
3 Normalized resistance interpretation [3]	3
4 Craig on interlaboratory variability [4]	4
5 Turnidge [5]	5
6 Kassteele [6]	5
7 Meletiadiis [7]	6
8 Cantón [8]	7
9 Jaspers [9]	7

1 Error rate-bound method [1]

Aim Determine CBPs for inhibition zone diameters.

Required MIC cut-off, desired sensitivity and specificity.

Model Classification according to the MIC cut-off is treated as the ground truth. CBPs for inhibition zone diameters are set such that desired sensitivity and specificity are achieved for a minimal intermediate zone. Ref. [10] presents a cost function and an algorithm for its optimisation to apply the error rate-bound method in cases with two different MIC cut-offs.

Abstract Ref. [1] In considerations of susceptibility testing of bacteria, linear regression has been used to relate minimal inhibitory concentration and zone size. Although lacking in statistical validity, this technique has been used for aerobic bacteria. Reports of anaerobic testing show much greater variation about the regression line; the errors of misclassification have not been analyzed.

A bivariate “error rate-bounded” classification scheme is proposed for relating minimal inhibitory concentration and zone size of bacteria. This method requires the following clinical input: (1) specification of breakpoints of susceptibility as determined by minimal inhibitory concentrations and the pharmacology of the antibiotic, (2) determination of the relative importance of two types of errors, false resistant classification and false susceptible classification, and (3) determination of the acceptable rate of error of false classification. Based on this input the classification scheme determines the zone size, which gives the minimal error of misclassification with the known data on zone diameter and minimal inhibitory concentration.

Abstract Ref. [10] In antimicrobial susceptibility testing, minimum inhibitory concentration (MIC) susceptibility break points are defined by correlation of bacteriologic—clinical outcome data with MIC data for the infecting organisms. Disk diffusion [that is, zone-diameter (Z)] correlates are then established that provide for the prediction of organism susceptibility, while misclassification errors are kept to a minimum. The determination of Z break points through an error-rate-bounded classification scheme was first proposed by Metzler and DeHaan (1972). This method involves one MIC break point that separates susceptible and resistant strains. More recently, researchers have preferred to use two MIC break points (susceptible and resistant) that separate susceptible, moderately susceptible, and resistant strains. There is no known methodology for determining the Z break points for this latter situation, other than enumerating solutions for all feasible Z break-point pairs and choosing among the results. Our interest lay in presenting a methodology for determining the Z break points once the MIC break points are established. By deriving an index as a function of Z break points, a search method for finding the optimal Z break points is given. For the data set examined, our index interval solution required only a small percentage of solutions to be examined.

2 Craig [2]

Motivation It is difficult to assess the consistency of the error rate-bound method since repeat experiments have rarely been performed.

Model

- The distribution of the true MIC is a mixture of normal distributions.
- The true diameter is given by

$$d = \beta_0 \frac{e^{\beta_1 - \beta_2 m}}{1 + e^{\beta_1 - \beta_2 m}},$$

where $m = \log_2(MIC)$. A comment on p. 194 explains why this form is preferred to the usual linear or quadratic relationship.

- The observed MIC and diameter are given by $\lceil m + \varepsilon \rceil$ and $\lceil d + \delta \rceil$, where ε and δ are independent of m and d and normally distributed with zero mean.

Probability of correct classification Formulas are given for the probability that a classification is correct as a function of the *true* diameter, which is not a forecast probability.

Abstract In determining zone diameter breakpoints, the error-rate bounded method focuses directly on the observed discrepancy percentages (very major, major, and minor). These percentages, however, are quite variable due to the number of isolates investigated, the drug-specific relationship between MIC and zone diameter, the location of the isolates relative to the MIC intermediate zone, and the inherent variability of each test. To overcome potential sampling problems, a hierarchical model is proposed which explicitly accounts for each of these factors and probabilities from this model are used to determine diameter breakpoints. A simulation study is performed to demonstrate the improved consistency of this model-based procedure. Application to three published scatterplots demonstrate its interpretability advantages.

Follow-up study [11] Focus on bias introduced by rounding MIC values. Model (Gaussian error and mixture of normal distributions) remains the same.

Lamy [12] Focus on the impact of susceptibility prevalence. Almost the same model as in Ref. [2] is used. However, the conditional probability of the true diameter for a given observed diameter is erroneously modelled as Gaussian.

DePalma [13, 14] Software tool for determining breakpoints for diameters based on breakpoints for MIC. Uses Craig’s model, but also allows for different relationship between MIC and MGIT (still not physically motivated).

3 Normalized resistance interpretation [3]

Aim Determine the lower end of the susceptible population.

Model The diameters of the susceptible population are assumed to be normally distributed.

Fitting First, the histogram is smoothed in order to identify the position of the peak associated with the susceptible population. Data points with diameters greater than the position of this peak are then used to fit a normal distribution using linear regression of empirical and theoretical quantiles of a standard normal distribution. Note that the uppermost 2% of the data are omitted for this.

Abstract *Objective:* To evaluate a calibration method for disk diffusion antibiotic susceptibility tests, using zone diameter values generated in the individual laboratory as the internal calibrator for combinations of antibiotic and bacterial species.

Methods: The high-zone side of zone histogram distributions was first analyzed by moving averages to determine the peak position of the susceptible population. The accumulated percentages of isolates for the high zone diameter

values were calculated and converted into probit values. The normal distribution of the ideal population of susceptible strains was then determined by using the least-squares method for probit values against zone diameters, and the ideal population was thereby defined, including mean and standard deviation. Zone diameter values were obtained from laboratories at the Karolinska Hospital (KS) and Växjö Hospital (VX), and from two laboratories (LabA, LabB) in Argentina. The method relies on well standardized disk tests, but is independent of differences in MIC limits and zone breakpoints, and does not require the use of reference strains. Resistance was tentatively set at below 3 SD from the calculated, ideal mean zone diameter of the susceptible population.

Results: The method, called normalized interpretation of antimicrobial resistance, was tested on results from the KS and VX clinical microbiology laboratories, using the disk diffusion method for antimicrobial susceptibility tests, and for two bacterial species, *Staphylococcus aureus* and *Escherichia coli*. In total, 114 217 test results were included for the clinical isolates, and 3582 test results for control strains. The methodology at KS and VX followed the standard of the Swedish Reference Group for Antibiotics (SRGA). Zone diameter histograms for control strains were first analyzed to validate the procedure, and a comparison of actual means with the calculated means showed a correlation coefficient of $r = 0.998$. Results for clinical isolates at the two laboratories showed an excellent agreement for 54 of 57 combinations of antibiotic and bacterial species between normalized interpretations and the interpretations given by the laboratories. There were difficulties with *E. coli* and mecillinam, and *S. aureus* and tetracycline and rifampicin. The method was also tested on results from two laboratories using the NCCLS standard, and preliminary results showed very good agreement with quality-controlled laboratory interpretations.

Conclusions: The normalized resistance interpretation offers a new approach to comparative surveillance studies whereby the inhibition zone diameter results from disk tests in clinical laboratories can be used for calibration of the test.

4 Craig on interlaboratory variability [4]

Aim Quantify the interlaboratory variability of MIC measurements.

Model $\log_2(MIC_{i,j}) = \lceil \theta + \delta_i + \varepsilon_{i,j} \rceil$, where $\delta_i \sim N(0, \sigma_L^2)$ is the i -th laboratory effect and $\varepsilon_{i,j} \sim N(0, \sigma_e^2)$ is the experimental error.

Fitting Laboratory-specific maximum likelihood estimates of the mean are computed using the EM algorithm (taking censorship into account). "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."

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%.

5 Turnidge [5]

Aim Determine the extent of the susceptible population.

Model The MICs of the wild-type population are assumed to be log-normally distributed.

Fitting I do not understand their procedure.

Abstract MIC distribution data were obtained from a variety of international sources, and pooled after selection by a defined criterion. Sixty-seven of these datasets were subjected to a range of statistical goodness-of-fit tests. The log-normal distribution was selected for subsequent modelling. Cumulative counts of MIC distribution data were fitted to the cumulative log-normal distribution using non-linear least squares regression for a range of data subsets from each antibiotic–bacterium combination. Estimated parameters in the regression were the number of isolates in the subset, and (the log₂ values of) the mean and standard deviation. Optimum fits for the cumulative log-normal curve were then used to determine the wild-type MIC range, determined by calculating the MICs associated with the lower and upper 0.1% of the distribution, rounding to the nearest two-fold dilution, and calculating the probabilities of values higher and lower than these values. When plotted logarithmically, histograms of MIC frequencies appeared normal (Gaussian), but standard goodness-of-fit tests showed that the two-fold dilution grouping of MICs fits poorly to a log-normal distribution, whereas non-linear regression gave good fits to population (histogram) log-normal distributions of log₂ MIC frequencies, and even better fits to log-normal cumulative distributions. Optimum fits were found when the difference between the estimated and true number of isolates in the fitted subset was minimal. Sixteen antibiotic–bacterium datasets were fitted using this technique, and the log₂ values of the means and standard deviations were used to determine the 0.1% and 99.9% wild-type cut-off values. When rounded to the nearest two-fold dilution, $\geq 98.5\%$ of MIC values fall within the cut-off value range. Non-linear regression fitting to a cumulative log-normal distribution is a novel and effective method for modelling MIC distributions and quantifying wild-type MIC ranges.

6 Kasstele [6]

Aim Quantify the interlaboratory variability of MIC measurements.

Model For the i -th measurement: $\log_2(MIC) \sim [N(\mu_i, \sigma^2)]$, where $\mu_i = X_i\beta + b_{lab,j[i]}$, X_i is a binary vector encoding strain, antimicrobial agent, and agar medium (= fixed-effect terms), and $b_{lab,j[i]}$ is a lab-specific random effect.

Fitting Bayesian setting, WinBUGS.

Abstract Seventeen laboratories participated in a cooperative study to validate the regional susceptibility testing of *Neisseria gonorrhoeae* in The Netherlands. International reference strains were distributed. Each laboratory determined the MICs of ciprofloxacin, penicillin, and tetracycline, for each strain by Etest. To explore a more transparent assessment of quality and comparability, a statistical regression model was fitted to the data that accounted for the censoring of the MICs. The mean MICs found by all of the laboratories except three were closer than one 2-fold dilution step to the overall mean, and the mean MICs of each antimicrobial agent were close to the MICs for the international reference strains. This approach provided an efficient tool to analyze the performance of the Dutch decentralized gonococcal resistance monitoring system and confirmed good and comparable standards.

7 Meletiadis [7]

Aim Determine cut-off for MIC distribution.

Model The distribution of the MICs is assumed to be a normal mixture of two components with potentially different variances.

Fitting Regression to parts of Q-Q plot.

Abstract Epidemiological cutoff values (ECV) are commonly used to separate wild-type isolates from isolates with reduced susceptibility to antifungal drugs, thus setting the foundation for establishing clinical breakpoints for *Aspergillus fumigatus*. However, ECVs are usually determined by eye, a method which lacks objectivity, sensitivity, and statistical robustness and may be difficult, in particular, for extended and complex MIC distributions. We therefore describe and evaluate a statistical method of MIC distribution analysis for posaconazole, itraconazole, and voriconazole for 296 *A. fumigatus* isolates utilizing nonlinear regression analysis, the normal plot technique, and recursive partitioning analysis incorporating cyp51A sequence data. MICs were determined by using the CLSI M38–A2 protocol (CLSI, CLSI document M38–A2, 2008) after incubation of the isolates for 48 h and were transformed into log₂ MICs. We found a wide distribution of MICs of all azoles, some ranging from 0.02 to 128 mg/liter, with median MICs of 32 mg/liter for itraconazole, 4 mg/liter for voriconazole, and 0.5 mg/liter for posaconazole. Of the isolates, 65% (192 of 296) had mutations in the cyp51A gene, and the majority of the mutants (90%) harbored tandem repeats in the promoter region combined with mutations in the cyp51A coding region. MIC distributions deviated significantly from normal distribution (D’Agostino-Pearson omnibus normality test P value, <0.001), and they were better described with a model of the sum of two Gaussian distributions (R²,

0.91 to 0.96). The normal plot technique revealed a mixture of two populations of MICs separated by MICs of 1 mg/liter for itraconazole, 1 mg/liter for voriconazole, and 0.125 mg/liter for posaconazole. Recursive partitioning analysis confirmed these ECVs, since the proportions of isolates harboring *cyp51A* mutations associated with azole resistance were less than 20%, 20 to 30%, and >70% when the MICs were lower than, equal to, and higher than the above-mentioned ECVs, respectively.

8 Cantón [8]

Aim Determine cut-off for MIC distribution.

Model The smoothed distribution of the MICs is assumed to be a normal mixture of an arbitrary number of components with potentially different variances. It is not clear whether the model is multivariate.

Fitting Mclust().

Abstract The Sensititre YeastOne (SYO) method is a widely used method to determine the susceptibility of *Candida* spp. to antifungal agents. CLSI clinical breakpoints (CBP) have been reported for antifungals, but not using this method. In the absence of CBP, epidemiological cutoff values (ECVs) are useful to separate wild-type (WT) isolates (those without mechanisms of resistance) from non-WT isolates (those that can harbor some resistance mechanisms), which is the goal of any susceptibility test. The ECVs for five agents, obtained using the MIC distributions determined by the SYO test, were calculated and contrasted with those for three statistical methods and the MIC₅₀ and modal MIC, both plus 2-fold dilutions. The median ECVs (in mg/liter) (% of isolates inhibited by MICs equal to or less than the ECV; number of isolates tested) of the five methods for anidulafungin, micafungin, caspofungin, amphotericin B, and flucytosine, respectively, were as follows: 0.25 (98.5%; 656), 0.06 (95.1%; 659), 0.25 (98.7%; 747), 2 (100%; 923), and 1 (98.5%; 915) for *Candida albicans*; 8 (100%; 352), 4 (99.2%; 392), 2 (99.2%; 480), 1 (99.8%; 603), and 0.5 (97.9%; 635) for *C. parapsilosis*; 1 (99.2%; 123), 0.12 (99.2%; 121), 0.25 (99.2%; 138), 2 (100%; 171), and 0.5 (97.2%; 175) for *C. tropicalis*; 0.12 (96.6%; 174), 0.06 (96%; 176), 0.25 (98.4%; 188), 2 (100%; 209), and 0.25 (97.6%; 208) for *C. glabrata*; 0.25 (97%; 33), 0.5 (93.9%; 33), 1 (91.9%; 37), 4 (100%; 51), and 32 (100%; 53) for *C. krusei*; and 4 (100%; 33), 2 (100%; 33), 2 (100%; 54), 1 (100%; 90), and 0.25 (93.4%; 91) for *C. orthopsilosis*. The three statistical methods gave similar ECVs (within one dilution) and included $\geq 95\%$ of isolates. These tentative ECVs would be useful for monitoring the emergence of isolates with reduced susceptibility by use of the SYO method.

9 Jaspers [9]

Aim Determine the extent of the susceptible population.

Model A two-component mixture model is assumed, where the first component corresponds to the wild-type and follows a log-normal or gamma distribution. A clear separation between the two components is required.

Fitting Maximum-likelihood approach. The non-wild-type component is treated non-parametrically.

Abstract Antimicrobial resistance has become one of the main public health burdens of the last decades, and monitoring the development and spread of non-wild-type isolates has therefore gained increased interest. Monitoring is performed based on the minimum inhibitory concentration (MIC) values, which are collected through the application of dilution experiments. In order to account for the unobserved population heterogeneity of wild-type and non-wild-type isolates, mixture models are extremely useful. Instead of estimating the entire mixture globally, it was our major aim to provide an estimate for the wild-type first component only. The characteristics of this first component are not expected to change over time, once the wild-type population has been confidently identified for a given antimicrobial. With this purpose, we developed a new method based on the multinomial distribution, and we carry out a simulation study to study the properties of the new estimator. Because the new approach fits within the likelihood framework, we can compare distinct distributional assumptions in order to determine the most suitable distribution for the wild-type population. We determine the optimal parameters based on the AIC criterion, and attention is also paid to the model-averaged approach using the Akaike weights. The latter is thought to be very suitable to derive specific characteristics of the wild-type distribution and to determine limits for the wild-type MIC range. In this way, the new method provides an elegant means to compare distinct distributional assumptions and to quantify the wild-type MIC distribution of specific antibiotic–bacterium combinations.

First follow-up study [15] Penalized mixture approach: The non-wild type population is modelled as a mixture of equidistant Gaussian densities with data-independent variances. I do not see a conceptual difference to a kernel density estimate. Fitting involves a penalty term. Bootstrapping is used to quantify the uncertainty of the model. Model based classification (et vs non-wt) is suggested as a possibility but not performed.

Second follow-up study [16] Same model as in Ref. [15]. Improved algorithm for fitting. Plots of $\text{prob}(\text{wt}|\text{MIC})$ with confidence intervals are shown.

References

- [1] C. M. Metzler and R. M. DeHaan. Susceptibility tests of anaerobic bacteria: statistical and clinical considerations. *J. Infect. Dis.*, 130(6):588–594, 1974.
- [2] B. A. Craig. Modeling approach to diameter breakpoint determination. *Diagn. Microbiol. Infect. Dis.*, 36(3):193–202, 2000.

- [3] G. Kronvall, G. Kahlmeter, E. Myhre, and M. F. Galas. A new method for normalized interpretation of antimicrobial resistance from disk test results for comparative purposes. *Clin. Microbiol. Infect.*, 9(2):120–132, 2003.
- [4] D. H. Annis and B. A. Craig. The effect of interlaboratory variability on antimicrobial susceptibility determination. *Diagn. Microbiol. Infect. Dis.*, 53(1):61–64, 2005.
- [5] J. Turnidge, G. Kahlmeter, and G. Kronvall. Statistical characterisation of bacterial wild-type MIC value distributions and the determination of epidemiological cut-off values. *Clin. Microbiol. Infect.*, 12(5):418–425, 2006.
- [6] J. van de Kastele, M. G. van Santen-Verheuve, F. D. H. Koedijk, A. P. van Dam, M. A. B. van der Sande, and A. J. de Neeling. New statistical technique for analyzing MIC-based susceptibility data. *Antimicrob. Agents Chemother.*, 56(3):1557–1563, 2012.
- [7] J. Meletiadis, E. Mavridou, W. J. G. Melchers, J. W. Mouton, and P. E. Verweij. Epidemiological cutoff values for azoles and *Aspergillus fumigatus* based on a novel mathematical approach incorporating cyp51A sequence analysis. *Antimicrob. Agents Chemother.*, 56(5):2524–2529, 2012.
- [8] E. Cantón, J. Pemán, D. Hervás, C. Iñiguez, D. Navarro, J. Echeverría, J. Martínez-Alarcón, D. Fontanals, B. Gomila-Sard, B. Buendía, L. Torroba, J. Ayats, A. Bratos, F. Sánchez-Reus, Fernández-Natal, and the FUNGEMYCA Study Group. Comparison of three statistical methods for establishing tentative wild-type population and epidemiological cutoff values for echinocandins, amphotericin B, flucytosine, and six *Candida* species as determined by the colorimetric Sensititre YeastOne method. *J. Clin. Microbiol.*, 50(12):3921–3926, 2012.
- [9] S. Jaspers, M. Aerts, G. Verbeke, and P.-A. Beloeil. Estimation of the wild-type minimum inhibitory concentration value distribution. *Statist. Med.*, 33(2):289–303, 2014.
- [10] M. N. Brunden, G. E. Zurenko, and B. Kapik. Modification of the error-rate bounded classification scheme for use with two MIC break points. *Diagn. Microbiol. Infect. Dis.*, 15(2):135–140, 1992.
- [11] D. H. Annis and B. A. Craig. Statistical properties and inference of the antimicrobial MIC test. *Statist. Med.*, 24(23):3631–3644, 2005.
- [12] B. Lamy, G. Carret, J.-P. Flandrois, and M. L. Delignette-Muller. How does susceptibility prevalence impact on the performance of disk diffusion susceptibility testing? *Diagn. Microbiol. Infect. Dis.*, 49(2):131–139, 2004.
- [13] G. DePalma, J. Turnidge, and B. A. Craig. Determination of disk diffusion susceptibility testing interpretive criteria using model-based analysis: development and implementation. *Diagn. Microbiol. Infect. Dis.*, in press(?):??–??, 2016.
- [14] G. DePalma, J. Turnidge, and B. A. Craig. dBETS: An online software package for diffusion test breakpoint determination. *user manual*, in press(?):??–??, 2016.

- [15] S. Jaspers, M. Aerts, G. Verbeke, and P.-A. Beloeil. A new semi-parametric mixture model for interval censored data, with applications in the field of antimicrobial resistance. *Comput. Stat. Data Anal.*, 71:30–42, 2014.
- [16] S. Jaspers, G. Verbeke, D. Böhning, and M. Aerts. Application of the vertex exchange method to estimate a semi-parametric mixture model for the MIC density of *Escherichia coli* isolates tested for susceptibility against ampicillin. *Biostatistics*, 17(1):94–107, 2016.