Integrating forecast probabilities in antibiograms – a way to guide antimicrobial prescriptions more reliably?

Florian P. Maurer¹, Patrice Courvalin², Erik C. Böttger¹, Michael Hombach²‡

¹) Institut für Medizinische Mikrobiologie, Universität Zürich, 8006 Zürich, Schweiz
²) Unite des Agents Antibactériens, Institut Pasteur, 75724 Paris, France

Short title: Reporting forecast uncertainties in antibiograms

Keywords: BSBL, ESBL, ampC, resistotype, clinical breakpoints, antibiogram

Corresponding author:
Michael Hombach, M.D.
Institut für Medizinische Mikrobiologie
Universität Zürich
Gloriastr. 30/32
8006 Zürich
Switzerland
Phone: 0041 44 634 27 00
Fax: 0041 634 49 06
Email: mhombach@imm.uzh.ch
Abstract

Antimicrobial susceptibility testing (AST) assigns pathogens to “susceptible” or “resistant” clinical categories based on clinical breakpoints (CBPs) derived from minimal inhibitory concentrations or inhibition zone diameters, and indicates the likelihood for therapeutic success. AST reports do not provide quantitative measures for reliability of such categorization. Thus, it is currently impossible for clinicians to estimate the technical forecast uncertainty of an AST result regarding clinical categorization. AST error rates depend on the localization of pathogen populations in relation to CBPs. Bacterial species are, however, not homogeneous, and subpopulations behave differently with respect to AST results. We addressed how AST reporting errors differ between isolates with or without acquired drug resistance determinants. Using as example the beta-lactams and their most important resistance mechanisms, we analyzed different pathogen populations for their individual reporting error probabilities. Categorization error rates were significantly higher for bacterial populations harboring resistance mechanisms as compared to the wild-type population. Reporting errors for amoxicillin/clavulanic acid and piperacillin/tazobactam in Escherichia coli were almost exclusively due to broad-spectrum and extended spectrum beta-lactamase (ESBL)-producing microorganisms (79% and 20% of all errors, respectively). Clinicians should be aware of the significantly increased risk of erroneous AST reports for isolates producing beta-lactamases, particularly ESBL and AmpC. Including probability indicators for interpretation would improve AST reports.
Introduction

Reports of antimicrobial susceptibility testing (AST) data have been used for decades to translate *in vitro* laboratory results into predictions of clinical outcome by interpretative categorization of pathogens as “susceptible” or “resistant” based on clinical breakpoints (CBP) (1, 2). The susceptible category implies a high likelihood for therapeutic success if standard dosing is applied, whereas the resistant category implies probable therapeutic failure (3, 4). In contrast to other laboratory tests, such as quantitative PCRs or serologic parameters, for which variation coefficients indicating measurement precision are usually reported, or can at least be requested, AST reports do not provide such indicators (5, 6). As a consequence, it is currently impossible for the clinician to estimate the forecast uncertainty of an AST categorization and the related reliability of the predicted clinical outcome. AST categorizations based on MIC or inhibition zone measurements close to the CBPs are reported as “susceptible” or “resistant” without further comment, although the statistical error probability increases the closer a MIC or zone diameter measurement is to the CBP (7). In consequence, providing levels of reliability for AST categorization would improve clinical decisions in antimicrobial therapy.

For a long time, AST categorization was not only based on MIC and zone diameter measurements but also on the detection of individual resistance mechanisms, i.e. interpretative reading (8-10). Even if *in vitro* results indicated susceptibility to a drug, the reported category was edited to resistant if the presence of a resistance mechanism was confirmed, e.g. in the case of extended spectrum beta-lactamases (ESBLs). While this strategy is retained, e.g. for methicillin resistance in *Staphylococcus aureus*, CLSI and EUCAST recently abandoned editing of AST reports based on the detection of ESBLs (3, 11-13). AST categorization now depends...
solely on MIC and/or zone diameter measurement. However, AST methods show considerable technical and biological variation (2, 7, 14-16). The rate of major, very major or minor clinical categorization errors for a given antibiotic depends on i) the presence, width, or absence of a “grey” (or intermediate) zone between the susceptible and resistant category, ii) the relative position of a population to the CBP, and iii) the sum of methodological imprecision and biological variation (7). Since it is reasonable to assume that subpopulations may behave differently with respect to the latter two aspects, we addressed to what extent AST reporting error rates differ between wild type isolates (no acquired drug resistance determinants) and isolates encoding a resistance mechanism (termed “resistotypes” in this work). Beta-lactam resistance in the three most prevalent *Enterobacteriaceae* species was chosen as a model, since treatment of ESBL producers with penicillin-inhibitor combinations and/or newer cephalosporins is subject of an ongoing debate (17-21).
Materials and Methods

Bacterial strains. In total, 7,148 non-duplicate clinical isolates recovered in our laboratory between 2010 and 2013 were included in the study, comprising 4,287 \textit{Escherichia coli}, 1,886 \textit{Klebsiella pneumoniae}, and 975 \textit{Enterobacter cloacae}. Isolates of the same species were considered duplicate(s) if they i) originated from the same patient, and ii) showed at most one major (susceptible/resistant) AND two minor (susceptible/intermediate and/or intermediate/resistant) differences in AST interpretation. The numbers of various phenotypes for a species can be retrieved from Table 1. For some isolates not all zone diameters were available resulting in lower numbers of data for certain drug/species combinations (Table 1).

Susceptibility testing. Susceptibility testing was done by disk diffusion according to EUCAST recommendations (22). Antibiotic disks and Mueller-Hinton agar were obtained from Becton Dickinson, Franklin Lakes, NJ. Inhibition zone diameters were recorded using the Sirweb/Sirsan system (i2a, Montpellier, France) to standardize reading precision (23). ESBL and AmpC detection was done as previously described (24, 25). Reading precision of the Sirscan instrument was determined by 100 repeat measurements of the same plate/inhibition zone (see Table 2).

Interpretation errors. Differences in clinical categorization, referred to as "discrepancies" are split according to therapeutic implications: Those resulting in erratic assignment of bacterial isolates to adjacent interpretative categories (susceptible to intermediate, intermediate to susceptible, intermediate to resistant, resistant to intermediate) are referred to as "minor errors". Erroneous categorisation...
of true-susceptible isolates as resistant are referred to as “major errors” leading to unnecessary restriction of therapeutic options. The most serious clinical implications result from “very major errors”, i.e. categorisation of true-resistant isolates as susceptible, as there is a high likelihood of therapeutic failure.

**Phenotype/resistotype definition:** *E. coli* were considered wild-type, i.e. devoid of resistance mechanisms to beta-lactams, if the isolates were susceptible to ampicillin (according to EUCAST CBPs) to exclude the presence of broad-spectrum beta-lactamases (BSBLs) such as TEM-1/2 or SHV-1, and to cefoxitin (according to the EUCAST CBPs) to exclude porin deficiency and/or overexpression of efflux systems. To exclude cephalosporinases, isolates were required to be cephalothin wild-type as per in house epidemiological cut-off (ECOFF) of 10 mm (30 µg disk, data not shown). *K. pneumoniae* isolates were considered wild-type (i.e. chromosomally encoded SHV-1 beta-lactamase-producers only) if ESBL, AmpC, and carbapenemase production had been excluded (24-27) and if cefoxitin as a marker for porin deficiency was susceptible according to EUCAST CBPs. *E. cloacae* was considered wild-type, if ESBL and carbapenemase production had been excluded (24-26, 28), and if cefpodoxime was found susceptible according to EUCAST CBPs to exclude isolates with a derepressed chromosomal AmpC. *E. cloacae* isolates susceptible to cefpodoxime were checked for resistance to ceftiraxone, cefotaxime, ceftazidime, and cefepime and found susceptible. Any colony within the inhibition zone of a 3rd generation cephalosporin was counted for zone diameter measurement. BSBL-producing *E. coli* were defined as ampicillin resistant, non-ESBL, non-AmpC, carbapenemase-negative and cefoxitin susceptible. *E. cloacae* with derepressed AmpC were defined as cefpodoxime resistant, non-ESBL, and carbapenemase-negative.
Statistical analysis. Error probabilities were calculated separately for all diameter values adjacent to the interpretative category borders as described previously (7). In brief, theoretical upper and lower tails of the standard normal distribution were used to calculate probability distributions for actual diameter measurements which can be assumed to be normally distributed around the mean diameters obtained by disk diffusion with respect to different cut-offs (29, 30). Standard deviations of 120 independent measurements of *E. coli* ATCC 25922 (EUCAST quality control strain) are listed in Table 2. Standard deviations of *E. coli* ATCC 25922 were in general agreement with those derived from EUCAST quality control (QC) tables and were thus used as the basis for probability calculations (31). In addition 100 independent measurements of nine clinical *E. coli* strains (three wild-type isolates, two BSBL-producing isolates, three ESBL, and one AmpC-producing isolates were included to test for the influence of the genotype on measurement precision. To tests for statistical significance the Mann-Whitney-U- test was applied.

Software. All calculations were done using the IBM SPSS statistics software version 20 (IBM Corporation, Armonk, NY) and the Microsoft Excel 2010 software (Microsoft Corporation, Redmond, WA).
Results

Disk diffusion diameters for a total of 7,148 non-duplicate clinical isolates were used in the study. Statistical probabilities for antibiogram misclassifications were determined for individual sub-populations, i.e. resistotypes (wild-type, BSBL, ESBL, AmpC) as previously described (7).

The total number of interpretation errors for a given species/drug combination depends on four variables: i) technical measurement precision, ii) range of population distributions, iii) distance of the population’s lowest diameter or MIC value from the next CBP, and iv) relative resistotype prevalence. A paradigmatic example is depicted in Figure 1: Higher measurement precision lowers the probability of assigning a population’s tail area to the false category (Figure 1A); larger diameter ranges of a population lead to higher numbers of isolates in the area at risk (Figure 1 B); a low distance of a population from the CBP increases the probability of assigning a population’s tail area to the false category (Figure 1 C); the higher the prevalence of a resistotype clustering around the CBP, the higher the number of isolates in the area at risk for interpretation errors (Figure 1 D).

Technical measurement precision was similar for all resistotypes as reflected in average standard deviations of resistotypes (see Table 2). In the critical range for interpretation errors, i.e. diameter values flanking CBPs (15-25 mm) one-fold standard deviations of all resistotype/drug combinations ranged from 1.0 to 1.5 mm. This value is consistent with EUCAST methodological precision requirements (QC ranges of 4 to 6 reflecting a zone of two-fold standard deviation around the target value, i.e. a one-fold standard deviation of 1 to 1.5 mm) (31). Therefore, standard deviations of the EUCAST QC strain E. coli ATCC 25922 were considered a reasonable basis for error calculations.
Figure 2 depicts the above-mentioned factors for the *E. coli* study population and amoxicillin-clavulanic acid: The closer the median diameter of a resistotype to the CBP the more isolates in the zone of increased error risk. While few wild-type isolates (mean diameter 25 mm, distance to CBP 8 mm) are located in the area at risk for AST errors (Figure 2A), significantly more BSBL (mean diameter 20 mm, distance to CBP 3 mm, Figure 2B) and ESBL isolates (mean diameter 17 mm, distance to CBP 0 mm, Figure 2C) are found in the area at risk. Finally, the lower mean diameter of the AmpC resistotype (10 mm, Figure 2D) increased the distance to the CBP to -7 mm, decreasing the number isolates in the zone of highest risk for misclassification. The influence of population mean diameters distances to CBPs was also reflected in our analysis: The distances of the median diameter of the wild-type *E. coli* population to the next EUCAST CBP ranged from 7 mm for piperacillin/tazobactam and ceftazidime to 12 mm for cefotaxime. The corresponding rates of expected interpretation errors were low (0.3% minor errors for ceftazidime, < 0.1% errors for the other drugs, Table 1). In contrast, the distance of the median diameter of the *E. coli* ESBL resistotype to the next EUCAST CBP varied significantly for different beta-lactam drugs, ranging from 0 mm for amoxicillin/clavulanic acid (i.e. the median diameter of ESBL positive *E. coli* was equal to the EUCAST CBP) to -11 mm from the resistant EUCAST CBP for cefotaxime and ceftriaxone. Thus, the resulting interpretation error rates for ESBL-producing *E. coli* were significantly higher, ranging from 2.4% and 1.3% minor errors for cefotaxime and ceftriaxone to 5.4% major and very major errors for amoxicillin/clavulanic acid (Table 1).

*K. pneumoniae* and *E. cloacae*, which naturally produce chromosomal beta-lactamases, showed lower median wild-type diameters than *E. coli*, resulting in higher expected interpretation errors rates as the same susceptible CBP applies to all *Enterobacteriaceae* species (Table 1). Amoxicillin/clavulanic acid,
piperacillin/tazobactam, and ceftazidime median zone diameters, were 20 mm, 22 mm, and 26 mm for wild-type *K. pneumoniae* (harboring a chromosomal SHV-type beta-lactamase), but 25 mm, 27 mm, and 28 mm for wild-type *E. coli*, respectively. Consequently, corresponding distances of median *K. pneumoniae* diameters from CBPs were lower than those for *E. coli*, leading to higher error rates (0.8%, 9.2%, and 1.7% versus < 0.1%, < 0.1%, and 0.3%, for amoxicillin/clavulanic acid, piperacillin/tazobactam, and ceftazidime, respectively Table 1). The highest wild-type error rates were found for *E. cloacae*, which harbors a chromosomal AmpC type beta-lactamase. *E. cloacae* wild-type error rates ranged from 1.6% for cefepime to 8.6% for ceftazidime (Table 1).

Categorization error rates were significantly higher for some resistotypes as compared to the wild-type of the same species (Table 3.). The relative risk of interpretation errors for the *E. coli* ESBL resistotype as compared to the *E. coli* wild-type significantly increased for ceftriaxone (13-fold, \( p = 0.002 \)), piperacillin/tazobactam (> 107-fold, \( p < 0.001 \)), and amoxicillin/clavulanic acid (> 54-fold, \( p < 0.001 \)). Similarly, the interpretation error risk was significantly increased for *E. coli* expressing an AmpC beta-lactamase (e.g. > 154-fold increase for piperacillin/tazobactam, \( p < 0.001 \), Table 3). In addition, the *K. pneumoniae* and *E. cloacae* ESBL and AmpC resistotypes showed a significant increase in expected interpretation errors as compared to their wild-type populations (Table 3).

For *K. pneumoniae*, the wild-type population accounted for the majority of errors regarding all beta-lactams except for cefotaxime (Table 1). A comparable situation was found for *E. cloacae* (Table 1). Total errors for *Enterobacteriaceae* species (i.e. the wild-type plus all resistotypes) were, to a considerable extent, dependent on the number of interpretation errors of non-wild type populations (Table 1). For instance, in *E. coli* the ESBL resistotype was the primary source of interpretation errors for all
cephalosporins and ertapenem (47% to 82% of all errors, respectively, Table 1). Errors for amoxicillin/clavulanic acid and piperacillin/tazobactam in *E. coli* were almost exclusively due to the BSBL and ESBL resistotypes (79% and 20% of all errors, respectively, Table 1).

Interestingly, ESBL and AmpC resistotypes interpretation error rates were also significantly increased for ertapenem (Table 1), despite the fact that ertapenem is considered a therapy of choice for ESBL and AmpC producers (32). For *E. coli*, 79% of all expected ertapenem interpretation errors were caused by the ESBL resistotype (Table 1). Expected error rates for ertapenem ranged from < 0.1% for wild-type *E. coli* to > 7% for the ESBL and AmpC resistotypes, for wild-type *K. pneumoniae* (natural BSBL producer) and *E. cloacae* (natural AmpC producer) ertapenem expected error rates were 2.5 % and 6 %, respectively (Table 1).
The suitability of amoxicillin/clavulanic acid, piperacillin/tazobactam, or newer cephalosporins for the treatment of ESBL-producing *Enterobacteriaceae* infections is still controversial. Clinical data for treatment outcome reveal an ambiguous picture (19, 33-35). Until 2009 CLSI and EUCAST recommended to report *in vitro* susceptible and intermediate AST results of penicillins, cephalosporins, and monobactams for ESBL producers either as resistant (CLSI) or to modify interpretation from susceptible to intermediate and from intermediate to resistant (EUCAST) (36, 37). Such editing for beta-lactams and ESBL-producing isolates (i.e. “interpreative reading”) has been abandoned (3, 38). However, EUCAST expert rules contain a warning on uncertain therapeutic outcome for ESBL-producing isolates and inhibitor combinations (rule 9.1) or newer cephalosporins, and monotherapy of *Enterobacter* spp. with newer cephalosporins is discouraged (rule 9.2) (13). The new CLSI and EUCAST “report as found” strategy for ESBL producers and beta-lactams emphasizes the role of inhibition zone diameter or MIC value relative to a CBP as the single parameter for clinical categorization, i.e. the prediction for the clinical outcome.

In a previous work we have demonstrated that CBP setting and measurement precision significantly influence the rate of reporting errors for individual species/drug combinations (7). Accepted error rates in AST classification systems are < 5% for minor and < 1% for major and very major errors (2). Such error rates for species/drug combinations are calculated from mixed wild-type and non-wild-type populations, so that they represent mean error rates for all isolates of an individual species. Species are, however, not homogeneous, but rather represent sub-populations of wild-type isolates and one or more resistotypes. Thus, average error rates calculated from species distributions represent a statistical mean, but do not apply to an individual isolate.
Following the revised CLSI/EUCAST reporting strategy we analyzed reporting error probabilities for various resistotypes. Subpopulations carrying ESBL, BSBL, or AmpC beta-lactamases displayed significantly different probabilities of erroneous clinical categorization (Table 1). Most importantly, the categorization average reliability was significantly lower for non-wild type populations, in particular ESBL and AmpC-producing isolates (Table 3), which accounted for the majority of expected errors (Table 1). This may explain, in part, ambiguous reports on the therapeutic applicability of certain resistotype/drug combinations, e.g. ESBL positive isolates and amoxicillin/clavulanic acid and piperacillin/tazobactam (21, 39).

In its 2014 update of AST guidelines EUCAST altered the amoxicillin/clavulanic acid susceptible/resistant CBP to $\geq 16 / < 16$ mm for uncomplicated urinary tract infections, and $\geq 19$ mm $/ < 19$ mm for all other cases (40). This CBP modification hardly changes the picture of expected errors, e.g. applying the new CBPs, expected error rates for ESBL and *E. coli* would be at 5.3 % for urinary tract isolates, and at 6 % for other isolates versus 5.4 % for all isolates applying the former uniform CBP of $\geq 17$ mm $/ < 17$ mm. This example indicates that the problem of resistotype dependent reporting errors cannot be solved by CBP changes alone. This study analysed data from disk diffusion testing. However, by logical consequence, the problem of interpretation errors and overlapping or adjacent wild-type and non-wild type populations may be extrapolated also to MIC-based methods as disk diffusion diameters and MICs correlate.

This study found only minor errors for drug/species combinations with an intermediate zone (see Table 1). However, minor errors may nonetheless influence clinical decision making as they will either prompt the selection of another drug limiting therapeutic options (susceptible to intermediate error), or prevent the use of high-dose or combination therapy (intermediate to susceptible). The majority of minor errors expected were “susceptible to intermediate”, especially for the *E. coli* wild-type, the
BSBL-producing *E. coli* and the *K. pneumoniae* wild-type (Table 1). The ESBL genotypes showed a balanced proportion of susceptible/intermediate and intermediate/resistant minor errors, whereas AmpC-producing isolates tended to produce more “intermediate to resistant” and “resistant to intermediate errors”. These findings are consistent with the distance of the median diameters of the resistotypes in relation to the susceptible/intermediate and intermediate/resistant CBPs.

The CBPs of the drugs analyzed in this study are identical for all *Enterobacteriaceae* (22). However, many authors point out the importance of setting species-specific CBPs to avoid erroneous AST reports due to different diameter/MIC distributions of species (2, 7, 41). Our results reinforce these statements as the same CBPs applied to wild-type *E. coli*, *K. pneumoniae*, and *E. cloacae* lead to significantly different error rates for most beta-lactams (Tables 1 and 3). Error rates for wild-type *E. coli* were close to 0%, whereas those for wild-type *K. pneumoniae* and *E. cloacae* reached 9.2% (*K. pneumoniae* with piperacillin/tazobactam). Interpretation error rates were dependent on resistotypes rather than being species-related, i.e. error rates in species with natural resistance mechanisms were analogous to those of species with acquired similar resistance mechanisms, e.g. *K. pneumoniae* wild-type harboring an SHV-1 beta-lactamase (BSBL-type) and the *E. coli* BSBL resistotype show similar error rates (Tables 1 and 3). Furthermore, AST error characteristics of *K. pneumoniae* and *E. coli* AmpC resistotypes equaled those of the *E. cloacae* wild-type, which produces a chromosomal AmpC. From the perspective of antibiogram reliability, knowledge of the resistotype may, therefore, be at least as relevant as correct species identification.

Our results underline that total error rates in antibiograms for species/drug combinations depend on the prevalence of individual resistotypes, e.g. the number of ESBL-producing organisms in relation to the wild-type as illustrated by Figure 1 D. As ESBL-producers bear a higher likelihood for reporting errors, error rates will increase in
parallel to ESBL prevalence. In our study population, *E. coli* ESBL had a prevalence of 9% (Table 1), which is in line with many countries in Central Europe, and the resulting total error rate for all *E. coli* isolates with respect to amoxicillin/clavulanic acid was 2.5% (Table 1) (42). If ESBL prevalence would be as high as 60%, as reported for many regions in Asia, the total error rate would increase to 4.8% (43).

What are the practical implications of our findings for AST reporting and therapeutic decision making? To ensure an adequate therapy by improving antibiograms three possible options exist: i) for the short-term, a simple and practical way would be to pursue detection and reporting of ESBL and AmpC production (currently recommended by CLSI and EUCAST for epidemiological purposes only). Clinicians should be aware of the significantly increased risk of erroneous AST reports for isolates harboring beta-lactamases, in particular ESBL and AmpC. This also applies to drugs that are recommended for the treatment of ESBL and AmpC producers such as ertapenem; ii) AST reports may indicate MIC and diameter measurements in combination with reference ranges and laboratory measurement precision enabling clinicians to assess the reliability of antibiotic susceptibility classification, or iii) AST reports could benefit from including an indicator of interpretation reliability, e.g. percentages of probability for correct clinical classification based on CLSI and/or EUCAST CBPs. These probabilities would reflect the resistotype/drug combination-dependent AST forecast reliability for clinical outcome facilitating the selection of the most adequate drug for treatment.
Funding

This work was supported by the University of Zurich.

Acknowledgments

We are grateful to the laboratory technicians of the Institute of Medical Microbiology, University of Zurich for their dedicated help, to Forouhar Mouttet for performing proper statistical analyses, and to Reinhard Zbinden and Guido Bloemberg, for valuable discussions.

Conflict of interest

None to declare.


European Committee on Antimicrobial Susceptibility Testing. Disk diffusion manual v. 3.0.

http://www.eucast.org/eucast_susceptibility_testing/disk_diffusion_methodology
(11th June 2014, date last accessed).


   http://www.eucast.org/expert_rules (11th June 2014, date last accessed)


   http://www.eucast.org/clinical_breakpoints/ (11th June 2014, date last accessed).


<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>S/&lt;R: 17/17 mm</th>
<th>≥S/&lt;R: 21/19 mm</th>
<th>≥S/&lt;R: 20/17 mm</th>
<th>≥S/&lt;R: 23/20 mm</th>
<th>≥S/&lt;R: 24/21 mm</th>
<th>&quot; ≥S/&lt;R: 25/22 mm</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amoxicillin/clavulanic acid</td>
<td>2440</td>
<td>385</td>
<td>66</td>
<td>799</td>
<td>1790</td>
<td>384</td>
<td>799</td>
</tr>
<tr>
<td>Ceftazidime</td>
<td>660</td>
<td>115</td>
<td>17</td>
<td>2448</td>
<td>5244</td>
<td>911</td>
<td>5244</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>58</td>
<td>10</td>
<td>1</td>
<td>79</td>
<td>237</td>
<td>42</td>
<td>237</td>
</tr>
<tr>
<td>Cefepime</td>
<td>23</td>
<td>4</td>
<td>0</td>
<td>66</td>
<td>220</td>
<td>38</td>
<td>220</td>
</tr>
</tbody>
</table>

| Proportion of errors (%) | 0 79 36 1 100 90 8 2 100 83 17 0 100 34 66 0 100 78 19 2 100 |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| All errors (%)           | 6 0 83 21 1 105 0 77 41 10 128 2 20 32 6 60 0 0 9 5 14 0 2 5 3 11 0 5 32 2 39 0 2 27 5 34 |
| Proportion of errors (%) | 50 4 46 100 80 0 20 100 68 2 30 100 52 4 44 100 32 7 61 100 58 3 39 100 |

| All errors (%)           | 63 36 1 100 90 8 2 100 83 17 0 100 34 66 0 100 78 19 2 100 |
| Proportion of errors (%) | 0 79 36 1 100 90 8 2 100 83 17 0 100 34 66 0 100 78 19 2 100 |

| All errors (%)           | 63 36 1 100 90 8 2 100 83 17 0 100 34 66 0 100 78 19 2 100 |
| Proportion of errors (%) | 0 79 36 1 100 90 8 2 100 83 17 0 100 34 66 0 100 78 19 2 100 |

| All errors (%)           | 63 36 1 100 90 8 2 100 83 17 0 100 34 66 0 100 78 19 2 100 |
| Proportion of errors (%) | 0 79 36 1 100 90 8 2 100 83 17 0 100 34 66 0 100 78 19 2 100 |
Definition of error type: minor errors, erratic assignment to adjacent interpretative categories (susceptible to intermediate, intermediate to susceptible, intermediate to resistant, resistant to intermediate); major errors, erroneous categorisation of true-susceptible isolates as resistant; very major errors, erroneous categorisation of true-resistant isolates as susceptible.

1) For *E. cloacae*, which carries a chromosomal AmpC (wild-type, ampC not derepressed), this column indicates isolates with derepressed ampC.

2) Wild-type *E. coli* do not express a beta-lactamase. Wild-type *K. pneumoniae* carry a chromosomal BSBL; results are shown in the "wild-type" column. The *E. cloacae* wild-type carries a chromosomal AmpC.

NR, natural resistance; CBP, clinical breakpoint; BSBL, broad spectrum beta-lactamase; ESBL, extended spectrum beta-lactamase.
Table 2: Measurement precision (in mm) of disk diffusion testing displayed as median diameter values and one-fold standard deviations of 100 repeat measurements for clinical wild-type, BSBL, ESBL, and AmpC producing *E. coli* strains and 120 repeat measurements of *E. coli* ATCC 25922

<table>
<thead>
<tr>
<th>Isolate number/statistical parameters</th>
<th>Genotype</th>
<th>Amoxicillin/ clavulanic acid</th>
<th>Piperacillin/ tazobactam</th>
<th>Cefazidime</th>
<th>Cefotaxime</th>
<th>Ceftriaxone</th>
<th>Cefepime</th>
<th>Ertapenem</th>
<th>Average standard deviation isolates</th>
<th>Average standard deviation resistotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>263032</td>
<td>wild-type</td>
<td>Median diameter: 39</td>
<td>34</td>
<td>38</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 2.2</td>
<td>1.2</td>
<td>1.6</td>
<td>1.5</td>
<td>0.9</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
</tr>
<tr>
<td>47660</td>
<td>wild-type</td>
<td>Median diameter: 27</td>
<td>25</td>
<td>28</td>
<td>34</td>
<td>31</td>
<td>54</td>
<td>32</td>
<td>32</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.4</td>
<td>1.0</td>
<td>1.2</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td><em>E. coli</em> ATCC 25922</td>
<td>wild-type</td>
<td>Median diameter: 22</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>30</td>
<td>32</td>
<td>33</td>
<td>33</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 1.4</td>
<td>1.6</td>
<td>1.3</td>
<td>1.3</td>
<td>1.2</td>
<td>1.3</td>
<td>1.8</td>
<td>1.8</td>
<td>1.4</td>
</tr>
<tr>
<td>249330</td>
<td>Wild-type</td>
<td>Median diameter: 20</td>
<td>25</td>
<td>26</td>
<td>33</td>
<td>30</td>
<td>33</td>
<td>32</td>
<td>32</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 0.6</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.4</td>
<td>1.2</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>36619</td>
<td>BSBL (TEM-1)</td>
<td>Median diameter: 19</td>
<td>26</td>
<td>32</td>
<td>37</td>
<td>34</td>
<td>37</td>
<td>34</td>
<td>34</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 1.0</td>
<td>1.1</td>
<td>1.1</td>
<td>1.3</td>
<td>1.2</td>
<td>1.4</td>
<td>1.4</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>261313</td>
<td>BSBL (TEM-1)</td>
<td>Median diameter: 14</td>
<td>18</td>
<td>24</td>
<td>28</td>
<td>29</td>
<td>28</td>
<td>29</td>
<td>29</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 0.8</td>
<td>0.7</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.1</td>
<td>1.3</td>
<td>1.0</td>
</tr>
<tr>
<td>264478</td>
<td>ESBL (CTX-M VI)</td>
<td>Median diameter: 22</td>
<td>26</td>
<td>20</td>
<td>24</td>
<td>11</td>
<td>24</td>
<td>30</td>
<td>29</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 0.8</td>
<td>1.0</td>
<td>1.0</td>
<td>2.2</td>
<td>1.0</td>
<td>1.5</td>
<td>1.5</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>251055</td>
<td>ESBL (CTX-M I)</td>
<td>Median diameter: 17</td>
<td>23</td>
<td>10</td>
<td>16</td>
<td>6</td>
<td>16</td>
<td>26</td>
<td>26</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 0.9</td>
<td>0.7</td>
<td>0.9</td>
<td>0.0</td>
<td>0.6</td>
<td>0.8</td>
<td>1.1</td>
<td>1.1</td>
<td>0.7</td>
</tr>
<tr>
<td>100446</td>
<td>ESBL (CTX-M I)</td>
<td>Median diameter: 11</td>
<td>18</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>12</td>
<td>23</td>
<td>23</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 1.1</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
<td>1.4</td>
<td>1.4</td>
<td>0.7</td>
</tr>
<tr>
<td>67886</td>
<td>AmpC (promoter mutation)</td>
<td>Median diameter: 10</td>
<td>24</td>
<td>25</td>
<td>35</td>
<td>30</td>
<td>55</td>
<td>31</td>
<td>31</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard deviation: 0.7</td>
<td>0.9</td>
<td>0.9</td>
<td>1.1</td>
<td>1.0</td>
<td>1.3</td>
<td>1.3</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Reading precision of the Sirscan instrument was determined by 100 repeat measurements of the same plate.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotic</td>
<td>E. coli (wild-type)</td>
<td>E. coli (BSBL)</td>
<td>E. coli (ESBL)</td>
<td>E. coli (AmpC)</td>
<td>K. pneumoniae (wild-type)</td>
<td>K. pneumoniae (BSBL)</td>
<td>K. pneumoniae (ESBL)</td>
<td>K. pneumoniae (AmpC)</td>
<td>E. cloacae (wild-type)</td>
<td>E. cloacae (BSBL)</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>----------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Amoxicillin/clavulanic acid</td>
<td>&gt;34 &gt;54</td>
<td>&gt;12</td>
<td>&gt;43 &gt;107</td>
<td>&gt;154</td>
<td>&gt;24 &gt;70</td>
<td>&gt;24 &gt;70</td>
<td>&gt;24 &gt;70</td>
<td>&gt;24 &gt;70</td>
<td>&gt;24 &gt;70</td>
<td>&gt;24 &gt;70</td>
</tr>
<tr>
<td>Piperacillin/tazobactam</td>
<td>4 28 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3 0</td>
<td>3 0</td>
<td>3 0</td>
<td>3 0</td>
<td>3 0</td>
</tr>
<tr>
<td>Ceftazidime</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cefotaxime</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cefepime</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ertapenem</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NR: natural resistance; BSBL: broad spectrum beta-lactamase; ESBL: extended spectrum beta-lactamase.
Figure 1: Schematic representation of variables influencing interpretation error rates
Figure 2: Error probabilities, diameter distributions, and distance to clinical breakpoints for *E. coli* and amoxicillin/clavulanic acid.
Figure Legends

Figure 1:
Variables influencing interpretation errors: technical measurement precision (A), population distributions (B), distance of a populations median diameter from the next CBP (C), and relative resistotype prevalence (D1/2): A: Higher measurement precision, i.e. a lower standard deviation will result in a smaller area at risk for errors around the CBP (blue dotted curve) as compared to lower measurement precision (red dotted curve). B: Higher population variation (red bars) will result in higher error risk as compared to lower population variation (blue bars) as more isolates are located in the area at risk for errors (dotted black curve). C: The closer a population to the CBP the more isolates are at risk for errors. CBP 1 (blue) will result in fewer errors than CBP 2 (red). D1 and D2: Influence of the prevalence of sub-populations/resistotypes on error rates, ESBL-producing organisms (red bars) and wild-type (blue bars). Black vertical line: EUCAST 2010-2013 clinical breakpoint for amoxicillin-clavulanic acid (≥17 mm/ < 17 mm susceptible/resistant, respectively). Dotted curve indicates the error probability around the clinical EUCAST breakpoint. D1: E. coli ESBL prevalence 9%, E. coli wild-type prevalence 32%, other resistotypes (BSBL, AmpC, total prevalence 59%) are not shown. Total error rate for all E. coli with respect to amoxicillin/clavulanic 2.5% (see Table 1). D2: ESBL prevalence modeled at 60%, wild-type prevalence at 1% (other resistotypes, i.e. BSBL, AmpC, total prevalence 39% are not shown). The total error rate increased to 4.8%.
Zone diameter distributions of amoxicillin/clavulanic acid (grey bars) for (A) *E. coli* wild-type, (B) BSBL resistotype, (C) ESBL resistotype, and (D) AmpC resistotype. Median diameter values are indicated by vertical red lines. The error probability around the clinical EUCAST breakpoint (≥17 mm, susceptible; <17 mm, resistant, vertical black lines) corresponding to methodological imprecision (Table 2) is indicated by the dotted curve. The overlapping area (black shaded) of error probability and moving diameter average (solid line) is proportional to the cumulated relative error risk, which is dependent on the distance of median diameters CBP (double-headed arrows). (A) Wild-type *E. coli* show only a small overlap of curves (low error risk). The overlap of curves increases for the BSBL resistotype (B). The overlap of curves is maximal for the ESBL resistotype (C), and decreases for the AmpC resistotype (D) in agreement with the relative distance of the populations mean diameter to the clinical breakpoint. Note that the AmpC population is not unimodal, but shows two peaks at 6 mm and 10 mm. The according subpopulations are most probably related to plasmid encoded AmpC that show lower mean diameters, and ampC promoter/attenuator mutants that usually display higher mean diameters (24).