

# Estimation of the wild-type minimum inhibitory concentration value distribution

Stijn Jaspers,<sup>a,\*†</sup> Marc Aerts,<sup>a</sup> Geert Verbeke<sup>b</sup> and Pierre-Alexandre Beloeil<sup>c</sup>

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. Copyright © 2013 John Wiley & Sons, Ltd.

**Keywords:** antimicrobial resistance; censoring; model averaging; multinomial distribution; wild-type distribution

## 1. Introduction

Antimicrobial resistance (AMR) is the main undesirable side effect of antimicrobial use in both humans and animals. Because of the continuous positive selection of resistant bacterial clones, whether pathogenic, commensal, or even environmental bacteria, the population structure of microbial communities is modified. AMR has become one of the main public health burdens of the last decades, and it is therefore extremely important to study and monitor the emergence of isolates with reduced susceptibility against antimicrobials [1, 2]. This may be performed by determining the minimum inhibitory concentration (MIC), defined as the smallest concentration of a given antimicrobial substance that inhibits the visible growth of a microorganism. The MIC is commonly measured via a broth dilution method, in which a standardised amount of the isolate is exposed to successive two-fold concentrations of the antimicrobial (i.e. 0.25, 0.5, 1, 2 mg/L, ...). The MIC is defined as the lowest concentration with no visible growth after a prescribed incubation period. Consider, for example, a bacterial isolate that is subjected to an antimicrobial at concentrations 0.5, 1, 2, and 4 mg/L. In case the isolate shows inhibition of growth at values of 2 and 4 mg/L, but growth at lower values, the reported MIC value is equal to

<sup>a</sup>Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium

<sup>b</sup>Interuniversity Institute for Biostatistics and statistical Bioinformatics, KU Leuven, Leuven, Belgium

<sup>c</sup>European Food Safety Authority (EFSA), Parma, Italy

\*Correspondence to: Stijn Jaspers, Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Agoralaan 1, 3590 Diepenbeek, Belgium.

†E-mail: stijn.jaspers@uhasselt.be

2 mg/L. However, the true inhibition occurs between the concentrations of 1 and 2 mg/L, so the obtained MIC value is interval censored. This additional data complexity needs to be taken into account when estimating the wild-type density of interest.

The common way of representing a dilution experiment is by drawing an MIC distribution, that is, the frequency of occurrence of each given MIC plotted against the MIC value. For a given bacterial species, the multi-modal pattern of the MIC distribution can usually enable the separation of the wild-type population of microorganisms (typically on the left of the MIC distribution) from the non-wild-type populations that show a reduced susceptibility to the antimicrobial in question (on the right of the MIC distribution).

Typically, MIC values are dichotomised using epidemiological cut-off values (ECOFFs) to define non-wild-type isolates [3, 4]. As such, the MIC distribution is collapsed into a binary variable, and a large amount of information is lost. One of the main shortcomings of this approach is that the characteristics of both wild-type and non-wild-type isolates cannot be studied in detail. In addition, the broth dilution method implicitly rounds data upwards, and inference based on unadjusted measurements is biased and overestimates bacterial resistance to a drug. Annis and Craig [5] provide a distribution free estimate, which is still suffering from the shortcoming that a trend above the ECOFF cannot be detected. One can remedy this disadvantage via the exploration of the MIC distribution on the full continuous scale, in which interest emerged during the process of harmonisation of breakpoints [4]. Craig [6] proposed an appealing maximum likelihood estimate based on a censored normal mixture.

Mixture models are ideally suited as they offer a natural framework for modelling unobserved population heterogeneity. In our context, a two-component mixture

$$f(x) = \pi f_1(x|\theta_1) + (1 - \pi) f_2(x|\theta_2) \quad (1)$$

is assumed, in which  $f_1$  and  $f_2$  respectively represent the wild-type and non-wild-type components of the MIC distribution and the prevalence of wild-type isolates is denoted by  $\pi$ . The wild-type susceptible population, typically located on the left of the MIC distribution, is assumed to have no acquired or mutational resistance. It commonly shows a unimodal distribution reflecting a slight biological variability around a mode, which is not altered by changing circumstances over time. Therefore, we can assume the first component in (1) to be of a fixed parametric form, such as the log-normal or gamma distribution [7, 8]. The second component, representing the non-wild-type isolates, is often multi-modal, suggesting that it is itself a mixture of different non-wild-type subpopulations that are characterised by different degrees of reduced susceptibility conferred by different mechanisms. To allow proper modelling of the different possible characteristics of the non-wild-type distribution, we can model the second component by using another mixture of  $m$  components:

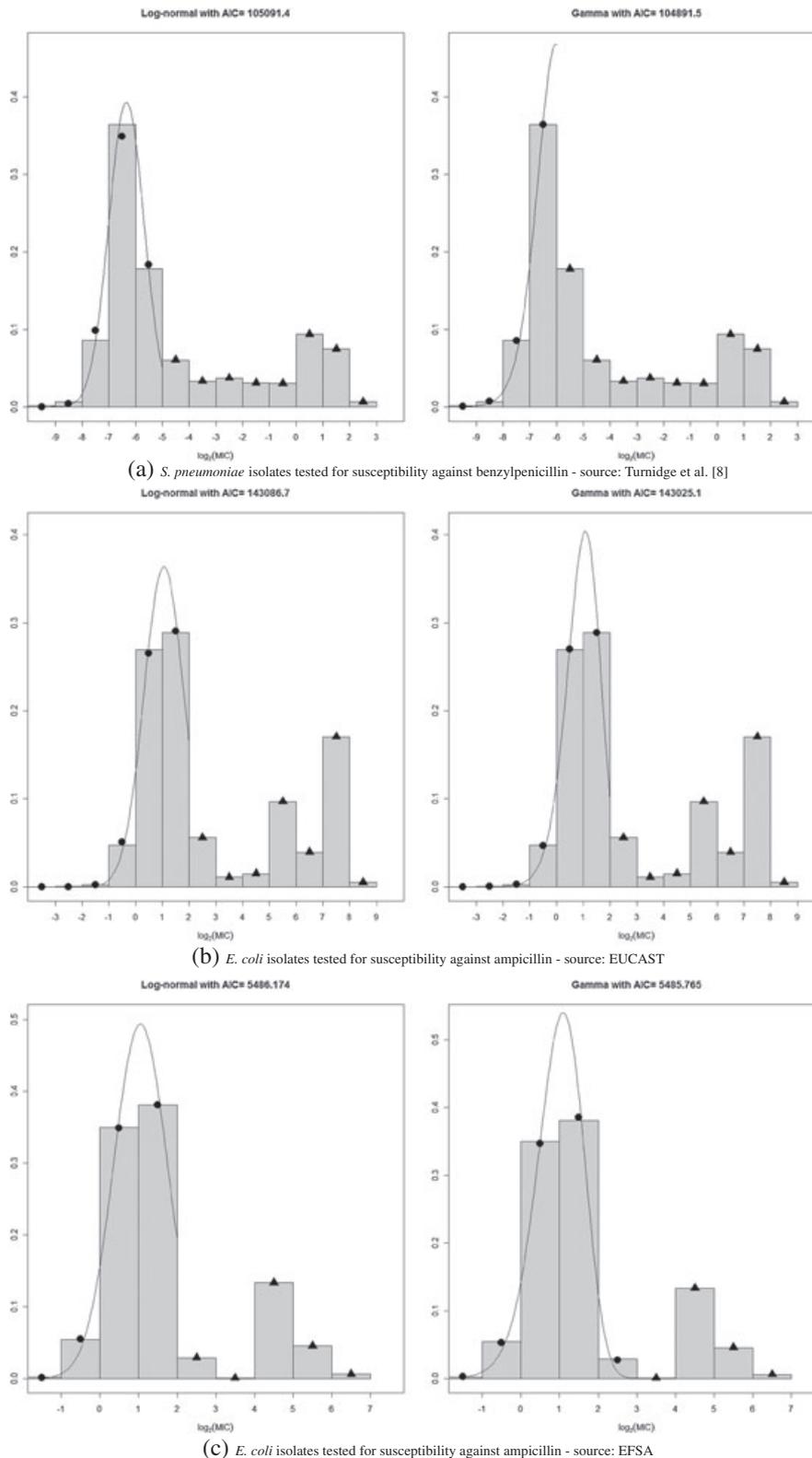
$$f_2(x|\theta_2) = \sum_{\ell=1}^m \tau_{\ell} f_{\ell}(x|\theta_{2\ell}).$$

At this stage, we can follow several approaches. Adopting a global view, we can estimate the mixture in (1) as a whole using a non-parametric maximum likelihood approach in combination with the Vertex Exchange Method [9] or through the application of a penalised mixture approach [10]. However, in this paper, we will apply a local view and focus on the wild-type component only. In a similar fashion, Turnidge *et al.* [8] provided a mathematical description of the wild-type MIC distribution through the application of non-linear least squares regression. We propose an alternative method that is based on the multinomial distribution in combination with likelihood inference, thereby allowing for a direct comparison between different parametric assumptions for the wild-type distribution.

In Section 2, we present three sources of data to which the method developed in Section 3 is applied. We carry out the actual data analysis in Section 4 and compare the performance of the new method with the method of Turnidge *et al.* [8] through a simulation study in Section 5. A discussion will end the paper in Section 6.

## 2. Data

Minimum inhibitory concentration is the lowest concentration of an antimicrobial agent that inhibits the visible growth of a microorganism, and these values are commonly obtained from dilution experiments. In this paper, we will study two antibiotic–bacterium combinations in more detail: *Streptococcus pneumoniae* versus benzylpenicillin and *Escherichia coli* versus ampicillin.



**Figure 1.** Barplot of the distribution of minimum inhibitory concentration (MIC) for the three data examples. The full line indicates the estimated density according to the multinomial-based method. The triangles represent the multinomial probabilities, while the dots represent the parametric counterparts in (3). (a) *S. pneumoniae* isolates tested for susceptibility against benzylpenicillin – source: Turnidge *et al.* [8], (b) *E. coli* isolates tested for susceptibility against ampicillin – source: EUCAST, and (c) *E. coli* isolates tested for susceptibility against ampicillin – source: EFSA.

## 2.1. *Streptococcus pneumoniae* and benzylpenicillin

*Streptococcus pneumoniae* is a Gram-positive bacterium that can cause pneumonia, paranasal sinusitis and otitis media. The organism is currently the leading cause of invasive bacterial disease in children and the elderly. Benzylpenicillin belongs to the group of  $\beta$ -lactams, which are by far the most widely used and efficacious of all antibiotics. Over the past few decades, however, widespread resistance has evolved among most common pathogens, including *S. pneumoniae* (e.g. [11, 12]). One of the considered datasets in Turnidge *et al.* [8] concerns *S. pneumoniae* tested for susceptibility against benzylpenicillin. The MIC distribution data were collected from a variety of international sources and pooled using a well-defined selection criterion. The resulting pool consisted of 17 independently generated datasets (for more information, see [8]). We present a graphical representation of the corresponding MIC value distribution in Figure 1(a).

In total, 26 403 observations were collected, with MIC values ranging from 0.002 to 8 mg/L. A first mode is located around the value of 0.016 mg/L, the concentration that probably represents the modal value of the wild-type component. The authors argued that this is an example where the non-wild-type isolates seem to merge with the susceptible wild-type strain. Therefore, it is expected that there is a lot of contamination by the non-wild-type isolates distribution.

## 2.2. *Escherichia coli* and ampicillin

*Escherichia coli* are Gram-negative bacteria that commonly live in the intestines of humans and animals worldwide. The bacteria can cause severe abdominal cramps and are a leading cause of bloody diarrhoea. Ampicillin is a semi-synthetic penicillin with an additional amino chain synthesised onto the penicillin molecule. This allows the ampicillin to be effective against Gram-negative organisms as well as the Gram-positive organisms covered by penicillin. Although the treatment of *E. coli* infections with ampicillin was initially effective, it is no longer the preferred drug because of changes in susceptibility of certain *E. coli* subpopulations. The susceptibility of *E. coli* against ampicillin will constitute the second dataset of interest. Data are obtained from two European institutions concerned with the gathering and analysis of AMR data.

The European Committee on Antimicrobial Susceptibility Testing (EUCAST) is an organisation that deals with breakpoints and technical aspects of phenotypic *in vitro* antimicrobial susceptibility testing. Most antimicrobial MIC breakpoints (e.g. ECOFFs) in Europe have been harmonised by EUCAST. On their website, the committee provides MIC distributions for a wide range of organisms and antimicrobial agents. These distributions are based on collated data from a total of almost 20 000 MIC distributions from worldwide sources. Among them, data are present on the desired combination: *E. coli* and ampicillin. The resulting MIC distribution consists of 39 220 isolates that were obtained from 48 distinct sources. The observed MIC values ranged from 0.125 to 512 mg/L, with the first mode being located around the value of 2 mg/L. We give a graphical representation of the data by the barplot in Figure 1(b). Two large peaks are clearly visible at the values of 2 and 4, probably representing the center of the wild-type component. Towards the larger MIC values, two smaller peaks are located at the values of 64 and 256, which could represent distinct strains of the non-wild-type isolates.

The European Food Safety Authority (EFSA) is the keystone of European Union (EU) risk assessment regarding food and feed safety. One of its main tasks is the monitoring of AMR in zoonotic and indicator bacteria from food-producing animals and food in the EU. EFSA coordinates the annual reporting of AMR data from the member states, analyses the data collected and issues the results of this analysis. EFSA and European Centre for Disease Prevention and Control jointly prepare and publish an annual EU summary report on AMR in zoonotic and indicator bacteria from humans, animals and food on a yearly basis. For the purpose of this modelling study, EFSA has provided an exemplary MIC distribution summarising the results of ampicillin susceptibility testing of indicator *E. coli* isolates. Figure 1(c) presents a graphical representation of the MIC values of the 1890 isolates. The mode of the wild-type component is again located between the values of 2 and 4 mg/L. However, there is presumably a unique non-wild-type population, for which the modal MIC is located at 32 mg/L.

## 3. Methodology

### 3.1. Method of Turnidge *et al.* [8]

Turnidge *et al.* [8] developed a method for characterising the wild-type MIC distribution, from which one can derive the ECOFF. The proposed method is based on the assumption that the wild-type component

of the MIC distribution follows a log-normal curve. They perform a non-linear least squares regression on the cumulative counts for a range of data subsets. Starting with the subset that includes values that are either at the first mode or at one unit higher, they fit the cumulative counts to the cumulative log-normal curve. They estimate three parameters, namely the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of the employed distribution function and the total number of observations ( $N$ ) in the presumed subset. They estimate  $N$  rather than take  $N$  as a constant in the regression because of the desire to fit the data to the distribution without assuming that  $N$  only contains wild-type isolates. Hence, the nonlinear function to be minimised is

$$h(N, \mu, \sigma) = \sum_i \left[ y_i - N * \Phi \left( \frac{x_i - \mu}{\sigma} \right) \right]^2, \quad (2)$$

where  $y_i$  represents the cumulative number of isolates up to MIC category  $x_i$  (on the  $\log_2$  scale). After having obtained model estimates for the first subset of the data, they repeat the procedure with an augmented subset of the data (i.e. they add MIC values that are one dilution higher to the previous subset). The fit was optimum when the absolute value of the ratio of the estimate of the parameter ( $\mu$  or  $\sigma$ ) to the asymptotic standard error was maximal. While optimum fits for each of the estimated parameters could occur with different subsets, most reasonable values for the parameters occurred when the difference between the observed and the estimated numbers of isolates in the fitted subset was minimal. Hence, the authors considered this criterion to be most suitable. One can consecutively apply the selected mean and standard deviation from the optimum fit to derive the ECOFF.

### 3.2. The multinomial-based method

Despite the fact that the method discussed in the previous subsection seems to perform well in practice (see [8]), there remain some issues about the underlying assumptions. First, the strongest assumption is about the distributional form of the wild-type component. The authors suggest the use of the log-normal cumulative distribution function, implying that the wild-type component always belongs to a log-normal distribution. Other distributions (such as the gamma) may be more appropriate. Although the log-normal can be replaced by any other distribution, the method does not allow direct comparison of various assumptions. In addition, the non-linear least squares regression approach requires appropriate starting values to ensure convergence of the used algorithm. We propose likelihood-based inference instead. Of course, because the area of application stays the same, there still remain the complexities of having censored data and a region of overlap between the wild-type and non-wild-type components. We address the latter difficulty with an idea similar to that of Turnidge *et al.* [8], namely constructing the likelihood in a cumulative fashion.

Denoting by  $Z_i$  the number of times MIC value  $i$  was observed over the  $n$  trials, the observed MIC groupings can be seen as possible outcomes of  $Z = (Z_1, \dots, Z_k) \sim \text{Mult}(n, p)$ , where  $p = (p_1, \dots, p_k)$  such that  $p_1 + \dots + p_k = 1$ . The maximum likelihood estimates for the multinomial probabilities  $p_i$  are just the observed relative frequencies  $\frac{Z_i}{n}$  (i.e. in a saturated model). Nevertheless, the main interest remains in identifying the most suitable parameters of the continuous wild-type distribution rather than those of the discrete multinomial distribution. We can achieve this by exploiting the fact that the observed groupings are actually the result of the censored readings of the dilution experiment. Hence, the multinomial probabilities corresponding to a certain outcome  $i$  can be rewritten as

$$\begin{cases} \tilde{p}_i = \pi * F(u_i; \theta) \dots \text{if } i = 1, \\ \tilde{p}_i = \pi * [F(u_i; \theta) - F(l_i; \theta)] \dots \text{if otherwise.} \end{cases} \quad (3)$$

where  $u_i$  and  $l_i$  are the respective upper and lower values of the  $i$ th MIC category and  $F(\cdot)$  represents the wild-type cumulative distribution function under consideration, with  $\theta$  its according parameters. In addition, the unknown parameter  $\pi$  accounts for the fact that the true MIC distribution is a mixture of the wild-type and non-wild-type components.

The length of the parameter vector  $\phi = (\theta, \pi)$  determines the number of MIC categories the procedure starts with. More specifically, compared to the length of  $\phi$ , one additional category is needed to render an unsaturated model. The idea is now to tentatively replace some of the multinomial probabilities with their parametric counterparts in (3). The probabilities of the remaining outcomes are left unchanged and

are thus to be estimated similar to those of the saturated model (i.e. the observed relative frequencies). The resulting sequence of likelihoods is specified by

$$l_j(\phi, p_{k_j+1}, \dots, p_k) = \sum_{i=1}^{k_j} z_i \log \tilde{p}_i + \sum_{i=k_j+1}^k z_i \log p_i, \quad (4)$$

with  $j = 1, \dots, k - 3$  and where  $k_j$  indicates how many of the original multinomial probabilities are replaced:  $k_j = j + \text{length of } \phi$ , with  $\phi = (\theta, \pi)$ . We can maximise this sequence to obtain several proposal estimates for the parameters of interest. A direct result of the parametric assumption is that less parameters are used in the construction of the likelihood when  $j$  increases. Because all, and hence the same, data are used in all steps, we can apply the AIC criterion to select the most appropriate parameter estimates.

It is a common practice to select the most suitable parameter estimates according to the minimum value of  $AIC = -2\log\text{likelihood} + 2P$ , with  $P$  the number of parameters. Alternatively, one can apply model averaging, especially in case several good fitting models are obtained with similar AIC values [13]. The idea is to combine the obtained estimates from  $M$  models into an averaged estimate as follows:

$$\hat{\theta}_a = \sum_{m=1}^M w_m \hat{\theta}_m,$$

where the  $w_m$  represents the so-called Akaike weights, defined as

$$w_m = \frac{\exp(-\frac{1}{2}\Delta_m)}{\sum_{j=1}^M \exp(-\frac{1}{2}\Delta_j)}.$$

These Akaike weights use the difference between the AIC value corresponding to a specific estimate  $\hat{\theta}_m$  and the minimum observed AIC value:  $\Delta_m = AIC_m - AIC_{min}$ . When this difference is small, the Akaike weight is large, and the corresponding estimate is allowed to contribute more to the averaged estimate. On the other hand, large differences are an indication of a less suitable estimate, and hence, the small Akaike weight reduces the influence of that estimate to the averaged one. Through the application of these Akaike weights, estimates are based on a range of models. Therefore, we obtain a more realistic indication of the variability because the uncertainty about the true model is accounted for.

The proposed model is based on two assumptions. First of all, we require a clear separation between the wild-type and non-wild-type population. More specifically, applying the transformation in (3), it is assumed that the two components of the MIC mixture distribution are relatively well separated and that the contribution of the non-wild-type component to the first  $k_j$  MIC categories is nearly negligible. Of course, this assumption of having a pure wild-type subset becomes less tenable when  $j$  increases and will result in an elevated AIC value. Second, we assume that the observed data arise from a true mixture distribution in the sense that the mixing weight  $\pi$  does not equal 1. This second assumption follows from a biological argument. Because of a constant positive selection, more and more bacterial isolates are becoming less susceptible and are therefore transferred to the non-wild-type. Because data result from random sampling, a non-negligible amount of isolates in the datasets under investigation can be attributed to the non-wild-type class, such that  $\pi$  does not equal 1.

### 3.3. Determination of cut-off values

The primary goal of the new method is to identify the distribution of the wild-type MIC value distribution and to estimate its corresponding parameters. Once the most suitable distribution for the wild-type component is determined, we can derive the cut-off values to distinguish between wild-type and non-wild-type populations by the construction of confidence intervals. Following Turnidge *et al.* [8], we select cut-offs at 0.1% (lower tail) and 99.9% (upper tail). The obtained values are consecutively rounded up or down to the nearest two-fold dilution, and the resulting 99.8% CI contains MIC values of the isolates that can almost surely be considered to be wild-types. Attention will only be paid to the upper limit of the interval as this value represents the boundary with the non-wild-type isolates. While Turnidge *et al.* [8] only focused on the normal distribution on the  $\log_2$  scale, we can follow a similar approach for other distributions. The idea is to consider the rounded MIC value corresponding to the 99.9% percentile as being the upper bound of the wild-type distribution.

An additional advantage of the newly developed multinomial-based method is that the cut-off value (and other percentiles) can be calculated for various distributional assumptions. Instead of considering only the most appropriate estimate based on the AIC values, we can obtain an averaged estimate across the different distributions following the procedure of the Akaike weights. Nevertheless, the obtained cut-off values should be interpreted with care as they are based on the extrapolated estimate of the wild-type distribution.

#### 4. Data analysis

In this section, we apply the introduced procedures to the three datasets presented in Section 2. Initially, we characterised the wild-type MIC distribution by using the method proposed by Turnidge *et al.* [8]. We compare the resulting estimates to those obtained through the application of the multinomial-based method in combination with the AIC criterion and the Akaike weights.

##### 4.1. *Streptococcus pneumoniae* and benzylpenicillin

The upper panel of Table I shows the typical outline of the procedure presented by Turnidge *et al.* [8].

The initial subset contains all isolates that have an MIC value less than or equal to 0.016 mg/L (i.e. the value corresponding to the first mode). It is observed that the absolute difference between the observed and the estimated numbers of isolates decreases in the second model fit and consecutively deteriorates again. Hence, the most optimal mean and standard deviation on the  $\log_2$  scale were found to be  $-6.34$  (0.01) and 0.62 (0.01), respectively. For comparison purposes, the lower panel of Table I shows an equivalent procedure, replacing the Gaussian CDF in (2) with a gamma distribution and applying the model to the data on the original MIC scale. Optimal shape and scale parameters were again identified from the model fit on the second subset and were estimated to be 6.15 (0.07) and 0.002 ( $2.56e-5$ ), respectively. Note that the difference between the observed and the estimated numbers of observations in this optimal subset is somewhat smaller compared to the output from the log-normal assumption. Nevertheless, this does not necessarily imply that a gamma distribution is more appropriate. Rather, a direct comparison of the two distributional assumptions should be carried out. We can achieve this with our new multinomial-based method, for which the results are summarised in Table II.

The top row corresponds to the model in which the initial four multinomial probabilities are replaced with their parametric counterparts in (3). In this model, it is hence assumed that the non-wild-type component of the MIC distribution is nearly zero for these initial four MIC categories. From the second row

**Table I.** Parameter estimates according to the method of Turnidge *et al.* [8], applied to the benzylpenicillin-*Streptococcus pneumoniae* data.

Log-normal CDF								
Endpoint	Number of observations				Mean		Standard deviation	
	Observed	Est.	Diff.	Std. Err.	Est.	Std.Err.	Est.	Std.Err.
0.016	12 063	34 746.28	22 683.28	4094.67	-5.64	0.12	0.93	0.04
0.03	<i>16 761</i>	<i>17 028.09</i>	<i>267.09</i>	<i>106.60</i>	<i>-6.34</i>	<i>0.01</i>	<i>0.62</i>	<i>0.01</i>
0.06	18 351	17 895.56	-455.44	388.78	-6.28	0.05	0.68	0.06
0.125	19 221	18 498.06	-722.94	424.57	-6.24	0.06	0.74	0.09
0.25	20 197	19 088.65	-1108.35	480.62	-6.19	0.08	0.81	0.11
Gamma CDF								
Subset	Number of observations				Shape		Scale	
	Observed	Est.	Diff.	Std. Err.	Est.	Std.Err.	Est.	Std.Err.
0.016	12 063	18 951.21	6888.21	763.34	5.40	0.22	0.003	1.84e-4
0.03	<i>16 761</i>	<i>16 898.48</i>	<i>137.48</i>	<i>30.68</i>	<i>6.15</i>	<i>0.07</i>	<i>0.002</i>	<i>2.56e-5</i>
0.06	18 351	17 754.10	-596.90	424.28	5.39	1.08	0.003	5.69e-4
0.125	19 221	18 359.13	-861.87	466.14	4.73	1.20	0.003	8.66e-4
0.25	20 197	18 951.53	-1245.47	519.05	4.05	1.21	0.004	1.26e-3

Est., estimated; Diff., difference; Std. Err., standard error.

The row in italic corresponds to the optimal fit based on the smallest difference between the observed and estimated number of observations.

**Table II.** Parameter estimates of the multinomial-based method with a normal and gamma CDF, applied to the benzylpenicillin–*Streptococcus pneumoniae* data.

Endpoint	Log-normal CDF				Gamma CDF			
	AIC	$\pi$	Mean (SE)	SD (SE)	AIC	$\pi$	Shape (SE)	Scale (SE)
0.016	104 928.3	1.00	−5.91 (0.01)	0.84 (0.01)	<i>104 891.5</i>	<i>0.76</i>	<i>5.12 (0.18)</i>	<i>0.003 (1.97e-4)</i>
0.03	<i>105 091.4</i>	<i>0.65</i>	<i>−6.34 (0.01)</i>	<i>0.66 (0.01)</i>	104 905.6	0.64	5.99 (0.09)	0.002 (3.99e-5)
0.06	105 763.2	0.70	−6.21 (0.01)	0.79 (0.00)	106 305.9	0.70	3.93 (0.05)	0.004 (5.11e-5)
0.125	107 577.0	0.73	−6.09 (0.01)	0.95 (0.01)	110 536.5	0.73	2.47 (0.03)	0.007 (8.66e-5)
0.25	111 457.6	0.77	−5.92 (0.01)	1.21 (0.01)	119 080.6	0.76	1.36 (0.01)	0.019 (2.17e-4)
Model-averaged		0.65	−6.34 (0.01)	0.66 (0.01)		0.76	5.12 (0.18)	0.003 (1.98e-4)

SE, standard error; SD, standard deviation.

The row in italic corresponds to the optimal fit based on the smallest AIC value for fits where  $\pi < 1$ .

onwards, each model replaces one additional multinomial probability. As it is biologically known in this case that the wild-type distribution is part of a mixture, the estimated  $\pi$  cannot be equal to one. Hence, model 1 (which is fitted on the subset of MIC values smaller than 0.016) is discarded when determining the most optimal fit in case of a log-normal first component. Initially, the AIC values are employed in the selection procedure. For a log-normal wild-type component, this results into a mean and standard deviation (on the  $\log_2$  scale) of  $-6.34 (0.01)$  and  $0.66 (0.01)$ , respectively. In case of a gamma first component, the most optimal shape and scale parameters are  $5.12 (0.09)$  and  $0.003 (1.97e-4)$ , respectively. One can find a graphical representation of the estimated densities in Figure 1(a). Comparing both models, the gamma distribution seems to be the preferred one. Alternatively, we can obtain averaged estimates by applying the Akaike weights. In this example, both selection procedures yield the same estimates as the AIC values clearly show a single best model. An additional feature of the newly developed method is that it provides an estimate for the mixing weight of the first component, which indicates the prevalence of wild-type isolates. In this example, it is seen that the log-normal first component receives a weight of  $0.65 (3.12e-3)$ , while the estimated weight of the gamma first component is  $0.74 (0.04)$ . However, because data are collected from different worldwide sources, we cannot use them to infer rates of susceptibility, and the previous results are just exemplary.

Based on the obtained estimates, we can calculate the median and 99.9% percentile of the respective distributions. In contrast to the original distribution parameters, these additional percentiles are directly comparable on the original scale. Therefore, the estimates resulting from both distributional assumptions can be regarded simultaneously, and an overall averaged estimate of both percentiles can be derived. We present the results in Table III, where the 99.9% percentile is also rounded to the nearest two-fold dilution in order to represent a possible cut-off value to distinguish between wild-type and non-wild-type isolates.

The estimates for the median obtained from the different estimation procedures are all very similar, ranging between 0.012 and 0.014. The range of estimated 99.9% percentiles is slightly larger, that is, from 0.037 to 0.051. Finally, it is seen that overall averaging results into the same cut-off value compared to the method by Turnidge *et al.* [8], that is, 0.031 mg/L. Because the true distribution is in fact unknown, we recommend overall averaging.

#### 4.2. *Escherichia coli* and ampicillin

Because the analysis was carried out in a similar fashion, we present only the final estimates for these data in Table III. In case of the EUCAST sample, assuming a log-normal first component resulted into an estimated mean and standard deviation (on the  $\log_2$  scale) of  $1.06 (0.01)$  and  $0.74 (0.01)$ , respectively, while the estimated shape and scale for the corresponding gamma distribution were equal to  $5.56 (0.08)$  and  $0.38 (0.01)$ , respectively. The model-averaged approach did not influence the former estimates but had a minor impact on the shape and scale parameters. Similar results are obtained for the EFSA sample, with an estimated mean and standard deviation (on the  $\log_2$  scale) of  $1.05 (0.01)$  and  $0.69 (0.02)$ , respectively. The estimated shape and scale for the corresponding gamma distribution were equal to  $5.91 (0.25)$  and  $0.36 (0.02)$ , respectively. For both data sources, the AIC values indicate that the gamma assumption is better compared to the log-normal. However, as argued previously, the overall averaged estimates are preferred. Based on the EUCAST data, the cut-off value is estimated to be 8 mg/L, while this breakpoint is only 4 mg/L for the EFSA data. A possible explanation for this discrepancy is that the two MIC

**Table III.** Results for the different data application examples.

Data (source)	Method	Log <sub>2</sub> Mean (SE)	Log <sub>2</sub> SD (SE)	Shape (SE)	Scale (SE)	Median (SE)	P99.9 (SE)	Cut-off
<i>S. pneumoniae</i> -benzylpenicillin (Turnidge)	T1	-6.34 (0.01)	0.62 (0.01)	-	-	0.012 (1.80e-5)	0.047 (4.86e-4)	0.031
	T2	-	-	6.15 (0.07)	0.002 (2.56e-5)	0.013 (2.63e-5)	0.037 (2.01e-4)	0.031
	AIC1	-6.34 (0.01)	0.66 (0.01)	-	-	0.012 (1.10e-5)	0.051 (2.64e-4)	0.063
	AV1	-6.34 (0.01)	0.66 (0.01)	-	-	0.012 (1.10e-5)	0.051 (2.64e-4)	0.063
	AIC2	-	-	5.12 (0.18)	0.003 (1.97e-4)	0.014 (6.11e-5)	0.045 (3.81e-4)	0.031
	AV2	-	-	5.12 (0.18)	0.003 (1.98e-4)	0.014 (6.23e-5)	0.045 (3.87e-4)	0.031
	AIC3	-	-	5.12 (0.18)	0.003 (1.97e-4)	0.014 (6.11e-5)	0.045 (3.81e-4)	0.031
	AV3	-	-	-	-	0.014 (6.23e-5)	0.045 (3.87e-4)	0.031
<i>E. coli</i> -ampicillin (EUCAST)	T1	1.04 (2.60e-3)	0.71 (3.70e-3)	-	-	2.05 (7.33e-3)	9.45 (3.22e-1)	8
	T2	-	-	4.65 (0.42)	0.48 (0.05)	2.06 (3.45e-2)	6.76 (3.68e-1)	8
	AIC1	1.06 (0.01)	0.74 (0.01)	-	-	2.08 (1.90e-2)	10.21 (5.71e-1)	8
	AV1	1.06 (0.01)	0.74 (0.01)	-	-	2.08 (1.91e-2)	10.20 (5.71e-1)	8
	AIC2	-	-	5.56 (0.08)	0.38 (0.01)	1.98 (7.69e-3)	5.97 (5.26e-2)	4
	AV2	-	-	5.59 (0.14)	0.38 (0.02)	1.98 (1.55e-2)	5.93 (8.11e-2)	4
	AIC3	-	-	5.56 (0.08)	0.38 (0.01)	1.98 (7.69e-3)	5.97 (5.26e-2)	4
	AV3	-	-	-	-	1.98 (1.55e-2)	5.93 (8.11e-2)	4
<i>E. coli</i> -ampicillin (EFSA)	T1	1.00 (0.02)	0.62 (0.03)	-	-	2.00 (5.38e-2)	7.59 (2.02)	8
	T2	-	-	5.59 (0.09)	0.36 (0.01)	2.00 (4.13e-3)	5.86 (4.71e-2)	4
	AIC1	1.05 (0.01)	0.69 (0.02)	-	-	2.07 (6.43e-2)	9.17 (1.64)	8
	AV1	1.05 (0.02)	0.69 (0.02)	-	-	2.07 (6.45e-2)	9.17 (1.64)	8
	AIC2	-	-	5.91 (0.25)	0.36 (0.02)	2.01 (2.39e-2)	5.88 (1.28e-1)	4
	AV2	-	-	5.97 (0.30)	0.36 (0.02)	2.01 (2.65e-2)	5.85 (1.61e-2)	4
	AIC3	-	-	5.91 (0.25)	0.36 (0.02)	2.01 (2.39e-2)	5.88 (1.28e-1)	4
	AV3	-	-	-	-	2.03 (4.75e-2)	6.95 (1.66)	8

In this table, the applied methods are denoted with a shorthand notation: T1 and T2 represent the method of Turnidge *et al.* [8] with, respectively, a log-normal and gamma first component; AIC1 and AIC2 indicate the multinomial-based method using minimum AIC as selection criterion with, respectively, a log-normal and gamma first component; AV1 and AV2 represent the multinomial-based method using model averaging with, respectively, a log-normal and gamma first component, while AIC3 and AV3 indicate estimates based on overall minimum AIC and overall model averaging.

SD, standard deviation; SE, standard error; EUCAST, European Committee on Antimicrobial Susceptibility Testing; EFSA, European Food Safety Authority.

distributions include data from different sources, geographical areas, and time periods. Regarding the mixing weight of the first component, the previous remark still holds, and the following estimates are again exemplary. The log-normal first component receives a weight of 0.68 (0.004) and 0.86 (0.02) for the EUCAST and EFSA data, respectively, while the gamma first component has respectively a weight of 0.63 (0.003) and 0.81 (0.01).

## 5. Simulation study

Based on the example obtained from the EUCAST website, that is, *E. coli* tested for susceptibility against ampicillin, three mixtures were constructed. In each of them, the non-wild-type component was approximated by a mixture of three log-normal densities. They differ however in the way the wild-type component was defined. In the first mixture, the wild-type component is assumed to follow a log-normal distribution, while the second mixture has a gamma first component. The third mixture can be seen as a combination of the former two because the first component is a mixture of the log-normal and gamma distributions from those respective mixtures. In the following expressions,  $C_1 = \log \mathcal{N}(1.87, 0.31)$ ,  $C_2 = \log \mathcal{N}(3.81, 0.31)$  and  $C_3 = \log \mathcal{N}(5.20, 0.24)$  represent the three components of the non-wild-type distribution. The models considered are

$$g_1(x) = 0.63 \log \mathcal{N}(0.72, 0.5) + 0.05C_1 + 0.14C_2 + 0.18C_3 \quad (5)$$

$$g_2(x) = 0.63\Gamma(5.59, 0.38) + 0.05C_1 + 0.14C_2 + 0.18C_3 \quad (6)$$

$$g_3(x) = 0.63\{0.5\Gamma(5.59, 0.38) + 0.5 \log \mathcal{N}(0.72, 0.5)\} + 0.05C_1 + 0.14C_2 + 0.18C_3 \quad (7)$$

The method of Turnidge *et al.* [8] and the newly developed multinomial-based method are applied to a total of 1000 samples from each aforementioned example mixture. We make two distributional assumptions: a log-normal first component and a gamma first component. Note that although both assumptions are used for the method proposed by Turnidge *et al.* [8], no direct comparison is possible to determine which of the parametric forms is most appropriate. On the other hand, the newly developed multinomial-based method can exploit the benefits of the AIC criterion. Regarding the latter approach, we select estimates based on the minimum AIC value as well as through the application of the Akaike weights. We apply these selection procedures to estimates obtained from a single run of the multinomial-based method (i.e. assuming either a log-normal or gamma first component) as well as to the total pool of estimates obtained from both runs. We consecutively use the resulting density estimates for the first component to determine the median, the 99.9% percentile and the cut-off value.

We compare the performance of the applied methods based on the estimated bias, variance and mean squared error (MSE). We summarise the results in Tables IV, V and VI for the three mixtures, respectively. For ease of representation, we use the shorthand notations from Table III. Regarding the estimates for the median, it is observed that AIC1 and AV1 perform best in case of mixture (5), while AIC2 and AV2 are most optimal for mixture (6). In both examples, these optimal approaches correspond to the multinomial-based method, which assumes the correct underlying distribution. It is also apparent that the performance of the overall averaged selection procedure (AV3) is close to optimal. A similar observation applies to mixture (7), where the true underlying distribution could not be assumed. Because of a decreased bias, the overall-averaging approach performs best in the latter example. Regarding the other characteristics, we observe a similar performance of the respective procedures.

In order to obtain an idea of the overall estimation quality of the new method, we calculated the Kullback–Leibler distance as well. Denoting by  $h(x)$  the true first component and by  $g(x; \theta)$  the obtained estimate, the KL distance corresponds to

$$KL = \int h(x) \log \frac{h(x)}{g(x; \theta)} dx.$$

We calculate the KL distance for each of the samples, and we present an averaged value in Table VII. In all three examples, it is noted that the estimation performance increases with increasing sample sizes. In case the true underlying distribution was used for the wild-type component, it was also found to be the most appropriate in terms of reduction in averaged KL distance. Nevertheless, because in practice the underlying model is unknown, most attention should be given to the overall averaged approach (AV3).

**Table IV.** Simulation results for mixture (5).

Parameter	Method	Sample size = 5000			Sample size = 10000			Sample size = 20000			Sample size = 30000		
		Bias	Variance	MSE	Bias	Variance	MSE	Bias	Variance	MSE	Bias	Variance	MSE
Median	T1	0.110	0.001	0.013 (7)	0.109	0.000	0.012 (7)	0.108	0.000	0.012 (7)	0.107	0.000	0.012 (7)
	T2	0.125	0.001	0.016 (8)	0.125	0.000	0.016 (8)	0.126	0.000	0.016 (8)	0.125	0.000	0.016 (8)
	AIC1	0.013	0.001	0.002 (1)	0.010	0.001	0.001 (1)	0.010	0.000	0.000 (1)	0.010	0.000	0.000 (1)
	AV1	0.016	0.002	0.002 (2)	0.010	0.001	0.001 (2)	0.010	0.000	0.000 (2)	0.010	0.000	0.000 (2)
	AIC2	-0.072	0.001	0.006 (5.5)	-0.073	0.000	0.006 (5.5)	-0.072	0.000	0.005 (5.5)	-0.073	0.000	0.005 (5.5)
	AV2	-0.072	0.001	0.006 (5.5)	-0.073	0.000	0.006 (5.5)	-0.072	0.000	0.005 (5.5)	-0.073	0.000	0.005 (5.5)
	AIC3	0.000	0.002	0.002 (4)	0.004	0.001	0.001 (4)	0.009	0.000	0.000 (4)	0.009	0.000	0.000 (4)
	AV3	-0.002	0.002	0.002 (3)	0.002	0.001	0.001 (3)	0.008	0.000	0.000 (3)	0.009	0.000	0.000 (3)
	P99.9	T1	2.228	0.242	5.208 (6)	2.204	0.126	4.983 (6)	2.172	0.053	4.771 (6)	2.161	0.028
T2	-1.665	0.034	2.807 (5)	-1.663	0.017	2.783 (5)	-1.662	0.008	2.772 (5)	-1.664	0.005	2.773 (5)	
AIC1	0.147	0.336	0.357 (1)	0.104	0.149	0.160 (1)	0.095	0.073	0.082 (1)	0.101	0.050	0.060 (1)	
AV1	0.209	0.400	0.444 (2)	0.109	0.155	0.167 (2)	0.095	0.073	0.082 (2)	0.101	0.050	0.060 (2)	
AIC2	-3.736	0.022	13.977 (7.5)	-3.741	0.012	14.010 (7.5)	-3.743	0.006	14.014 (7.5)	-3.742	0.004	14.003 (7.5)	
AV2	-3.736	0.022	13.977 (7.5)	-3.741	0.012	14.010 (7.5)	-3.743	0.006	14.014 (7.5)	-3.742	0.004	14.003 (7.5)	
AIC3	-0.399	2.101	2.260 (4)	-0.157	1.063	1.087 (4)	0.038	0.293	0.295 (4)	0.074	0.154	0.159 (4)	
AV3	-0.596	1.205	1.560 (3)	-0.289	0.712	0.796 (3)	0.000	0.232	0.232 (3)	0.069	0.129	0.134 (3)	
Cut-off	T1	2.592	14.032	20.750 (8)	1.888	11.551	15.116 (8)	1.104	7.621	8.840 (6)	0.696	5.089	5.573 (6)
	T2	0.000	0.000	0.000 (1.5)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)
	AIC1	0.040	0.319	0.320 (4)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)
	AV1	0.000	0.000	0.000 (1.5)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)
	AIC2	-3.048	2.905	12.195 (6.5)	-3.332	2.228	13.330 (6.5)	-3.712	1.070	14.849 (7.5)	-3.856	0.556	15.425 (7.5)
	AV2	-3.048	2.905	12.195 (6.5)	-3.332	2.228	13.330 (6.5)	-3.712	1.070	14.849 (7.5)	-3.856	0.556	15.425 (7.5)
	AIC3	-0.408	1.755	1.992 (5)	-0.220	0.832	0.881 (5)	-0.060	0.237	0.240 (5)	-0.028	0.111	0.112 (5)
AV3	-0.028	0.111	0.112 (3)	-0.012	0.048	0.048 (4)	-0.004	0.016	0.016 (4)	-0.008	0.032	0.032 (4)	

The bias, variance and mean squared error (MSE) are calculated for the median, 99.9% percentile and the cut-off value (i.e. the 99.9% percentile rounded to the nearest two-fold dilution). Considered sample sizes were 5000, 10000, 20000 and 30000.

The italic numbers between brackets determine the order of the employed methods based on increasing AIC.

**Table V.** Simulation results for mixture (6).

Parameter	Method	Sample size = 5000			Sample size = 10000			Sample size = 20000			Sample size = 30000		
		Bias	Variance	MSE	Bias	Variance	MSE	Bias	Variance	MSE	Bias	Variance	MSE
Median	T1	0.073	0.000	0.006 (5)	0.072	0.000	0.005 (5)	0.072	0.000	0.005 (5)	0.071	0.000	0.005 (5)
	T2	0.081	0.001	0.007 (6)	0.081	0.000	0.007 (6)	0.082	0.000	0.007 (6)	0.081	0.000	0.007 (6)
	AIC1	0.103	0.002	0.013 (8)	0.107	0.001	0.013 (8)	0.110	0.000	0.013 (8)	0.110	0.000	0.012 (8)
	AV1	0.099	0.002	0.011 (7)	0.104	0.001	0.012 (7)	0.108	0.000	0.012 (7)	0.109	0.000	0.012 (7)
	AIC2	0.007	0.001	0.001 (1)	0.007	0.000	0.000 (1)	0.007	0.000	0.000 (1.5)	0.006	0.000	0.000 (1.5)
	AV2	0.007	0.001	0.001 (2)	0.007	0.000	0.000 (2)	0.007	0.000	0.000 (1.5)	0.006	0.000	0.000 (1.5)
	AIC3	0.022	0.002	0.002 (4)	0.012	0.001	0.001 (4)	0.008	0.000	0.000 (4)	0.007	0.000	0.000 (4)
AV3	0.028	0.001	0.002 (3)	0.015	0.001	0.001 (3)	0.008	0.000	0.000 (3)	0.007	0.000	0.000 (3)	
P99.9	T1	3.495	0.077	12.293 (6)	3.489	0.038	12.214 (6)	3.478	0.018	12.114 (6)	3.477	0.012	12.104 (6)
	T2	0.822	0.022	0.697 (3)	0.828	0.011	0.696 (3)	0.827	0.006	0.690 (5)	0.826	0.004	0.687 (5)
	AIC1	4.243	0.532	18.531 (8)	4.328	0.269	18.997 (8)	4.356	0.116	19.092 (8)	4.361	0.079	19.095 (8)
	AV1	4.176	0.423	17.860 (7)	4.266	0.253	18.456 (7)	4.325	0.128	18.831 (7)	4.348	0.087	18.988 (7)
	AIC2	0.033	0.027	0.028 (1)	0.035	0.013	0.014 (1)	0.031	0.007	0.007 (1.5)	0.030	0.005	0.006 (1.5)
	AV2	0.033	0.027	0.028 (2)	0.035	0.013	0.014 (2)	0.031	0.007	0.007 (1.5)	0.030	0.005	0.006 (1.5)
	AIC3	0.741	2.361	2.911 (5)	0.251	0.842	0.905 (5)	0.073	0.183	0.188 (4)	0.051	0.095	0.098 (4)
AV3	1.013	1.396	2.423 (4)	0.399	0.664	0.824 (4)	0.096	0.121	0.130 (3)	0.064	0.070	0.074 (3)	
Cut-off	T1	4.000	0.000	16.000 (5.5)	4.000	0.000	16.000 (5.5)	4.000	0.000	16.000 (6.5)	4.000	0.000	16.000 (6.5)
	T2	4.000	0.000	16.000 (5.5)	4.000	0.000	16.000 (5.5)	4.000	0.000	16.000 (6.5)	4.000	0.000	16.000 (6.5)
	AIC1	4.112	0.884	17.793 (7.5)	4.008	0.064	16.128 (7.5)	4.000	0.000	16.000 (6.5)	4.000	0.000	16.000 (6.5)
	AV1	4.112	0.884	17.793 (7.5)	4.008	0.064	16.128 (7.5)	4.000	0.000	16.000 (6.5)	4.000	0.000	16.000 (6.5)
	AIC2	2.216	3.957	8.868 (1.5)	2.500	3.754	10.004 (1.5)	2.512	3.742	10.052 (1.5)	2.540	3.712	10.164 (1.5)
	AV2	2.216	3.957	8.868 (1.5)	2.500	3.754	10.004 (1.5)	2.512	3.742	10.052 (1.5)	2.540	3.712	10.164 (1.5)
	AIC3	2.652	3.867	10.900 (3)	2.604	3.639	10.420 (3)	2.524	3.729	10.100 (3)	2.548	3.703	10.196 (3)
AV3	3.636	1.325	14.545 (4)	3.232	2.485	12.930 (4)	2.656	3.573	10.628 (4)	2.620	3.619	10.484 (4)	

The bias, variance and mean squared error (MSE) are calculated for the median, 99.9% percentile and the cut-off value (i.e. the 99.9% percentile rounded to the nearest two-fold dilution). Considered sample sizes were 5000, 10000, 20000 and 30000.

The italic numbers between brackets determine the order of the employed methods based on increasing AIC.

**Table VI.** Simulation results for mixture (7).

Parameter	Method	Sample size = 5000			Sample size = 10000			Sample size = 20000			Sample size = 30000		
		Bias	Variance	MSE	Bias	Variance	MSE	Bias	Variance	MSE	Bias	Variance	MSE
Median	T1	0.099	0.001	0.011 (7)	0.098	0.000	0.010 (7)	0.097	0.000	0.010 (7)	0.097	0.000	0.009 (7)
	T2	0.115	0.001	0.014 (8)	0.114	0.000	0.013 (8)	0.114	0.000	0.013 (8)	0.113	0.000	0.013 (8)
	AIC1	0.037	0.002	0.003 (2)	0.032	0.001	0.002 (2)	0.029	0.000	0.001 (2)	0.030	0.000	0.001 (2)
	AV1	0.042	0.002	0.004 (6)	0.034	0.001	0.002 (3)	0.029	0.000	0.001 (3)	0.030	0.000	0.001 (3)
	AIC2	-0.056	0.001	0.004 (4.5)	-0.056	0.000	0.003 (5.5)	-0.057	0.000	0.003 (5.5)	-0.057	0.000	0.003 (5.5)
	AV2	-0.056	0.001	0.004 (4.5)	-0.056	0.000	0.003 (5.5)	-0.057	0.000	0.003 (5.5)	-0.057	0.000	0.003 (5.5)
	AIC3	0.011	0.003	0.003 (3)	0.009	0.002	0.002 (4)	0.017	0.001	0.002 (4)	0.022	0.001	0.001 (4)
AV3	0.012	0.002	0.003 (1)	0.010	0.001	0.002 (1)	0.015	0.001	0.001 (1)	0.021	0.001	0.001 (1)	
P99.9	T1	2.394	0.190	5.920 (6)	2.364	0.074	5.661 (6)	2.345	0.034	5.532 (6)	2.349	0.021	5.542 (6)
	T2	-1.182	0.035	1.433 (1)	-1.184	0.016	1.418 (3)	-1.190	0.008	1.424 (3)	-1.188	0.006	1.417 (4)
	AIC1	1.065	0.510	1.644 (2)	0.975	0.190	1.141 (1)	0.936	0.082	0.958 (1)	0.957	0.058	0.973 (1)
	AV1	1.163	0.481	1.833 (3)	1.010	0.220	1.241 (2)	0.938	0.084	0.963 (2)	0.957	0.058	0.974 (2)
	AIC2	-2.980	0.024	8.906 (7.5)	-2.983	0.013	8.911 (7.5)	-2.991	0.006	8.953 (7.5)	-2.986	0.004	8.921 (7.5)
	AV2	-2.980	0.024	8.906 (7.5)	-2.983	0.013	8.911 (7.5)	-2.991	0.006	8.953 (7.5)	-2.986	0.004	8.921 (7.5)
	AIC3	-0.054	3.687	3.690 (5)	-0.013	3.040	3.041 (5)	0.360	2.004	2.134 (5)	0.624	1.252	1.641 (5)
AV3	-0.110	1.899	1.911 (4)	-0.034	1.746	1.747 (4)	0.286	1.410	1.492 (4)	0.565	0.984	1.303 (3)	
Cut-off	T1	0.528	3.949	4.228 (6)	0.112	0.884	0.897 (5)	0.008	0.064	0.064 (4)	0.000	0.000	0.000 (2.5)
	T2	0.000	0.000	0.000 (1)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2.5)
	AIC1	0.064	0.508	0.513 (4)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2.5)
	AV1	0.008	0.064	0.064 (2)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2)	0.000	0.000	0.000 (2.5)
	AIC2	-2.796	3.370	11.187 (7.5)	-3.008	2.987	12.035 (7.5)	-3.444	1.917	13.778 (7.5)	-3.588	1.480	14.353 (7.5)
	AV2	-2.796	3.370	11.187 (7.5)	-3.008	2.987	12.035 (7.5)	-3.444	1.917	13.778 (7.5)	-3.588	1.480	14.353 (7.5)
	AIC3	-0.752	3.022	3.587 (5)	-0.748	2.435	2.994 (6)	-0.524	1.823	2.098 (6)	-0.292	1.084	1.169 (6)
AV3	-0.100	0.486	0.496 (3)	-0.120	0.466	0.480 (4)	-0.108	0.421	0.432 (5)	-0.100	0.390	0.400 (5)	

The bias, variance and mean squared error (MSE) are calculated for the median, 99.9% percentile and the cut-off value (i.e. the 99.9% percentile rounded to the nearest two-fold dilution). Considered sample sizes were 5000, 10000, 20000 and 30000.

The italic numbers between brackets determine the order of the employed methods based on increasing AIC.

**Table VII.** Kullback–Leibler distances for the three simulation mixtures under consideration.

Sample size	T1	T2	AIC1	AV1	AIC2	AV2	AIC3	AV3
Mixture (5)								
5000	0.014	0.024	0.001	0.002	0.081	0.081	0.013	0.006
10 000	0.013	0.024	0.001	0.001	0.081	0.081	0.006	0.003
20 000	0.013	0.024	0.000	0.000	0.081	0.081	0.002	0.001
30 000	0.013	0.024	0.000	0.000	0.081	0.081	0.001	0.001
Mixture (6)								
5000	0.036	0.015	0.048	0.047	0.001	0.001	0.009	0.009
10 000	0.036	0.015	0.050	0.048	0.000	0.000	0.003	0.003
20 000	0.035	0.015	0.050	0.049	0.000	0.000	0.001	0.001
30 000	0.035	0.014	0.050	0.050	0.000	0.000	0.000	0.000
Mixture (7)								
5000	0.014	0.021	0.004	0.005	0.057	0.057	0.019	0.007
10 000	0.014	0.021	0.003	0.003	0.057	0.057	0.016	0.007
20 000	0.013	0.021	0.002	0.002	0.057	0.057	0.010	0.005
30 000	0.013	0.021	0.002	0.002	0.056	0.056	0.007	0.004

The latter approach is close to optimal in each of the examples and outperforms the method of Turnidge *et al.* [8]. Therefore, it seems that the new approach is a valuable alternative to the method of Turnidge *et al.* [8].

## 6. Discussion

In this paper, we considered a new approach for estimating the wild-type component of an MIC distribution. The entire MIC distribution was represented by a mixture to reflect the underlying population heterogeneity. Adopting a local view, we estimate the first component by using maximum likelihood, while the contribution of the presumable non-wild-type component was taken into account through the multinomial distribution. Because the new method fits within the likelihood framework, we can compare several distributional assumptions for the wild-type population based on the AIC criterion, or, alternatively, a model-averaged approach can be followed. In this way, we can quantify the wild-type distribution of the MIC distribution of interest when a representative sample of the desired antibiotic–bacterium combination is available. Special attention was paid to the 99.9% percentile, which was considered as a possible upper bound of the wild-type distribution. Rounding the obtained value to the nearest two-fold dilution provides a cut-off value to discriminate between the wild-type and the non-wild-type subpopulations. Especially for the latter characteristic, the model-averaged approach is thought to be very appealing as the influence of possible miss-specified distributional assumptions is minimised. In this regard, we believe that the new method is a valuable alternative to the approach described by Turnidge *et al.* [8], where the authors make use of a non-linear least square regression approach. The simulation study also showed promising results when comparing our method to the described competitor.

Important to note is that because data are collected across laboratories, the MIC values in the data examples are probably sensitive to between-lab variability. In addition, there is also variability inherent to the applied test. Both sources of variability could have an influence on the obtained estimates and are therefore a plausible explanation for the minor discrepancies observed in the *E. coli*–ampicillin example in Table III.

Although the procedure presented in this paper provides interesting information regarding the wild-type first component of the MIC distribution, it is of even more importance to obtain a global view of the entire MIC mixture. After all, the emergence of resistant isolates is of primary interest in the field of antimicrobial susceptibility testing, with the monitoring of isolates exhibiting reduced susceptibility being the major objective. While the current procedure already provides an estimate for the prevalence of non-wild-type isolates, there is still room for improvement. More specifically, the assumption of having a well-separated wild-type and non-wild-type component could be relaxed through the consideration of the entire mixture in (1). Accounting for the non-wild-type component in a non-parametric way, Jaspers *et al.* [14] provide an alternative estimation procedure in which the estimator developed in this paper

is used to construct a semi-parametric mixture model that is able to estimate the MIC distribution on the full continuous scale. We can use this mixture to determine the prevalence of non-wild-type isolates in a more accurate way. In addition, with the full mixture approach, we can use model-based classification to attribute an isolate to the wild-type or non-wild-type component, thereby solving the issue of extrapolation we mentioned in Section 3.3. Nevertheless, the developed procedures still require further refinements and extensions such as the modelling of time trends and the estimation of multivariate distributions.

## Acknowledgements

Research Foundation Flanders (Fonds Wetenschappelijk Onderzoek, FWO), grant 11E2913N, supported the research of the first author. We gratefully acknowledged the support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy). For the simulations and bootstraps, we used the infrastructure of the Vlaams Supercomputer Centrum (VSC) – Flemish Supercomputer Center, funded by the Hercules Foundation and the Flemish Government – department Economie, Wetenschappen en Innovatie (EWI). The authors are grateful to EFSA for the approval to use the *Ampicillin* data.

## References

1. Tenover F. Mechanisms of antimicrobial resistance in bacteria. *The American Journal of Medicine* 2006; **119**:S3–S10.
2. Strasfeld L, Chou S. Antiviral drug resistance: mechanisms and clinical implications. *Infectious Disease Clinics of North America* 2010; **24**:413–437.
3. Chen D, McGeer A, de Azavedo J, Low D. Decreased susceptibility of *Streptococcus pneumoniae* to fluoroquinolones in Canada. *New England Journal of Medicine* 1999; **341**:233–239.
4. Kahlmeter G, Brown DFJ, Goldstein FW, MacGowan AP, Mouton JW, Osterlund A, Rodloff A, Steinbakk M, Urbaskova P, Vatopoulos A. European harmonization of MIC breakpoints for antimicrobial susceptibility testing of bacteria. *Journal of Antimicrobial Chemotherapy* 2003; **52**:145–148.
5. Annis D, Craig B. Statistical properties and inference of the antimicrobial MIC test. *Statistics in Medicine* 2005; **24**:3631–3644.
6. Craig B. Modeling approach to diameter breakpoint determination. *Diagnostic Microbiology and Infectious Disease* 2000; **36**:193–202.
7. Lee MLT, Whitmore GA. Statistical inference for serial dilution assay data. *Biometrics* 1999; **55**:1215–1220.
8. Turnidge J, Kahlmeter G, Kronvall G. Statistical characterisation of bacterial wild-type MIC value distributions and the determination of epidemiological cut-off values. *Clinical Microbiology and Infection* 2006; **12**:418–425.
9. Böhning D. A vertex-exchange-method in D-optimal design theory. *Metrika* 1986; **33**:337–347.
10. Schellhase C, Kauermann G. Density estimation and comparison with a penalized mixture approach. *Computational Statistics* 2012; **27**:757–777.
11. Goldstein F. Penicillin-resistant *Streptococcus pneumoniae*: selection by both  $\beta$ -lactam and non- $\beta$ -lactam antibiotics. *Journal of Antimicrobial Chemotherapy* 1999; **44**:141–144.
12. Hakenbeck R, Grebe T, Zähner D, Stock JB.  $\beta$ -lactam resistance in *Streptococcus pneumoniae*: penicillin-binding proteins and non-penicillin-binding proteins. *Molecular Biology* 1999; **33**:673–678.
13. Burnham K, Anderson D. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer: New York, 2002.
14. Jaspers S, Aerts M, Verbeke G, Beloeil P. A new semi-parametric mixture model for interval censored data, with applications in the field of antimicrobial resistance. *Computational Statistics and Data Analysis* 2013; **nv**:np. DOI: 10.1016/j.csda.2013.01.024.