

NOTE

Precision limits for MIC data required to set epidemiological cut-off values

P. Smith^{1a}

¹ Department of Microbiology, School of Natural science, National University of Ireland, Galway, Ireland

Keywords: MIC data, Precision limits, epidemiological cut-off values, antimicrobial susceptibility data

<https://doi.org/10.48045/001c.38713>

Bulletin of the European Association of Fish Pathologists

Vol. 42, Issue 2, 2022

If the standard deviation of the distribution of MIC values for a set of wild-type isolates exceeds $1.18 \log_2 \mu\text{g/mL}$ when calculated by NRI analysis or $1.11 \log_2 \mu\text{g/mL}$ when calculated by ECOFFinder that data set should be considered imprecise and not be used to set an epidemiological cut-off value.

In the Aquatic Animal Health Code (<http://www.oie.int/en/international-standard-setting/aquatic-code/access-online/>) OIE has argued that, in order for the antimicrobial susceptibility data for bacteria isolated from aquatic animals generated in monitoring and surveillance programmes to be commensurate, two essential conditions must be met. The data must be obtained by the use of standard testing methods and their meaning must be interpreted by the application of consensus, internationally harmonised, epidemiological cut-off values.

With respect to the availability of standard methods, those produced by CLSI (2020a) are suitable for susceptibility testing of 36 of the 44 species that are most commonly isolated from aquatic animals (Smith 2019). However, when the susceptibility tests are performed at incubation temperatures $\leq 28^\circ\text{C}$, the latest edition of the CLSI guideline VET04 (2020b) gives epidemiological cut-off values (ECVs) for only four of these species. The generation of the data that would allow the setting of ECVs for more species is an urgent priority.

The CLSI guideline M23 (CLSI 2018) provides some guidance as to the quantity of data they require for setting ECVs for minimum inhibitory concentration (MIC) data. This guideline states that ECVs can only be determined from aggregations of MIC data sets sourced from at least three separate laboratories which include observations from greater than 100 isolates. However, it makes no recommendation with respect to the quality of the data required. Smith (2020) has suggested that the precision of susceptibility data is an important, if frequently ignored, measure of the quality of any MIC data set.

^a Corresponding author:
peter.smith@nuigalway.ie

Epidemiological cut-off values categorise isolates as wild-type (WT) that are fully susceptible members of their species or non-wild-type (NWT) which show reduced susceptibility and are assumed to possess some resistance mechanism (Silleby 2012). Two methods for calculating epidemiological cut-off values for MIC data, ECOFFinder (https://www.eucast.org/mic_distributions_and_ecoffs/) and NRI (<http://www.bioscand.se/nri/>) have been developed. Both are available as free downloads. Both methods calculate a mean and a standard deviation (*sd*) for the distributions they generate from the \log_2 transformed observations from putative WT isolates. It should be noted that, in calculating their cut-off values, these methods use different statistical approaches to characterising the WT distributions. ECOFFinder uses a curve fitting approach whereas NRI calculates a normalised distribution of the WT observations. As a result, the *sd* values they produce cannot be directly compared with each other.

Smith et al. (2018) have argued that the *sd* values generated by these methods can serve as measures of the precision of the data set being examined. They analysed the *sd* values calculated by NRI analysis for 59 MIC data sets generated at 35°C, 27 data sets generated at 28°C, 28 data sets generated at 22°C and 37 data sets generated at 18°C (Table 1). They demonstrated that the distributions of *sd* values were independent of the temperatures at which the tests were performed and calculated the mean of the 151 *sd* values as 0.76 $\log_2 \mu\text{g/mL}$ with a standard deviation of 0.22 $\log_2 \mu\text{g/mL}$. Analysis of these data sets by ECOFFinder gave a mean of 0.62 $\log_2 \mu\text{g/mL}$ with a standard deviation of 0.26 $\log_2 \mu\text{g/mL}$ (Smith 2019). Smith et al. (2018) suggested that an upper limit for acceptable *sd* values could be set by calculating the mean plus two standard deviations of the *sd* calculated for these 151 published MIC data sets. They suggested that data sets for which either method generated a *sd* value in excess of these limits should be considered as excessively imprecise and should not be used to calculate epidemiological cut off values. Applied to the 151 data sets this approach gave an upper limit for data analysed by NRI and ECOFFinder of 1.19 $\log_2 \mu\text{g/mL}$ and 1.14 $\log_2 \mu\text{g/mL}$, respectively. However, these limits were calculated from data sets that had been generated in single laboratories and the validity of their application to the multi-laboratory aggregations required by CLSI (2018) for the setting of ECVs has not been established.

In order to generate the empirical evidence needed to set acceptable limits for the precision of multi-laboratory aggregations, NRI and ECOFFinder analyses were applied to 51 aggregated data sets that had been used by EUCAST to set epidemiological cut-off values. These data sets, that had been generated at 35°C, were selected from the EUCAST website (http://www.eucast.org/mic_distributions_and_ecoffs/) to meet three quantitative criteria. One group was selected to include data sets aggregated from <10 laboratories that all contained < 400 WT observations. This group contained 13 data sets for six different bacterial species and seven different agents. The criteria for this

Table 1. Means, standard deviations (SD) and acceptable upper limits ($\log_2 \mu\text{g/mL}$) for the *sd* values calculated by NRI and ECOFFinder analysis of single source and aggregated multi-source MIC data sets.

Sources ^a	WT (median) ^b	Data sets	Temp °C	NRI		ECOFFinder	
				Mean (SD)	Upper limit ^c	Mean (SD)	Upper limit ^c
Aggregated data sets from multiple sources							
3 - 5	192	13	35	0.77 (0.17)	1.11	0.63 (0.18)	0.98
3 - 8	2774	17	35	0.80 (0.14)	1.09	0.66 (0.18)	1.03
16 - 82	14774	21	35	0.80 (0.14)	1.08	0.72 (0.14)	0.88
Sub total		51	35	0.79 (0.14)	1.11	0.68 (0.16)	1.00
Data sets from single sources ^d							
1		37	18	0.76 (0.19)	1.14	0.58 (0.25)	1.08
1		28	22	0.77 (0.20)	1.16	0.58 (0.21)	1.00
1		27	28	0.76 (0.26)	1.28	0.62 (0.25)	1.12
1		59	35	0.76 (0.20)	1.20	0.65 (0.27)	1.19
Sub total		151	18-35	0.76 (0.22)	1.19	0.62 (0.26)	1.13
All data							
Total		201	18-35	0.77(0.20)	1.18	0.64 (0.23)	1.11

^a The number of independent laboratories contributing MIC observations to the total data set.

^b The median of numbers of MIC observations from isolates categorised as wild-type in the data sets.

^c Upper limit values were calculated as the mean + 2*SD of the data sets.

^d The analyses of the data sets from single sources were previously reported by Smith et al. (2018) and Smith (2019).

group were chosen as they would capture the majority of multi-laboratory aggregation so far generated for bacteria isolated from aquatic animals. A second group was selected to include data sets aggregated from <10 laboratories that contained >400 WT observations. This group contained 17 data sets for five different bacterial species and ten different agents. The third group was selected to include data sets aggregated for >15 laboratories. This group contained 21 data sets for six different bacterial species and 14 different agents. The mean of the standard deviations (*sd*) values of the normalised distributions generated by NRI and for the best-fit curves generated by ECOFFinder were calculated for each group separately and for the total 51 data sets taken together (Table 1). Statistical analysis of the distributions of *sd* values calculated by NRI and ECOFFinder were performed using InStat 3.1a (GraphPad Software). For both methods each separate group and the total set all passed the Kolmogorov-Smirnov normality test ($p < 0.05$). Analysis of *sd* obtained by the NRI for the three groups using the student t-test detected no significant differences between them ($p < 0.05$). Equally there were no significant differences between the *sd* values for these groups generated by ECOFFinder analysis ($p < 0.05$). Thus, these *sd* values are not significantly influenced by the number of observations in the data sets analysed or by the number of laboratories generating those observations. It was, therefore, considered legitimate to use the total data set to calculate the parameters of the *sd* distributions from multi-laboratory aggregations. The mean of the 51 *sd* values of the normalised distributions calculated by NRI was $0.79 \log_2 \mu\text{g/mL}$ with a standard deviation of $0.14 \log_2 \mu\text{g/mL}$. The mean of the best-fit curves calculated by ECOFFinder was $0.68 \log_2 \mu\text{g/mL}$ with a standard deviation of

0.16 $\log_2 \mu\text{g/mL}$. The suggested upper limits for acceptable *sd* values obtained by analysis for multi-laboratory aggregations were calculated as the mean plus two standard deviations of these 51 values. For NRI and ECOFFinder these upper limits were 1.11 $\log_2 \mu\text{g/mL}$ and 1.00 $\log_2 \mu\text{g/mL}$, respectively.

The limit values calculated for the aggregated data sets were similar to but slightly lower than the limits previously calculated for data sets obtained from single laboratories of 1.19 $\log_2 \mu\text{g/mL}$ and 1.14 $\log_2 \mu\text{g/mL}$ (Table 1). However, the student t-test demonstrated that there were no significant differences between the distributions of the *sd* values calculated by either method for the data sets from single laboratories and those aggregated from multiple laboratories ($p < 0.05$). This suggests that in addition to being independent of the temperatures at which the MIC tests were performed (Smith et al. 2018) these *sd* distributions can also be treated as independent of the number of observations they contain or the number of laboratories contributing data to them. Thus, it was considered legitimate to calculate acceptable limit values from an analysis of all 201 available data sets. When these data sets were analysed by NRI the mean of the *sd* values was 0.77 $\log_2 \mu\text{g/mL}$ with a standard deviation of 0.20 $\log_2 \mu\text{g/mL}$ and the limit value was 1.18 $\log_2 \mu\text{g/mL}$. When ECOFFinder was used the mean of the *sd* values was 0.64 $\log_2 \mu\text{g/mL}$ with a standard deviation of 0.23 $\log_2 \mu\text{g/mL}$ and the limit value was 1.11 $\log_2 \mu\text{g/mL}$.

In conclusion, these analyses suggest that it is possible to set acceptable limits for the precision of MIC data sets that would be applicable to all data sets irrespective of the number of observations in them, the number of independent sources contributing to them, or the temperatures at which they were obtained. It is argued that any data set for which NRI calculates an *sd* value $> 1.18 \log_2 \mu\text{g/mL}$ or ECOFFinder a value $> 1.11 \log_2 \mu\text{g/mL}$ should be considered as excessively imprecise. Consequently, epidemiological cut-off values calculated from aggregated data sets that exceed either one or both of these limits should be considered as unreliable.

When the inclusion of the data from a single laboratory in a multi-laboratory aggregation results in the excessive *sd* value calculated for that aggregation, the censoring of the aggregation by the exclusion of the aberrant data set may be considered legitimate.

Submitted: April 05, 2022 CEST, Accepted: May 17, 2022 CEST



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

REFERENCES

- CLSI. 2018. *Development of In Vitro Susceptibility Testing Criteria and Quality Control Parameters CLSI Guideline M23*. 5th ed. Wayne, Pennsylvania: Clinical Laboratory Standards Institute.
- . 2020a. “Methods for Antimicrobial Broth Dilution and Disk Diffusion Susceptibility Testing of Bacteria Isolated from Aquatic Animals.” CLSI guideline VET03 (2nd ed.). Wayne, Pennsylvania: Clinical Laboratory Standards Institute.
- . 2020b. “Performance Standards for Antimicrobial Susceptibility Testing of Bacteria Isolated from Aquatic Animals.” CLSI guideline VET04 (3rd ed.). Wayne, Pennsylvania: Clinical Laboratory Standards Institute.
- Silley, P. 2012. “Susceptibility Testing Methods, Resistance and Breakpoints: What Do These Terms Really Mean?” *Scientific and Technical Review* 31: 33–41. <https://doi.org/10.20506/rst.31.1.2097>.
- Smith, Peter. 2019. “The Performance of Antimicrobial Susceptibility Testing Programmes Relevant to Aquaculture and Aquaculture Products.” FAO Fisheries and Aquaculture Circular No. 1191. Rome: Food and Agriculture Organization. <http://www.fao.org/3/ca6028en/ca6028en.pdf>.
- . 2020. “Eight Rules for Improving the Quality of Papers on the Antimicrobial Susceptibility of Bacteria Isolated from Aquatic Animals.” *Diseases of Aquatic Organisms* 139: 87–92. <https://doi.org/10.3354/dao03476>.
- Smith, Peter, William Finnegan, Thao Ngo, and Göran Kronvall. 2018. “Influence of Incubation Temperature and Time on the Precision of MIC and Disc Diffusion Antimicrobial Susceptibility Test Data.” *Aquaculture* 490: 19–24. <https://doi.org/10.1016/j.aquaculture.2018.02.020>.