A new rapid resazurin-based microdilution assay for antimicrobial susceptibility testing of *Neisseria gonorrhoeae*

Commented [UMULI1]: I have optimised also the format for JAC.

Sunniva Förster[1,2,3,4]*, Valentino Desilvestro[5], Lucy Hathaway[3], Nicola Low[1], Christian Althaus[1] and Magnus Unemo[2]

[1]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland; [2]WHO Collaborating Centre for Gonorrhoea and other STIs, Örebro University, Örebro, Sweden; [3]Institute for Infectious Diseases, University of Bern, Bern, Switzerland; [4]Graduate School for Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland; [5]World Trade Institute (WTI), University of Bern, Bern, Switzerland

*Corresponding author. Institute of Social and Preventive Medicine, University of Bern, Finkelhubelweg 11, 3012, Bern, Switzerland. Tel.: +41 31 631 56 97; E-mail: sunniva.foerster@ifik.unibe.ch

Running title: Microdilution susceptibility assay for gonococci

Word count: Synopsis: 300 words, Main text: 19091 words

Objectives: Rapid, cost-effective and objective methods for antimicrobial susceptibility testing of *Neisseria gonorrhoeae* would greatly enhance surveillance of antimicrobial resistance. Etest, disc diffusion or agar dilution methods are subjective, laborious for large-scale testing, and take 16-20 hours. We developed a rapid broth microdilution assay using resazurin (blue), which is converted to resorufin (pink fluorescence), to analyse dose-response curves for antimicrobials in gonococcal isolates.

Methods: The resazurin-based broth microdilution assay was established using the 2008 WHO gonococcal reference strains (n=8) and the antimicrobials ceftriaxone, cefixime, azithromycin, spectinomycin, ciprofloxacin, gentamicin, tetracycline, and penicillin G. Training data including 84 blinded gonococcal strains were used to develop a regression model for estimating the MIC after six hours incubation time. The assay was finally validated with 40 blinded gonococcal strains.

Results: Results were obtained in approximately 7.5 hours. The EC_{50} of the dose-response curves correlates linearly with Etest MIC values (R^2 = 0.87). Only one very major error was found for a resistant ceftriaxone strain misclassified as susceptible. Minor errors resulting from misclassifications of intermediary resistant strains were found for 9% of the samples. Major errors occurred for ceftriaxone (4.8%), cefixime (3.5%), azithromycin (0.6%) and tetracycline (0.2%). Overall the sensitivity of the assay was 97.13% (CI: 95.22-98.42) and the specificity 79.27 % (CI: 74.84-83.23).

Conclusions: A rapid, objective, high-throughput, quantitative and cost-effective broth microdilution assay was established for gonococci. For use in routine diagnostics without confirmatory testing, the specificity might remain suboptimal for ceftriaxone and cefixime. However, the assay can be an effective low-cost method to evaluate novel antimicrobials, for high throughput screenings, and expands the currently available methodologies for surveillance of antimicrobial resistance in gonococci.

Commented [UMULI2]: resorufin?

Commented [s3R2]: We should stay consistent with the title and other literature about resazurin. The fluorescence is due to resorufin but the curves are based on resazurin and resorufin mixtures in response to every concentrations so I think resazurin assay is an appropriate simplification.

Commented [s4]: I decided to report errors as a fraction of all data, number since it sounds better like this and the detailed numbers can be found in figure and main text.

Commented [s5]: I changed the calculation of this and assume now that intermediary resistant strains are resistant (explained in the methods section) to be able to account for all data in a simple way.

Keywords: Gonorrhoea, antimicrobial resistance, resazurin, broth microdilution, minimum inhibitory concentration, dose-response curve

Introduction

Neisseria gonorrhoeae is a very fastidious bacterium that causes the sexually transmitted infection gonorrhoea. Gonorrhoea is a public health concern globally,[1,2] and *N. gonorrhoeae* has developed resistance to all antimicrobials introduced for treatment.[3] Accordingly, enhanced surveillance of antimicrobial susceptibility in *N. gonorrhoeae* is imperative globally.[1] Ideally, this surveillance should be performed using methods determining the MICs of relevant antimicrobials. MIC-based methods are also valuable to directly inform treatment after laboratory results are available and evaluate *in vitro* efficacy of novel antimicrobials.

Due to the lack of any appropriate broth medium, MIC-based susceptibility testing of *N. gonorrhoeae* has been limited to disk diffusion, Etest and agar dilution method (gold standard). Essential agreement with the agar dilution method is defined as ±1 doubling dilution and should ideally be above 90% for diagnostic purposes where the same resistance breakpoints are applied.[4] Etest has shown excellent agreement with the agar dilution method in many settings.[4–7] However, discordant results have been found particularly when different growth media were used.[8] A multicentre study revealed that the intra method agreement was above 70%. Categorical agreement between Etest and agar dilution was ≥ 88% but was very poor for disk diffusion.[9] Unfortunately, these methods are relatively slow (~24 hours), subjective, require expertise, and/or are expensive. Faster methods that allow results to be obtained on the same day have been developed in the past for other bacteria[10,11], but are not available for *N. gonorrhoeae*.

For many bacterial species, broth microdilution is the reference method due to accuracy, low costs and high versatility.[12,13] Several attempts have been made to develop a broth microdilution method also for *N. gonorrhoeae* but none of these have been particularly accurate

Commented [UMULI6]: 1.Unemo M, Shafer WM. Antimicrobial resistance in *Neisseria gonorrhoeae* in the 21st century: past, evolution, and future. *Clin Microbiol Rev* 2014; **27**: 587-613.

Commented [UMULI7]: Renumber all references last!

and suitable for routine use.[14–16] It is difficult to synchronize the growth of different *N. gonorrhoeae* strains and effects such as autolysis occur when the bacteria enter the stationary phase.[17–19] Chemically defined Graver-Wade (GW) broth[20] supports the growth of phylogenetically diverse auxotypes and clinical isolates and might be a suitable medium for susceptibility testing.[21,22]

Unfortunately, MIC values based on doubling dilution series are left, interval, or right censored discrete data which makes error statistics challenging.[23] The potency of drugs in pharmacology is frequently measured with dose-response curves, as this allows the estimation of the effective concentration (EC) at a specified response level. Furthermore, EC values on a continuous scale take the variability of the data into account by calculating model based confidence intervals (CIs). In the field of toxicology the lower confidence interval is defined as non-toxic concentration. This so called benchmark dose approach (BMD) has largely replaced methods that rely on dense dose spacing because of its statistical superiority and reduction of animal use.[24–27] Furthermore, the slope of the pharmacodynamic curve can provide additional valuable information on the compounds being tested.[28] The Hill coefficient can provide information about the pharmacodynamic properties of an antimicrobial and has been used in modelling studies of single and dual antimicrobial effects.[21,22,29–31] However, the interpretation and significance of the Hill slope has been unclear in previous studies and laborious colony counting limited these studies to few strains.

The biological response to a compound can be measured using different readouts. Traditionally the MIC is defined as the concentration of an antimicrobial that inhibits visual growth but methods to quantify the number of bacterial cells more objectively are available. Measuring the optical density (at e.g. OD_{600} or OD_{450}), resazurin (Alamar blue), 3-(4,5-dimethyethiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), luciferase (ATP levels) and lactate dehydrogenase are widespread methods where readouts correlate with the number of

cells.[32] Resazurin is a blue dye that is converted to pink fluorescent resorufin in the presence of metabolically active cells.[33,34] Unlike optical density, a measure of growth inhibition, it reflects the viability of cells and is potentially suitable for time-kill assays. Resazurin has an excellent signal to noise ratio and has been used previously in screenings for toxicity testing[35], high throughput applications[36],biofilm screening[37] and MIC testing.[33,38–40]

The aim of the present study was to develop a resazurin-based broth microdilution assay for antimicrobial susceptibility testing of *N. gonorrhoeae* that is rapid, objective, scalable, quantitative and inexpensive. Three datasets were generated in this study. A panel of reference strains (n=8) was studied to ensure the reproducibility of the assay and compare multiple measurement endpoints in a time-course between 0-15 hours. Training data consisting of 84 strains were used to develop a predictive model for estimating the MIC from dose-response curves. Finally, a panel of forty strains with blinded MICs was used to validate the prediction.

Commented [UMULI8]: resorufin?

Material and methods

Bacterial strains, culture and broth microdilution assay

The 4 gonococcal strains was used to develop a regression model for estimating the MIC after six hours incubation time (one replicate). The assay was finally validated with 40 blinded gonococcal strains (one replicate). The blinded strains were selected to represent a wide variety of antibiograms. The strains were preserved in glycerol stocks at -80°C. All strains were subsequently cultured on Chocolate agar PolyViteX (Biomerieux, Marcy l'Etoile, France) at 37°C in a humid 5% CO_2-enriched atmosphere for 16-18 hours and then sub-cultured once for 16 hours. A McFarland standard of 0.5 was prepared for each strain and 1 mL further diluted to approximately $1x10^7$ CFU/mL in 15 mL heated (37°C) GW broth.[20] A volume of 90 µL of this suspension was added to 96-well round bottom microtiter plates (360 µL wells) with each well containing 10 µL of a previously prepared dilution series. Dilution series of the

antimicrobials were prepared in GW medium. Positive control (GW medium containing 1% TritonX-100) and negative control (10 µL GW medium) were added to the first and last well, respectively. The plates were incubated for 6 hours at 37°C, humid 5% CO_2-enriched atmosphere.

Resazurin readouts

Resazurin powder (Sigma Aldrich, China) was diluted in PBS (pH 7.4) to a final concentration of 0.1 mg/mL. It was ensured that the pH of the highest antimicrobial concentration was neutral in all samples to avoid artefacts. After incubation of the broth microdilution plates, 50 µL of the dye was added to each well and mixed using an electronic multichannel dispenser. The plates were incubated for 75 minutes at 37°C. Fluorescence was then measured at 560 and 590 nm excitation in a plate reader (Varioskan Flash, Thermo Scientific).

Etest MIC

The Etest MICs (bioMérieux) were determined in accordance with the manufacturer´s instructions, on GCRAP agar plates (3.6% Difco GC Medium Base agar [BD, Diagnostics, Sparks, MD, USA] supplemented with 1% haemoglobin [BD, Diagnostics] and 1% IsoVitalex [BD, Diagnostics]).

Dose-response modelling

The antimicrobial effect on the different bacterial strains was quantified with dose-response curves. Firstly background fluorescence resulting from dead bacteria in the positive control wells was subtracted. A sigmoidal model was fit to the viability data of each antimicrobial-strain combination[41,42]:

$$f(x) = u + \frac{l-u}{1+e^{H(x-EC_{50})}} \quad \text{(Eq. 1).}$$

Commented [UMULI9]: GW medium including 1% TritonX-100?

Commented [s10R9]: yes

Commented [s11]: I rewrote this paragraph entirely. I changed from the log logistic model to a logistic model with log transformed dose values. The output is not exactly the same but similar and makes the bootstrapping procedure easier.

Four parameters describe the curve, the lower model asymptote *l*, the upper model asymptote *u*, the Hill coefficient *H* (slope) and the EC_{50}. Next, the data were divided by *u* to normalize all dose response curves to 100% viability. Hill coefficient differences across antimicrobials were tested with pairwise t-test. Hierarchical complete linkage clustering was used to compare antimicrobial similarity.[43]

Samples were considered to be above the limit of detection, and therefore categorized as resistant, if the antibiotic, at its highest concentration, reduced viability by less than 50%. This was the case for 6 samples in the training data (n=672), 9 samples in the validation data (n=320) and 3 reference strain samples (n=192). Excluding Etest MICs that were above or below limit of detection and samples where EUCAST resistance breakpoints are not available to date (gentamicin) resulted in 571 evaluable samples in the training data, 269 samples in the validation data and 137 in the reference strain data.

The relationship between EC_{50} and Etest was analysed for the training data by log transforming both values and fitting a linear regression:

$$\log(Etest) = \alpha + \beta \log(EC_{50}) + \varepsilon \quad \text{(Eq. 2).}$$

Slope and intercept of this regression were used to predict the MIC from the EC_{50} values. Confidence intervals (CI) for each predicted MIC were calculated using bootstrapping in order to take in account estimation error in both sigmoidal and linear regression models. The EC_{50} and its standard deviation from the sigmoidal model were used as parameters in the normal distribution used to resample 10^5 EC_{50}. Similarly, 10^5 values for α and β in Eq.2 were obtained by resampling from a two dimensional normal distribution $N_2([\alpha, \beta], \mathrm{cov}(\alpha, \beta))$. The 0.025 and 0.975 percentiles of the resulting 10^5 predicted MICs distribution consist in the 95% bootstrapped CIs. The analysis pipeline and data are available from github (https://github.com/sunnivas/ResazurinMIC).

Commented [s12]: I will make this available before submission

Essential agreement was defined as the percentage of strains with predicted MICs that do not deviate more than ±1 doubling dilution from Etest MICs. Deviations from the Etest MICs were calculated as $\log_2$ differences from the predicted MIC (840 evaluable samples for training and validation data). Reference strain data were not included to avoid bias from replicate testing of these samples.

Categorical agreement with Etest

The strains were categorized as S (susceptible), I (intermediate), and R (resistant) to each antimicrobial in accordance with the EUCAST 2016 guidelines.[44] As previously described,[45] minor errors were defined as misclassifications of intermediate strains as susceptible or resistant. Major errors were susceptible strains misclassified as resistant. Very major errors were resistant strains that were misclassified as susceptible. The EC_{50} values are read on a continuous scale, therefore nearly identical values around a resistance breakpoint (e.g. 0.125 and 0.126) can result in categorical errors. Sensitivity and specificity of the assay were calculated as previously described[46], for the resistant (true positive samples), intermediary strains (true positive samples) and susceptible strains (true negative samples).

Commented [UMULI13]: For these error definitions: Include the reference: Clinical and Laboratory Standards Institute. Development of In Vitro Susceptibility Testing Criteria and Quality Control Parameters, 2nd edn. Approved Guideline M23-A2. CLSI, Wayne, PA, USA, 2001.

Results

Dose-response modelling

The 2008 WHO reference strains (n=8) were exposed to ceftriaxone, cefixime, azithromycin, spectinomycin, ciprofloxacin, gentamicin, tetracycline, and penicillin G for a time course from 0 to 15 hours (Figure S1). After six hours, the difference between dead and viable gonococcal cells was sufficiently pronounced to fit dose-response curves to the data. For this endpoint of six hours, the coefficient of variation was calculated for the EC_{50} of three independent experiments. The coefficient of variation (CV) ranged from 1.7% to 100%, the intra-assay CV

Commented [UMULI14]: I have not got Figure S1!

was 28% (n=64) (Figure S2). Dose-response curves were gradually shifted towards higher concentrations, indicating increased potency of the antimicrobials in the intermediary resistant and resistant strains compared to susceptible strains (Figure 1). There was a clear separation of susceptible and resistant strains for ciprofloxacin and spectinomycin. For the β-lactams ceftriaxone, cefixime and penicillin G the Hill coefficients (slopes) were more heterogeneous than for the other samples. The mean of this parameter gradually increased from ceftriaxone (1.8 ± 1.7) to cefixime (2 ± 1.9), tetracycline (2.1 ± 0.87), penicillin G (2.4 ± 1.6), azithromycin (2.6 ± 1.5), ciprofloxacin (2.7 ± 1.2), spectinomycin (2.9 ± 1.7) and was highest for gentamicin (3.3 ± 1.3). A pairwise *t*-test showed that the differences between the antimicrobials were significant (p-value <0.005) when the distance between the means was larger than 0.5 (Figure S3A). Furthermore, hierarchical clustering shows a high similarity of the hill coefficient for the β-lactams ceftriaxone, cefixime and penicillin G compared to the other antimicrobials (Figure S3B).

The correlation between the Etest MICs and EC_{50} values for all antimicrobials was 0.87 (Figure 2A). Compared to the Etest values, the EC_{50} values were systematically lower with a median deviation of -1.68 doubling dilutions (Figure 2B). The regression parameter α ($\hat{\alpha} = 1.10; \mathrm{sd}_{\hat{\alpha}} = 0.048$) and β ($\hat{\beta} = 1.00; \mathrm{sd}_{\hat{\beta}} = 0.016$) of the linear log-log regression were used to predict the 840 MICs of training and validation data. The deviation of the predicted MIC from Etest followed a normal distribution with a median of -0.004, 95% of the deviations ranged between -2.28 and 4.00. Outliers can be attributed to the β-lactams penicillin G (overestimation in beta lactamase producing strains), cefixime and ceftriaxone (potentially biphasic or triphasic curves with large confidence intervals). One example for a biphasic curve was studied in detail (Figure S3)[47]. The 75% percent quartiles for the deviations were larger for azithromycin, cefixime and ceftriaxone compared to ciprofloxacin, penicillin G, spectinomycin and tetracycline (Figure 2C). The essential agreement between the Etest MICs

Commented [UMULI15]: I have not got Figure S2 either!

Commented [CA16]: What do you want to say by reporting the CoV? Do you calculate it from the linear or log values?

Commented [UMULI17]: Why have you excluded ciprofloxacin?

Commented [UMULI18]: Why have you excluded ciprofloxacin?

Commented [UMULI19]: But was the correlation not better for the 40 strains examined with the final assay? Is that correlation not what you should show?

Commented [s20R19]: I only need the correlation for the training data (84 strains). The validation data should be calculated independent from this to confirm that the prediction (shifting the distribution) works.

Commented [UMULI21]: why have you excluded gentamicin here?

and the predicted MICs was 47% for all antimicrobials, being lowest for cefixime (29%) and highest for penicillin G (61%).

Categorical agreement

The Etest and predicted MICs (n=868) were classified as susceptible, intermediate and resistant according to the EUCAST 2016 resistance breakpoints[30] (Figure 3). The sensitivity of the assay was 97.13% (CI: 95.22-98.42). One very major error (R to S), occurred for ceftriaxone (Etest MIC 0.19 mg/L vs. 0.053 mg/L). For penicillin G, spectinomycin and ciprofloxacin no major errors were identified. False positive misclassifications (S to R), i.e. major errors, occurred for tetracycline (0.2%), azithromycin (0.6%), cefixime (3.48%) and ceftriaxone (4.8%) for a total of 9% of the data. Minor errors resulting from misclassifications of intermediary resistant strains were found for 9% of the data. A high number of predicted MIC values (20%) had 95% CIs spanning two categories. The overall specificity of the assay was 79.27 % (CI: 74.84-83.23).

Discussion

The developed resazurin-based broth microdilution assay was able to discriminate between resistant and susceptible strains reliably, is faster (approximately 7.5 hours) than currently available MIC methods for *N. gonorrhoeae* and has an excellent sensitivity of 97.13% (CI: 95.22-98.42). The gold standard methods agar dilution and Etest are both based on subjective, visual readouts and are therefore limited to a relatively low throughput. Dose-response modelling allows the precise estimation of the EC_{50} of antimicrobials from a continuous scale and provides confidence intervals rather than having the precision limited by doubling dilutions. It is inherently difficult to apply resistance breakpoints that were designed for doubling dilution-based methods to dose-response curve based MICs. This was reflected by

Commented [UMULI22]: This sounds extremely bad and we need to discuss it in the discussion, i.e. essential agreement for exact MICs are low but still the accuracy in categorising into S or R is not so bad. This message we need to get a sclear as possible.

Commented [s23R22]: I checked this and found that it was a stupid bug (testing deviations before prediction). It slightly improved now for all antibiotics except cefixime (annoying double curves).

Commented [UMULI24]: Was this not also the case for spectinomycin?

Commented [s25R24]: Correct

Commented [s26]: I changed the calculation of this and assume now that intermediary resistant strains are resistant. Due to the large number of strains that have an MIC close to the breakpoint the specificity looks worse than it is....

Commented [UMULI27]: I would like to read and comment the discussion again after everyone commented.

Commented [UMULI28]: Consider to include the number of hours!

many categorical errors resulting from estimates that have CIs overlapping two SIR categories. The performance of the assay was excellent for ciprofloxacin, penicillin G and spectinomycin (no major errors) and acceptable for azithromycin (0.6% major errors) and tetracycline (0.2% major errors). For cefixime and ceftriaxone many false positive results and consequently an overestimation of resistance was measured. The complex mechanism of action of these antimicrobials is not well understood and involves several resistance determinants (*PenA, PenB, mtrR, ponA, factorX)*[48]. The correlation of EC_{50} and MIC has been previously shown to be largely influenced by different penicillin binding proteins in *Streptococcus pneumoniae*.[49] This might explain the strain dependent heterogeneity of hill coefficients (Figure S3) and dose response curves that are biphasic (Figure S4).[22,47] To address such complex effects with the simple four parameter hill function employed in this study is inaccurate and therefore contributed to the poor specificity of only 79.27 % (95% CI: 74.84-83.23). The deviation from Etest follows a normal distribution, outliers can be attributed to the β-lactams penicillin G (large overestimation in beta lactamase producing strains), cefixime and ceftriaxone (potentially biphasic or triphasic curves with large confidence intervals). The essential agreement was suboptimal and might reflect why it has been difficult to establish a broth microdilution assay for fastidious gonococci in the past. An endpoint of six hours provided only a snapshot of the antimicrobial properties and examining much more time-points, starting inocula, and very large number of strains might provide valuable data for improvements. Furthermore, obtaining significantly more data, possibly by scaling the assay to a robotic platform, would enable the regression analysis to be performed for the different antimicrobials separately and also allow fitting a biphasic model[47].

Despite these limitations, the developed rapid resazurin-based broth microdilution assay is highly objective (avoids visual subjective readout) and employs a standardized algorithm reducing operator bias, which can be especially valuable in multicentre studies. These

properties, and the low price of resazurin, are especially valuable when screening large libraries of new compounds, antimicrobials or antimicrobial combinations. Frequently, the question that needs to be answered is the potency of antimicrobials relative to each other rather than absolute numbers. The β-lactams cefixime, ceftriaxone and penicillin G displayed significantly lower Hill coefficients than the other antimicrobials. Information about this parameter is useful for research questions beyond susceptibility testing, such as combination therapy and pharmacodynamic modelling.

In summary, the developed resazurin-based broth microdilution assay is a rapid, objective, high-thoughput, quantitative and cost-effective new tool for studying *N. gonorrhoeae* in liquid culture. The Hill coefficient could be compared for a large number of strains highlighting differences between antimicrobials. The new assay opens up avenues for high-throughput synergy testing, evaluation of novel antimicrobials and surveillance of resistance.

Funding

The present study was funded through an Interdisciplinary PhD (IPhD) project from SystemsX.ch (The Swiss Initiative for Systems Biology), RaDAR-Go (RApid Diagnosis of Antibiotic Resistance in Gonorrhoea; funded by the Swiss Platform for Translational Medicine), and the Örebro County Council Research Committee and the Foundation for Medical Research at Örebro University Hospital, Sweden.

References

1. WHO. Global action plan to control the spread and impact of antimicrobial resistance in *Neisseria gonorrhoeae*. 2012. Available at: http://apps.who.int/iris/bitstream/10665/44863/1/9789241503501_eng.pdf. Accessed December 6, 2016.

2. Newman L, Rowley J, Hoorn SV, *et al*. global estimates of the prevalence and incidence of four curable sexually transmitted infections in 2012 based on systematic review and global reporting. PLOS ONE 2015; 10: e0143304.

3. Unemo M, Shafer WM. Antimicrobial resistance in *Neisseria gonorrhoeae* in the 21st century: Past, Evolution, and Future. Clin Microbiol Rev 2014; 27: 587–613.

4. Biedenbach DJ, Jones RN. Comparative assessment of Etest for testing susceptibilities of *Neisseria gonorrhoeae* to penicillin, tetracycline, ceftriaxone, cefotaxime, and ciprofloxacin: investigation using 510(k) review criteria, recommended by the Food and Drug Administration. *J Clin Microbiol* 1996; **34**: 3214–7.

5. Liu H, Taylor TH, Pettus K *et al*. Assessment of Etest as an alternative to agar dilution for antimicrobial susceptibility testing of *Neisseria gonorrhoeae*. J Clin Microbiol 2014; 52: 1435–40.

6. Singh V, Bala M, Kakran M *et al*. Comparative assessment of CDS, CLSI disc diffusion and Etest techniques for antimicrobial susceptibility testing of *Neisseria gonorrhoeae*: a 6-year study. BMJ Open 2012; 2: e000969.

7. Gose S, Kong CJ, Lee Y *et al*. Comparison of *Neisseria gonorrhoeae* MICs obtained by Etest and agar dilution for ceftriaxone, cefpodoxime, cefixime and azithromycin. J Microbiol Methods 2013; 95: 379–80.

8. Liao C-H, Lai C-C, Hsu M-S *et al*. Antimicrobial susceptibility of *Neisseria gonorrhoeae* isolates determined by the agar dilution, disk diffusion and Etest methods: comparison of results using GC agar and chocolate agar. Int J Antimicrob Agents 2010; 35: 457–60.

9. Ison CA, Martin IMC, Lowndes CM *et al*. Comparability of laboratory diagnosis and antimicrobial susceptibility testing of *Neisseria gonorrhoeae* from reference laboratories in Western Europe. J Antimicrob Chemother 2006; 58: 580–6.

10. Kelly MT, Leicester C. Evaluation of the Autoscan Walkaway system for rapid identification and susceptibility testing of gram-negative bacilli. *J Clin Microbiol* 1992; **30**: 1568–71.

11. Godsey JH, Bascomb S, Bonnette T, *et al.* Rapid antimicrobial susceptibility testing of gram-negative bacilli using Baxter MicroScan rapid fluorogenic panels and autoSCAN-W/A. *Pathol Biol (Paris)* 1991; **39**: 461–5.

12. Reller LB, Weinstein M, Jorgensen JH *et al*. Antimicrobial susceptibility testing: a review of general principles and contemporary practices. Clin Infect Dis 2009; 49: 1749–55.

13. Wiegand I, Hilpert K, Hancock REW. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. Nat Protoc 2008; 3: 163–75.

14. Takei M, Yamaguchi Y, Fukuda H *et al.* Cultivation of *Neisseria gonorrhoeae* in liquid media and determination of its in vitro susceptibilities to quinolones. J Clin Microbiol 2005; 43: 4321–7.

15. Geers TA, Donabedian AM. Comparison of broth microdilution and agar dilution for susceptibility testing of *Neisseria gonorrhoeae*. Antimicrob Agents Chemother 1989; 33: 233–4.

16. Shapiro MA, Heifetz CL, Sesnie JC. Comparison of microdilution and agar dilution procedures for testing antibiotic susceptibility of *Neisseria gonorrhoeae*. J Clin Microbiol 1984; 20: 828–30.

17. Dillard JP, Seifert HS. A peptidoglycan hydrolase similar to bacteriophage endolysins acts as an autolysin in *Neisseria gonorrhoeae*. Mol Microbiol 1997; 25: 893–901.

18. Elmros T, Burman LG, Bloom GD. Autolysis of *Neisseria gonorrhoeae*. J Bacteriol 1976; 126: 969–76.

19. Chan YA, Hackett KT, Dillard JP. The lytic transglycosylases of *Neisseria gonorrhoeae*. Microb Drug Resist 2012; 18: 271–9.

20. Wade JJ, Graver MA. A fully defined, clear and protein-free liquid medium permitting dense growth of *Neisseria gonorrhoeae* from very low inocula. FEMS Microbiol Lett 2007; 273: 35–7.

21. Foerster S, Golparian D, Jacobsson S et al. Genetic resistance determinants, in vitro time-kill curve analysis and pharmacodynamic functions for the novel topoisomerase II inhibitor ETX0914 (AZD0914) in *Neisseria gonorrhoeae*. Front Microbiol 2015; 6: 1377.

22. Foerster S, Unemo M, Hathaway LJ *et al*. Time-kill curve analysis and pharmacodynamic functions for in vitro evaluation of antimicrobials against *Neisseria gonorrhoeae*. BMC Microbiol 2016; 16: 216.

23. Kassteele J van de, Santen-Verheuvel MG van, Koedijk FDH *et al.* New statistical technique for analyzing MIC-based susceptibility data. Antimicrob Agents Chemother 2012; 56: 1557–63.

24. Slob W. Benchmark dose and the three Rs. Part I. Getting more information from the same number of animals. Crit Rev Toxicol 2014; 44: 557–67.

25. Slob W. Benchmark dose and the three Rs. Part II. Consequences for study design and animal use. Crit Rev Toxicol 2014; 44: 568–80.

26. Davis JA, Gift JS, Zhao QJ. Introduction to benchmark dose methods and U.S. EPA's benchmark dose software (BMDS) version 2.1.1. Toxicol Appl Pharmacol 2011; 254: 181–91.

27. Filipsson AF, Sand S, Nilsson J *et al*. The benchmark dose method-review of available models, and recommendations for application in health risk assessment. Crit Rev Toxicol 2003; 33: 505–42.

28. Sampah MES, Shen L, Jilek BL *et al*. Dose–response curve slope is a missing dimension in the analysis of HIV-1 drug resistance. Proc Natl Acad Sci U S A 2011; 108: 7613–8.

29. Regoes RR, Wiuff C, Zappala RM. *et al*. Pharmacodynamic functions: a multiparameter approach to the design of antibiotic treatment regimens. Antimicrob Agents Chemother 2004; 48: 3670–6.

30. Foucquier J, Guedj M. Analysis of drug combinations: current methodological landscape. *Pharmacol Res Perspect* 2015; **3**. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4492765/. Accessed November 17, 2015.

31. Yu G, Baeder DY, Regoes RR *et al*. Combination Effects of Antimicrobial Peptides. Antimicrob Agents Chemother 2016; 60: 1717–24.

32. Rampersad SN. Multiple applications of alamar blue as an indicator of metabolic function and cellular health in cell viability bioassays. Sensors 2012; 12: 12347–60.

33. Khalifa RA, Nasser MS, Gomaa AA *et al*. Resazurin microtiter assay plate method for detection of susceptibility of multidrug resistant *Mycobacterium tuberculosis* to second-line anti-tuberculous drugs. Egypt J Chest Dis Tuberc 2013; 62: 241–7.

34. Palomino J-C, Martin A, Camacho M *et al.* Resazurin microtiter assay plate: simple and inexpensive method for detection of drug resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2002; 46: 2720–2.

35. Zimmer B, Pallocca G, Dreser N, *et al.* Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery. *Arch Toxicol* 2014; **88**: 1109–26.

36. Lim KT, Zahari Z, Amanah A, *et al.* Development of resazurin-based assay in 384-well format for high throughput whole cell screening of *Trypanosoma brucei rhodesiense* strain STIB 900 for the identification of potential anti-trypanosomal agents. Exp Parasitol 2016; 162: 49–56.

37. Pettit RK, Weber CA, Pettit GR. Application of a high throughput Alamar blue biofilm susceptibility assay to *Staphylococcus aureus* biofilms. *Ann Clin Microbiol Antimicrob* 2009; **8**: 28.

38. Schmitt DM, Connolly KL, Jerse AE *et al*. Antibacterial activity of resazurin-based compounds against *Neisseria gonorrhoeae* in vitro and in vivo. Int J Antimicrob Agents 2016; 48: 367–72.

39. Elshikh M, Ahmed S, Funston S, *et al*. Resazurin-based 96-well plate microdilution method for the determination of minimum inhibitory concentration of biosurfactants. Biotechnol Lett 2016; 38: 1015–9.

40. Mann C m., Markham J l. A new method for determining the minimum inhibitory concentration of essential oils. J Appl Microbiol 1998; 84: 538–44.

41. Ritz C, Streibig J. Bioassay analysis using R. J Stat Softw 2005; 12: 1–22.

42. Ritz C, Baty F, Streibig J et al. Dose-Response analysis using R. PLOS ONE 2015; 10: e0146021.

43. Renaud Gaujoux, Cathal Seoighe (2010). A flexible R package for nonnegative matrix factorization. BMC Bioinformatics 2010, 11:367.

44. EUCAST. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. 2016.

415 45. CLSI, Wayne, PA, USA. Clinical and Laboratory Standards Institute. Development of In
416 Vitro Susceptibility Testing Criteria and Quality Control Parameters, 2nd edn. Approved
417 Guideline M23-A2. 2001. Available at:
418 http://shop.clsi.org/site/Sample_pdf/M23A3_sample.pdf. Accessed December 7, 2016.

419 46. Parikh R, Mathai A, Parikh S *et al*. Understanding and using sensitivity, specificity and
420 predictive values. Indian J Ophthalmol 2008; 56: 45–50.

421 47. Di Veroli GY, Fornari C, Goldlust I, *et al.* An automated fitting procedure and software
422 for dose-response curves with multiphasic features. Sci Rep 2015; 5: 14701.

423 48. Unemo M, Golparian D, Nicholas R *et al*. High-level Cefixime- and Ceftriaxone-resistant
424 *Neisseria gonorrhoeae* in France: Novel penA mosaic allele in a successful international
425 clone causes treatment failure. Antimicrob Agents Chemother 2012; 56: 1273–80.

426 49. Kocaoglu O, Tsui H-CT, Winkler ME *et al.* Profiling of β-Lactam selectivity for
427 penicillin-binding proteins in *Streptococcus pneumoniae* D39. Antimicrob Agents Chemother
428 2015; 59: 3548–55.

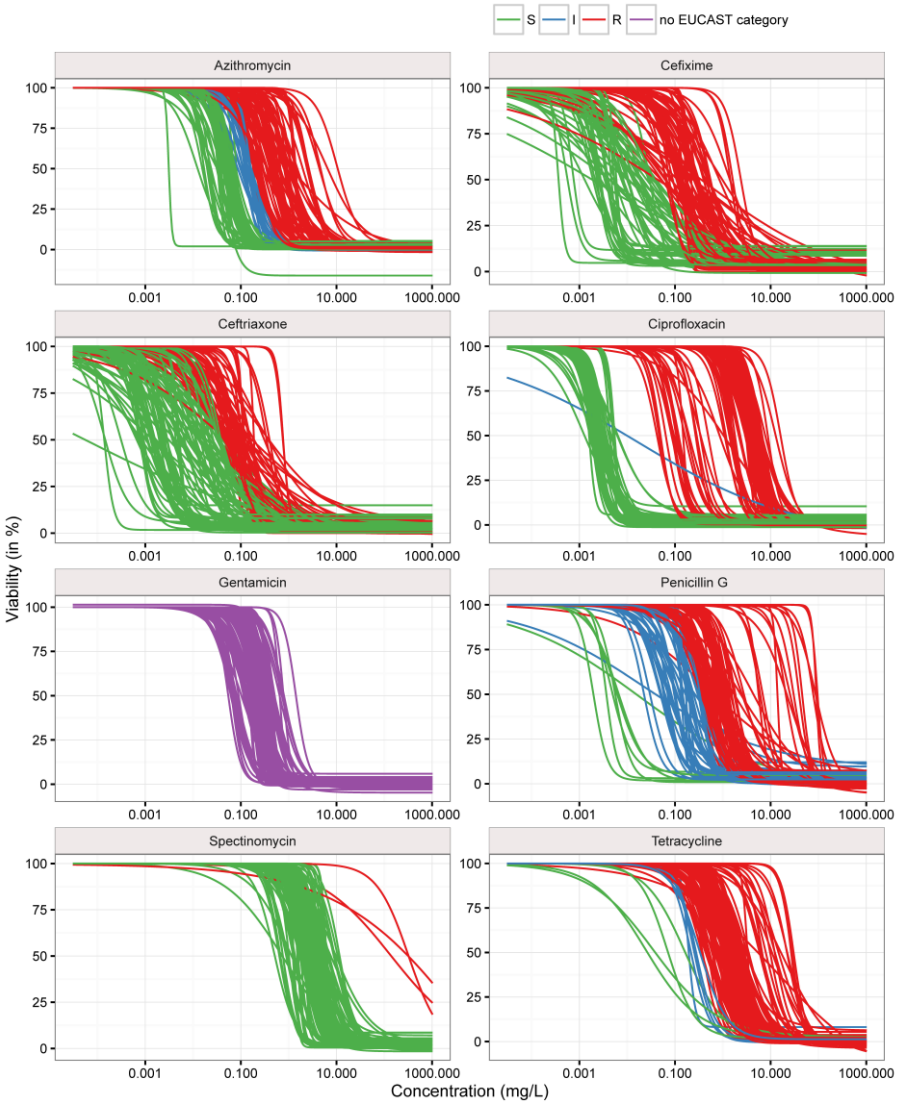

Figure 1. Potency shift of antimicrobials across different strains. Dose response curves for all strains and antimicrobials (without samples above limit of detection). Strains that were classified as susceptible according to EUCAST 2016 criteria were coloured in green, intermediary resistant strains in blue and resistant strains in red. No EUCAST criteria are defined for gentamicin (purple). The gradual shift of the potencies (EC_{50}) towards higher concentrations can be observed for all antimicrobials.

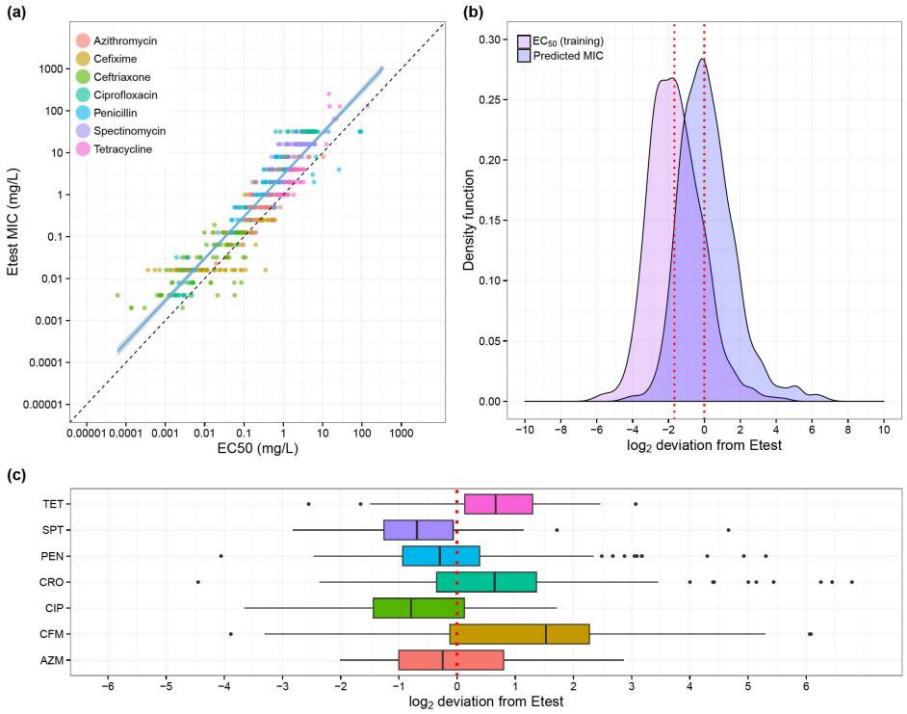

Figure 2. Correlation and deviations between the Etest and predicted MICs. (A) The correlation of Etest MIC and EC_{50} for the dataset used for developing the regression model (84 blinded strains examined) is shown for log transformed values. The Pearson's correlation coefficient for the linear regression (blue line) was 0.87 and the confidence interval highlighted in grey. Slope and intercept for a perfect correlation was drawn as dashed black line for comparison. (B) The kernel density function of the EC_{50} values the training data (n=269) drawn in red (median -1.68). The kernel density of the predicted MICs for training and validation data (n=840) is highlighted in purple (median -0.004). (C) Deviations of predicted MICs from Etest MIC per antimicrobial (n=840). The boxplots show the median and 25%-75% quartiles. The whiskers span the range from the bottom 5% to the highest 95% of the data. The essential agreement (EA) is written below the boxplots.

Commented [UMULI29]: A. Have this a total Square and not a rectangle. B. write EC50 with 50 as subscript on x-axis. C. Write Penicillin G

Commented [CA30]: Note that the values on the axes are not transformed, you just use a logarithmic scale.

Commented [s31R30]: ?? The values are log transformed in the regression and always have been log transformed. I don't understand.

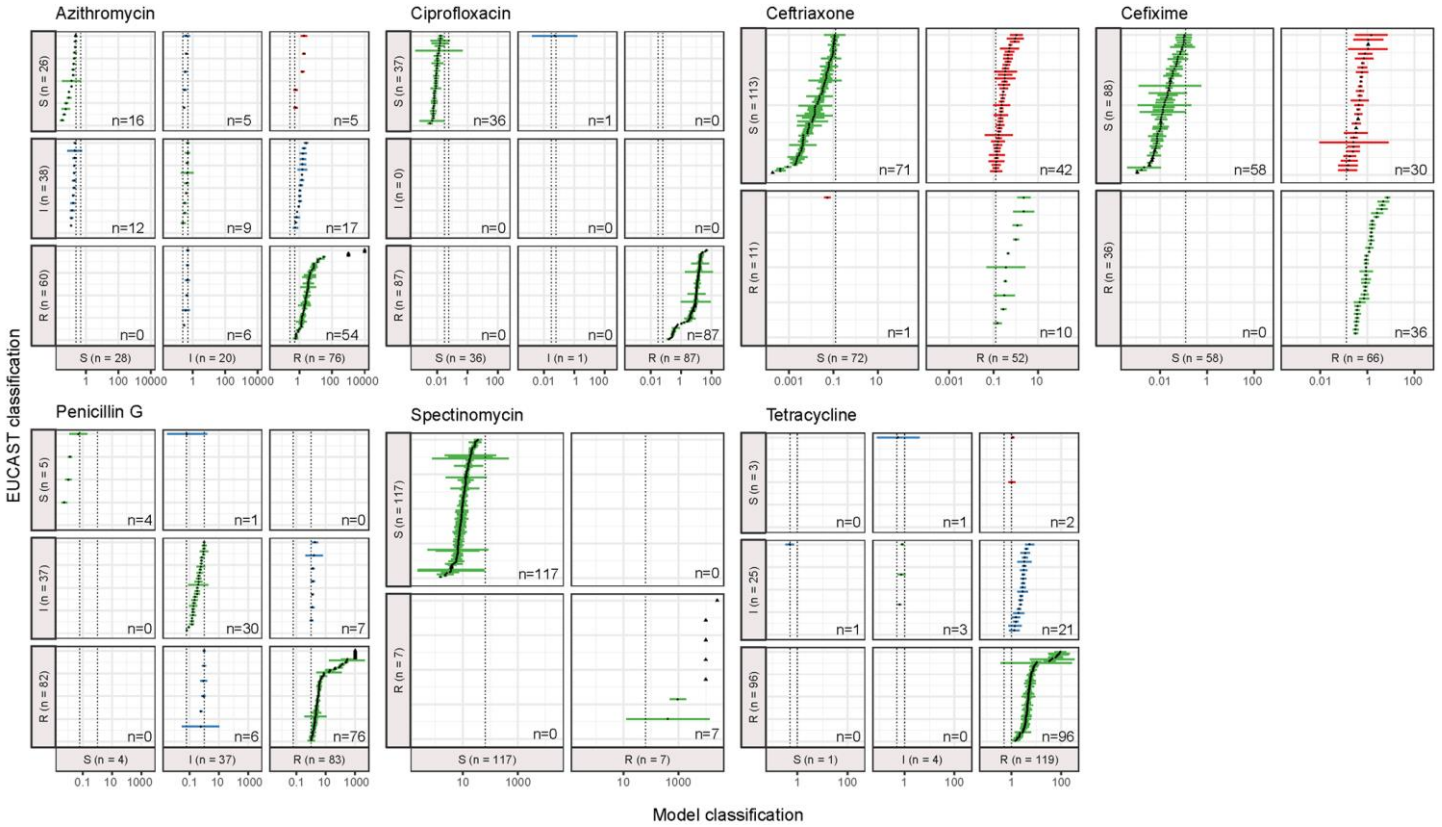

Figure 3. Contingency table with categorical errors of model predicted MICs. Etest MIC data were classified into the categories resistant (R), susceptible (S) and intermediate (I) according to the EUCAST 2016 criteria.[44] The cutoff values (mg/L) are shown as dashed black lines. Predicted

Commented [UMULI32]: A.Write Penicillin G, B. Does this not also look better with only the 40 strains for validation, i.e. when the model optimised?

Commented [s33R32]: It looks better but I would have a very bad feeling to "hide" misclassifications that come from the training dataset. At least we have statisticall meaningful numbers like that and the overall picture that ceftriaxone and cefixime are a problem will not change no matter how we show it.

MIC values (n=868) are shown as point estimate (black dots) with 95% confidence interval (colored dashes). For some Estimates no confidence interval could be calculated (limit of detection), those were drawn as triangles. Correctly classified strains are drawn in green green. One very major errors (R to S) was found for ceftriaxone (red). Major errors (S to R) and were found for ceftriaxone (n=42), cefixime (n=30), azithromycin (n=5) and tetracycline (n=2). Minor errors resulting from misclassifications of intermediate strains are drawn in blue. A high number of estimates (n=140) has confidence intervals spanning two categories.

Commented [UMULI34]: Would it be possible to distinguish these in the Figure, i.e. which will show that the very major errors are relatively few.