
Title: Relative Performance of Antimicrobial Susceptibility Assays on Clinical *Escherichia coli* Isolates from Animals

Authors: Skye Badger, Sam Abraham, Sugiyono Saputra, Darren J. Trott, John Turnidge, Tahlia Mitchell, Charles G.B. Caraguel, David Jordan

PII: S0378-1135(17)30701-0
DOI: https://doi.org/10.1016/j.vetmic.2017.12.008
Reference: VETMIC 7827

To appear in: *VETMIC*

Received date: 8-6-2017
Accepted date: 11-12-2017

Please cite this article as: Badger, Skye, Abraham, Sam, Saputra, Sugiyono, Trott, Darren J., Turnidge, John, Mitchell, Tahlia, Caraguel, Charles G.B., Jordan, David, Relative Performance of Antimicrobial Susceptibility Assays on Clinical Escherichia coli Isolates from Animals.Veterinary Microbiology https://doi.org/10.1016/j.vetmic.2017.12.008

# Relative Performance of Antimicrobial Susceptibility Assays on Clinical *Escherichia coli* Isolates from Animals

**Skye Badger[a,b], Sam Abraham[a,1], Sugiyono Saputra[a], Darren J. Trott[a,c], John Turnidge[d], Tahlia Mitchell[a, 2], Charles G.B. Caraguel[a,¶], David Jordan[a,b,e,¶, *]**

[a] School of Animal and Veterinary Sciences, The University of Adelaide, Roseworthy Campus, Mudla Wirra Rd, Roseworthy, 5371, Australia

[b] School of Veterinary and Life Sciences, Murdoch University, 90 South Street, Murdoch, Perth, Western Australia, 6150, Australia

[c] Australian Centre for Antimicrobial Resistance Ecology, The University of Adelaide, Roseworthy Campus, Mudla Wirra Rd, Roseworthy, 5371, Australia

[d] School of Biological Sciences, Department of Molecular and Cellular Biology, University of Adelaide, Adelaide, 5005, Australia

[e] Wollongbar Primary Industries Institute, NSW Department of Primary Industries, 1243 Bruxner Highway, Wollongbar, New South Wales, 2477, Australia

**Author email addresses:**

Skye Badger E: skye.badger@adelaide.edu.au

Sam Abraham E: S.abraham@murdoch.edu.au

Sugiyono Saputra E: Sugiyono.saputra@adelaide.edu.au

Darren Trott E: darren.trott@adelaide.edu.au

John Turnidge E: john.turnidge@safetyandquality.gov.au

Tahlia Mitchell E: tahlia.mitchell@gmail.com

Charles Caraguel E: charles.caraguel@adelaide.edu.au

David Jordan E: david.jordan@dpi.nsw.gov.au

¶ These authors contributed equally to this work.

[1] School of Veterinary and Life Sciences, Murdoch University, 90 South Street, Murdoch, Perth, Western Australia, 6150 Australia

* Corresponding author:

David Jordan (post-publication)

T: +61 (0)2 6626 1240

E: david.jordan@dpi.nsw.gov.au

Corresponding address to which proofs are sent:

Skye Badger

M: +61 (0) 427 805 717, E: skye.badger@adelaide.edu.au

## Highlights

- Disc diffusion was accurate for predicting the resistance status of clinical *E. coli.*
- Disc diffusion was less accurate for predicting the susceptible status of isolates.
- Breakpoints derived from statistical models improved disc diffusion performance.
- Disc diffusion data can contribute to surveillance for AMR.

## Abstract

The assessment of antimicrobial resistance in bacteria derived from animals is often performed using the disc diffusion assay. However broth-microdilution is the preferred assay for national antimicrobial resistance surveillance programs. This study aimed to evaluate the accuracy of disc diffusion relative to broth-microdilution across a panel of 12 antimicrobials using data from a collection of 994 clinical *Escherichia coli* isolates from animals. Disc diffusion performance was evaluated by diagnostic sensitivity, specificity, likelihood ratio pairs and receive-operating characteristic (ROC) analysis. Data was dichotomised using CLSI susceptible and resistant clinical breakpoints. In addition, disc diffusion breakpoints produced using diffusion Breakpoint Estimation Testing Software (dBETS) were evaluated. Analysis revealed considerable variability in performance estimates for disc diffusion susceptible and resistant breakpoints (AUC ranges: 0.78 - 0.99 and 0.92 - 1.0, respectively) across the panel of antimicrobials. Ciprofloxacin, tetracycline, and ampicillin estimates were robust across both breakpoints, whereas estimates for several antimicrobials including amoxicillin-clavulanic acid, cefoxitin and gentamicin were less favourable using susceptible breakpoints. Overall performance estimates were moderately improved when dBETS susceptible breakpoints were applied. For most antimicrobials, disc diffusion was accurate at predicting

resistance of clinical *E. coli* from animals that could otherwise be determined by broth-microdilution. While disc diffusion is suboptimal for assessing the proportion of fully susceptible isolates for some drugs, sensitivity and specificity estimates provided here allow for the use of standard formula to correct this. For this reason, disc diffusion has applicability in national surveillance provided the performance of the assay is taken into account.

**Keywords**

Disc diffusion, broth-microdilution, accuracy, ROC, antimicrobial resistance, surveillance

## Introduction

The emergence and spread of bacteria resistant to multiple antimicrobials including 'last-line of defence' drugs is a critical threat to the well-being of humans, animals and the environment. Strong international consensus for global action on antimicrobial resistance (AMR) has been established within the United Nations General Assembly (United Nations, 2016) and international agencies responsible for human health, animal health and agriculture (OIE, 2015; WHO, 2015b). National surveillance programs are the cornerstone in global efforts to contain the spread of AMR (WHO, 2015a). Integrated national surveillance involving the coordinated collection of data on AMR in humans, animals and the environment is critical for detecting emerging forms of resistance and evaluating the success of policies designed to contain AMR (Laxminarayan et al., 2013).

Surveillance of AMR in animal-derived bacteria is typically focussed on commensal and zoonotic bacteria from food-producing animals rather than clinical isolates from diseased animals. While zoonotic bacteria such as *Salmonella spp*. and *Campylobacter spp*. pose the

greatest health threat to humans, commensal organisms of the gastrointestinal tract such as *Escherichia coli* are also considered high-risk for the transmission of antimicrobial resistance genes to human bacteria via food products (Shaban et al., 2014). A barrier to achieving comprehensive surveillance of all AMR risks in animals is the acquisition of data from a sufficient number of clinical isolates. This could be overcome by collecting antimicrobial assay results from veterinary laboratories either as minimum inhibitory concentration (MIC) from dilution-based assays or millimetres of zone diameter from diffusion-based assays. The MIC is widely considered to be the superior measure for quantifying an isolate's susceptibility to antimicrobials (Turnidge and Paterson, 2007), and hence, broth-microdilution is the preferred susceptibility assay for national surveillance programs (ISO, 2006; OIE, 2017b). However, disc diffusion is often favoured by veterinary laboratories as it is affordable and readily customisable for a range of animal pathogens. There is considerable scope to merge susceptibility data acquired from disc diffusion from multiple laboratories into national surveillance provided the results are comparable to those from MIC assays.

The overall accuracy of disc diffusion relative to broth-microdilution remains inconclusive despite several previous studies having evaluated the assay's performance across a range of bacterial species and antimicrobials (Benedict et al., 2013; Hoelzer et al., 2011; Klement et al., 2005; Rhodes et al., 2014; Saini et al., 2011; Schumacher et al., 2001). This may be due to limitations of isolates entering such studies including small sample size, study validity (i.e. isolates are not obtained from an epidemiologically relevant population from which inferences can be drawn) and low prevalence of resistance to antimicrobials, particularly those that are critically important to humans. For instance, of those studies which include animal-derived *E. coli*, only Benedict et al (2013) (n= 3362), Klement et al (2005) (n=231) and Rhodes (2014) (n= 304) assessed more than 200 isolates. Many previous studies have also constrained the evaluation of test performance to descriptive measures such as

observed agreement of dichotomous results, simple linear regression and error-rate bounding without considering modern statistical approaches that fully exploit the data to aid interpretation of test performance.

Inevitably the assessment of diagnostic test accuracy relies on the reference test (usually broth-microdilution) and the cut-point (or breakpoint) used to dichotomise the data. In the context of AMR, the *clinical* breakpoint may define full susceptibility (susceptible breakpoint), resistance (resistant breakpoint) or the non-susceptible population (i.e. the combination of resistant and intermediate isolates) based on available pharmacokinetic data. In the evaluation of disc diffusion performance, some studies have applied the resistant breakpoint (Benedict et al., 2013; Hoelzer et al., 2011) while others applied the susceptible breakpoint (Klement et al., 2005; Saini et al., 2011). Inevitably different breakpoints will yield different estimates of test accuracy, with a resultant trade-off between the two types of misclassification errors – false negatives and false positives. While both misclassification errors have consequences, false negatives (i.e. classified susceptible when truly resistant) are the least desired in the clinical setting. Given the breakpoint is crucial for overall assessment of test performance, inconsistency in the use of breakpoints to dichotomise data across studies is likely to also be a key factor in the reported variable performance of disc diffusion relative to MIC-based assays. This is particularly relevant when the diagnostic test is used for different purposes as is the case in the clinical setting versus broad-scale surveillance. The receiver-operating characteristic (ROC) analysis addresses this by estimating the overall diagnostic accuracy of tests with continuous outcomes across all potential breakpoints.

Therefore, the aim of this study was to develop a robust statistical approach to evaluate the accuracy of zone diameter measurements obtained by disc diffusion relative to MIC measurements obtained by broth-microdilution. The approach uses ROC analysis to summarise the relative accuracy of zone diameter measurements compared to MIC results

(from the same isolates) across a large collection of clinical *E. coli* isolates from animals. Twelve antimicrobials relevant to animal health and public health were included for evaluation. For completeness, accuracy was evaluated using both susceptible and resistant clinical breakpoints recommended by the Clinical Laboratory Standards Institute (CLSI). In addition, new disc diffusion clinical breakpoints were produced using the model-based diffusion Breakpoint Estimation Testing Software (dBETS) and compared to CLSI breakpoints.

## Methods

### Isolate collection

Data used in this study were derived from the first nation-wide survey for antimicrobial resistance in veterinary pathogens, which took place between January 2013 and January 2014 with the cooperation of all veterinary diagnostic laboratories (n = 22) in Australia (Abraham et al., 2015). The data included disc diffusion and broth-microdilution results from 994 clinical *E. coli* isolates from canine (n = 510), feline (n = 338), equine (n = 28), and other species (n = 118), excluding food-producing animals.

### Antimicrobial susceptibility testing

*E. coli* isolates underwent disc diffusion and broth-microdilution testing according to CLSI VET01-A4 protocols (CLSI, 2013). The MIC results for the isolate collection were obtained from a previous study (Saputra et al, under review Vet Microbiol). Disc diffusion testing was performed independently and at a different point in time to when broth-microdilution testing occurred. Antimicrobial agents used in this study are listed in Table 1.

The dataset was dichotomised for each antimicrobial and both assays using the susceptible and resistant clinical breakpoints specified in CLSI performance standards VET01-S3 (CLSI, 2015a) and M100-S25 (CLSI, 2015b) (Table 1). For dichotomisation using the susceptible clinical breakpoint, isolates clinically referred to as 'intermediate' or 'resistant' were collectively classified as 'non-susceptible'. For dichotomisation using the resistant clinical breakpoint, isolates were classified as 'susceptible' if their measurement value fell in the susceptible or intermediate range. Where animal-specific clinical breakpoints were unavailable or did not have corresponding MIC and zone diameter breakpoints, human clinical breakpoints were used as indicated. The exception was cefovecin as there were no CLSI clinical breakpoints available, so MIC and zone diameter susceptible and resistant breakpoints were used according to the manufacturer's recommendations. In this paper, unless otherwise specified, reference to susceptible and resistant MIC and zone diameter breakpoints refer to the CLSI recommended *clinical* breakpoints.

**Statistical Analysis**

*Relative diagnostic accuracy*

The accuracy of disc diffusion classification relative to MIC (the reference method) was evaluated by estimating relative diagnostic sensitivity, diagnostic specificity, likelihood ratios of positive and negative results, and summarised using receiver-operating characteristic (ROC) analysis. MIC and zone diameters were compared using non-parametric ROC analysis since MIC data cannot be assumed to be normally distributed. For a given breakpoint, likelihood ratio pairs summarise how many times more (or less) likely a resistant isolate will be classified as resistant then an isolate that is fully susceptible. The likelihood ratio describes the direction and strength of evidence provided by a given test result. Details on likelihood

ratios and area-under the ROC-curve (AUC) estimations are given elsewhere (Greiner and Gardner, 2000).

*Agreement estimation*

Observed agreement was calculated as the proportion of isolates with the same AMR clinical classification by disc diffusion and broth-microdilution (i.e. both test results were within the susceptible breakpoint range, or within the resistant breakpoint range). McNemar's mid-*p* test (Fagerland et al., 2013) was used to assess significance (two-tailed $p < 0.05$) in the extent of disagreement between the two tests. The mid-*p* version of the McNemar's test was used instead of the conventional McNemar's test as the count of discordant results between the two methods was often less than 25. Prevalence adjusted, bias adjusted kappa (PABAK) was calculated as a measure of agreement to adjust for imbalances caused by extreme prevalence and bias between tests (Byrt et al., 1993).

*dBETS disc diffusion breakpoint values*

The recently published diffusion Breakpoint Estimation Testing Software (dBETS) program (https://dbets.shinyapps.io/dBETS/, accessed 25 April 2017) was used to generate zone diameter susceptible and resistant clinical breakpoints for the antimicrobials evaluated in this dataset (DePalma et al., 2017). The dBETS program was used to apply spline-based probability models to account for disc diffusion assay variability, providing an advantage over commonly used methods such as the modified error-rate bounded method.

Data was imported from MS excel files into Stata version 14.1 (Stata Corporation, College Station, TX) for analysis. For each isolate and each of the 12 antimicrobials tested, the broth-microdilution and disc diffusion results were paired in wide format.

**Results**

For eleven antimicrobial agents, 994 paired observations on zone diameter by disc diffusion and MIC by broth-microdilution were available for analysis. For cefovecin, 948 paired observations were available. The overall performance of disc diffusion relative to broth-microdilution was very strong for ten antimicrobials (two antimicrobials were not evaluated due to insufficient data) at the resistant breakpoints (AUC range: 0.92-1.0) (Table 2). However at susceptible breakpoints, overall performance for all 12 antimicrobials was appreciably lower (AUC range: 0.78-0.99) (Table 2). At the susceptible breakpoint, sensitivity and specificity (reflected by AUC) varied across the antimicrobial panel, and was suboptimal for amoxicillin-clavulanic acid (AUC, 0.82), cephalothin (AUC, 0.82), cefoxitin (AUC, 0.78) and gentamicin (AUC, 0.82). Performance estimates for ciprofloxacin, trimethoprim-sulfamethoxazole and tetracycline were relatively unaffected by the choice of breakpoint (Table 2). AUC estimates could not be determined for amikacin and imipenem as the isolates were all susceptible by the reference method.

Visual comparison of ROC plots for ciprofloxacin, ceftiofur, cefovecin, ceftiofur, cephalothin, tetracycline and cefoxitin are presented in Fig 1. Here, two ROC curves are plotted on each graph to demonstrate the accuracy of disc diffusion relative to broth-microdilution using the MIC susceptible and resistant breakpoints. For ciprofloxacin, ceftiofur, and tetracycline both susceptible and resistant ROC plots show near perfect test discrimination (both curves approach the top left corner of the graph). In contrast, cefovecin, cephalothin, and cefoxitin have higher levels of misclassification error (curves distant from the top left hand corner of the graph) (Fig 1).

Table 2 shows that when resistant breakpoints were applied, relative specificity was high across all antimicrobials (range, 0.95 – 1.0) while relative sensitivity was variable

(range, 0.72 – 0.99). When susceptible breakpoints were applied, relative specificity (range, 0.81 - 1.0) and sensitivity (range, 0.23 – 0.96) estimates were notably more variable. By these criteria, disc diffusion performed poorly for several antimicrobials especially amoxicillin-clavulanic acid, cefoxitin and gentamicin. When interpreting a positive disc diffusion result, using resistant breakpoints provided stronger evidence (large $LR^+$) compared to susceptible breakpoints ($LR^+$ ranges: 21-454.6 and 3.7-220.6, respectively) (Table 3). Similarly, the evidence provided by negative disc diffusion results were stronger (small $LR^-$) when using resistant breakpoints compared to susceptible breakpoints ($LR^-$ ranges: 0.01-0.28 and 0.04-0.79, respectively). Evidence from a positive disc diffusion result was weakest for cephalothin and ampicillin (lowest $LR^+$) and strongest for ciprofloxacin (highest $LR^+$) regardless of the breakpoint (Table 3). Evidence from a negative disc diffusion result was weakest for amoxicillin-clavulanic acid (highest $LR^-$) and strongest for ciprofloxacin (lowest $LR^-$) (Table 3).

Two-graph receiver-operating characteristic (TG-ROC) plots for disc diffusion relative to broth-microdilution shows the impact of breakpoint on sensitivity and specificity and hence the level of misclassification error (Fig 2). Sensitivity and specificity are equal at the point where the two lines intersect on the TG ROC plot, however the point of intersection does not always equate to the optimal breakpoint since the cost of misclassification errors almost always differs. CLSI and dBETS zone diameter breakpoints are plotted for comparison. For ciprofloxacin, CLSI and dBETS susceptible and resistant breakpoints correspond to almost perfect specificity with optimal sensitivity estimates (Tables 2 and 5). Similarly using both approaches, breakpoints for cefovecin and trimethoprim-sulfamethoxazole target the highest specificity and albeit with correspondingly lower sensitivity (Tables 2 and 5).

Observed agreement estimates were strong for most antimicrobials on resistant breakpoints (range, 0.94 - 1.0), but highly variable using susceptible breakpoints (range, 0.39 - 0.99) (Table 4). (Supplementary Tables 3 and 4 outline the contribution of positive agreement and negative agreement towards overall observed agreement estimates using susceptible and resistant breakpoints). Antimicrobials with greater than 1% difference between proportion resistant by broth-microdilution and proportion resistant by disc diffusion recorded a statistically significant ($p$ <0.05) mid-$p$ value McNemar's test (Table 4). Amoxicillin-clavulanic acid, cephalothin and cefoxitin recorded excessively large differences between the proportions resistant by broth-microdilution and disc diffusion based on susceptible breakpoints. These three antimicrobials also performed sub-optimally when inter-test agreement was measured by PABAK (Table 4). Antimicrobials with the lowest disc diffusion performance estimates also had increased overlapping susceptible and non-susceptible populations (Fig. 3). Disc diffusion estimates of accuracy are optimised when there is clear separation of 'susceptible' and 'non-susceptible' populations as demonstrated on the zone diameter histograms for ciprofloxacin, tetracycline, and ceftiofur (Fig. 3). However, disc diffusion estimates are weaker when susceptible and non-susceptible populations overlap (e.g. amoxicillin-clavulanic acid, cephalothin, and cefoxitin).

Improved disc diffusion performance estimates were produced when dBETS zone diameter susceptible breakpoints were applied (Table 5). This was particularly evident for amoxicillin-clavulanic acid where sensitivity went from 0.23 using the CLSI susceptible breakpoint to 0.61. However cefoxitin (CLSI: 0.33; dBETS 0.43) and gentamicin (CLSI: 0.50, dBETS: 0.50) estimates were minimally improved. At the resistant breakpoint, disc diffusion performance was relatively unchanged when the dBETS values were applied. At dBETS susceptible breakpoint, observed agreement for many of the antimicrobials evaluated was improved (Table 5) compared to CLSI susceptible breakpoints (Table 4).

## Discussion

Inferences made in this work are based on a large number of clinical *E. coli* isolates (n=994) from multiple animal species, and procured from a formal survey involving all major veterinary laboratories in Australia. The most notable finding of this study is the marked superiority in the performance of disc diffusion relative to broth-microdilution when assessed on resistant breakpoints compared to susceptible breakpoints. When resistant breakpoints are applied to broth-microdilution results, a very high level of disc diffusion relative accuracy is evident for the majority of antimicrobials evaluated, particularly for critically important antimicrobials (i.e. fluoroquinolones and third-generation cephalosporins). In comparison, disc diffusion performance was lower for most antimicrobials at susceptible breakpoints. This study also provides dBETS zone diameter breakpoints which have a greater objective basis than the current approach used to establish CLSI zone diameter breakpoints. The performance of disc diffusion for amoxicillin-clavulanic acid, cefoxitin and gentamicin was particularly sensitive to the choice of breakpoints, resulting in highly variable sensitivity estimates and large discrepancies in observed agreement. Cephalothin and trimethoprim-sulfamethoxazole had poor disc diffusion performance estimates regardless of the breakpoint used to dichotomise the data.

Observations arising from this study demonstrate that disc diffusion is appropriate to differentiate a population of clinical *E. coli* isolates derived from animals using CLSI or dBETS zone diameter *resistant* breakpoints for the majority of antimicrobials assessed in this study. However, for several antimicrobials including amoxicillin-clavulanic acid, cefoxitin and gentamicin, disc diffusion has limitations when differentiating a population of clinical *E. coli* isolates using CLSI zone diameter *susceptible* breakpoints. Susceptible zone diameter

breakpoints generated by dBETS are sometimes superior and should be considered when breakpoints are established. These findings also inform on the selection of antimicrobials for inclusion in national surveillance, with disc diffusion estimates for ciprofloxacin ceftiofur, ampicillin and tetracycline proving robust across breakpoints.

The study outcomes also support improved clinical decision-making by providing robust estimates of sensitivity and specificity for disc diffusion that hitherto have been rarely reported. These parameters, along with likelihood ratio pairs and ROC analysis, are key metrics relied upon in evidence-based approaches to clinical decision-making and the assessment of diagnostic test performance (Dohoo et al., 2009; OIE, 2017a). Moreover in a surveillance setting, the 'true' prevalence (Rogan and Gladen, 1978) of resistance in a population can be estimated if sensitivity and specificity are known. Calculating true prevalence from sensitivity and specificity will adjust for the inaccuracy of disc diffusion (i.e. apparent prevalence) and allow for comparison of zone diameter prevalence with MIC prevalence. This will improve the validity of surveillance data obtained from clinical *E. coli* isolates from animals. Thus, the quantitative estimates of test performance provided here for a broad panel of antimicrobials stands to benefit both population health and clinical medicine.

ROC analysis is useful to determine test accuracy and assist in defining breakpoint values however, only a small number of microbiology studies have utilised ROC analysis for determination of performance of phenotypic susceptibility assays in veterinary isolates (Jean et al., 2015; Klement et al., 2005; Saini et al., 2011; Schumacher et al., 2001). Hanzcar et al (2010) identified the need for large sample sizes in ROC estimation of assay performance (Hanczar et al., 2010) which has been achieved in this study. Although efforts have been made to utilise ROC analysis for veterinary pathogens, the sample size in such studies has been small, for example Saini et al (2011) perform ROC analysis for disc diffusion using a

sample of 25 *E. coli* isolates, and Klement et al (2005) used 231 *E. coli* isolates from bovine milk samples.

Discrepancies in disc diffusion performance estimates for some antimicrobials found here are in agreement with other studies (Hombach et al., 2013; Klement et al., 2005). While variable performance estimates may be attributed to biological differences, technical limitations (including laboratory error), or true variation in the disc diffusion test, the appropriateness of the breakpoints must also be considered. Not all antimicrobials evaluated in this study have breakpoints specific for veterinary isolates, making it necessary to use human breakpoints. This has likely resulted in variable disc diffusion performance estimates for some drugs. Additionally for trimethoprim-sulfamethoxazole, the trailing endpoint phenomenon seen with MIC assays (Jorgenson and Turnidge, 2015) may be responsible for variability in disc diffusion performance results. Epidemiological cut off points (ECOFFs) are often used as the basis for performing surveillance (Silley, 2012). However, owing to the existing complexity of this study (involving 12 antimicrobials and use of two breakpoints) ECOFFs were not included in the analysis. Nevertheless, ECOFFs for a given drug are often similar to, or lower than CLSI susceptible breakpoints and the conclusion of reduced test accuracy for disc diffusion compared to broth microdilution will also hold for interpretations based on ECOFFs. It was also evident in this study that overlapping susceptible and non-susceptible populations resulted in misclassification errors. In this study, misclassification errors were retained to replicate the imperfections that would likely occur if the veterinary laboratory network were to submit routine disc diffusion data for use in national surveillance. The dBETS method appeared relatively robust to outliers for most of the antimicrobials assessed.

Limitations associated with this study should be considered. This study only examined clinical *E. coli* isolates therefore the findings should not be generalized to non-

pathogenic (commensal) *E. coli* from healthy animals typically included in AMR surveillance. Data for this study was generated in a single laboratory and does not accommodate the possibility of laboratory-to-laboratory variation (reproducibility) in test performance. Broth-microdilution is an imperfect reference test and the performance estimates for disc diffusion can never exceed those of broth-microdilution. Theoretically, better disc diffusion accuracy estimates can be obtained by latent class analysis (Pepe and Janes, 2007) which is not reliant on a perfect reference test, however the assumptions that underlie this approach precludes its use in this study. While accuracy measures such as sensitivity, specificity, and AUC provide the best available evidence of inter-test compatibility, agreement measures such as observed agreement, McNemars test, and inter-test agreement have been reported in this study to facilitate comparison with previous studies. In the future, the existing isolate collection will be expanded to aid in the development of clinical breakpoints unique for animal species, disease syndromes or combinations of these.

**Conclusion**

We have demonstrated that for most antimicrobials, disc diffusion was shown to be accurate at predicting the resistance status of animal-derived clinical *E. coli* that could otherwise be obtained by broth-microdilution. However, for a sub-set of antimicrobials disc diffusion demonstrated inferior performance relative to broth-microdilution and this warrants further investigation. Although disc diffusion performance at the susceptible breakpoint is suboptimal, standard equations can be applied to correct this. Moreover, these findings inform on the selection of antimicrobials for inclusion in national surveillance, with disc diffusion performing well for critically important antimicrobial classes such as fluoroquinolones and third-generation cephalosporins. For these reasons disc diffusion

appears to have applicability in national surveillance provided performance of the assay, as defined in this work, is taken into account.

**Acknowledgements**

The authors thank Alysha Thriscutt and Quinn Mackie for technical assistance, and Dr Jeff Watts for his constructive comments. The authors acknowledge the assistance and support of private, government and university veterinary diagnostic laboratories within Australia for the provision of the isolates.

**Declaration of Interest statement**

This work was supported by an Australian Research Council Grant (LP130100736) and Zoetis Australia. Sam Abraham and Darren Trott have received research grants from Zoetis and Novartis; David Jordan has received funds from Meat and Livestock Australia for research advising on food safety issues in red meat production.

**Ethical approval**

None required.

**References**

Abraham, S., Jordan, D., Wong, H.S., Johnson, J.R., Toleman, M.A., Wakeham, D.L., Gordon, D.M., Turnidge, J.D., Mollinger, J.L., Gibson, J.S., Trott, D.J., 2015. First detection of extended-spectrum cephalosporin- and fluoroquinolone-resistant

Escherichia coli in Australian food-producing animals. Journal of Global Antimicrobial Resistance 3, 273-277.

Benedict, K.M., Gow, S.P., Checkley, S., Booker, C.W., McAllister, T.A., Morley, P.S., 2013. Methodological comparisons for antimicrobial resistance surveillance in feedlot cattle. BMC Vet Res 9, 216.

Byrt, T., Bishop, J., Carlin, J.B., 1993. Bias, Prevalence and Kappa. Journal of Clinical Epidemiology 46, 423-429.

CLSI, 2013. Performance Standards for Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated From Animals (VET01-A4), 4th Edition. Clinical and Laboratory Standards Institute, Wayne, PA.

CLSI, 2015a. Performance Standards for Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated From Animals (VET01-S3). 3rd Edition. Clinical and Laboratory Standards Institute Wayne, PA.

CLSI, 2015b. Performance Standards for Antimicrobial Susceptibility Testing; (M100-S25). 25th Informational Supplement. Clinical and Laboratory Standards Institute, Wayne, PA.

DePalma, G., Turnidge, J., Craig, B.A., 2017. Determination of disk diffusion susceptibility testing interpretive criteria using model-based analysis: development and implementation. Diagn Microbiol Infect Dis 87, 143-149.

Dohoo, I., Martin, W., Stryhn, H. 2009. Screening and Diagnostic Tests, In: Veterinary Epidemiologic Research. VER Inc, Charottetown, 91-127.

Fagerland, M.W., Lydersen, S., Laake, P., 2013. The McNemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. BMC Med Res Methodol 13, 91.

Greiner, M., Gardner, I.A., 2000. Epidemiologic issues in the validation of veterinary diagnostic tests. Preventive Veterinary Medicine 45, 3-22.

Hanczar, B., Hua, J., Sima, C., Weinstein, J., Bittner, M., Dougherty, E.R., 2010. Small-sample precision of ROC-related estimates. Bioinformatics 26, 822-830.

Hoelzer, K., Cummings, K.J., Warnick, L.D., Schukken, Y.H., Siler, J.D., Groehn, Y.T., Davis, M.A., Besser, T.E., Wiedmann, M., 2011. Agar Disk Diffusion and Automated Microbroth Dilution Produce Similar Antimicrobial Susceptibility Testing Results for Salmonella Serotypes Newport, Typhimurium, and 4,5,12:i-, But Differ in Economic Cost. Foodborne Pathogens and Disease 8, 1281-1288.

Hombach, M., Bottger, E.C., Roos, M., 2013. The critical influence of the intermediate category on interpretation errors in revised EUCAST and CLSI antimicrobial susceptibility testing guidelines. Clinical Microbiology and Infection 19, E59-E71.

ISO 2006. Clinical laboratory testing and in vitro diagnostic test systems -Susceptibility testing of infectious agents and evaluation of performance of antimicrobial susceptibility test devices. In Part 1: Reference method for testing the in vitro activity of antimicrobial agents against rapidly growing aerobic bacteria involved in infectious diseases. 20776-1:2006. (International Organisation for Standardisation).

Jean, S.S., Liao, C.H., Sheng, W.H., Lee, W.S., Hsueh, P.R., 2015. Comparison of commonly used antimicrobial susceptibility testing methods for evaluating susceptibilities of clinical isolates of Enterobacteriaceae and nonfermentative Gram-negative bacilli to cefoperazone-sulbactam. J Microbiol Immunol Infect.

Jorgenson, J.H., Turnidge, J. 2015. Susceptibility Test Methods: Dilution and Disk Diffusion Methods, In: Jorgensen, J.H., Pfaller, Michael A. (Ed.) Manual of Clinical Microbiology. ASM Press, Washington, 1253-1273.

Klement, E., Chaffer, M., Leitner, G., Shwimmer, A., Friedman, S., Saran, A., Shpigel, N., 2005. Assessment of accuracy of disk diffusion tests for the determination of antimicrobial susceptibility of common bovine mastitis pathogens: A novel approach. Microbial Drug Resistance-Mechanisms Epidemiology and Disease 11, 342-350.

Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A.K., Wertheim, H.F., Sumpradit, N., Vlieghe, E., Hara, G.L., Gould, I.M., Goossens, H., Greko, C., So, A.D., Bigdeli, M., Tomson, G., Woodhouse, W., Ombaka, E., Peralta, A.Q., Qamar, F.N., Mir, F., Kariuki, S., Bhutta, Z.A., Coates, A., Bergstrom, R., Wright, G.D., Brown, E.D., Cars, O., 2013. Antibiotic resistance-the need for global solutions. Lancet Infect Dis 13, 1057-1098.

OIE 2015. Resolution 26. Combating Antimicrobial Resistance and Promoting the Prudent Use of Anitmicrobial Agents in Animals. 83rd General Assembly (Paris, http://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/AMR/A_RESO_AMR_2015.pdf).

OIE. 2017a. Chapter 1.1.6. Principles and Methods of Validation of Diagnostic Assays for Infectious Diseases. Adopted 2013, In: Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. 7th Edition. http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/1.01.06_VALIDATION.pdf.

OIE. 2017b. Chapter 3.1 Laboratory Methodologies for Bacterial Antimicrobial Suceptibility Testing. Adopted 2012, In: Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. 7th Edition. http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.1_ANTIMICROBIAL.pdf.

Pepe, M.S., Janes, H., 2007. Insights into latent class analysis of diagnostic test performance. Biostatistics 8, 474-484.

Rhodes, N.J., Richardson, C.L., Heraty, R., Liu, J., Malczynski, M., Qi, C., Scheetz, M.H., 2014. Unacceptably high error rates in Vitek 2 testing of cefepime susceptibility in extended-spectrum-beta-lactamase-producing Escherichia coli. Antimicrob Agents Chemother 58, 3757-3761.

Rogan, W.J., Gladen, B., 1978. Estimating prevalence from the results of a screening test. American journal of epidemiology 107, 71-76.

Saini, V., Riekerink, R.G., McClure, J.T., Barkema, H.W., 2011. Diagnostic accuracy assessment of Sensititre and agar disk diffusion for determining antimicrobial resistance profiles of bovine clinical mastitis pathogens. J Clin Microbiol 49, 1568-1577.

Schumacher, H., Hoffmann, S., Holmboe, C., Moller, J.K., 2001. A procedure for evaluation and documentation of susceptibility test methods using the susceptibility of Klebsiella pneumoniae to ciprofloxacin as a model. J Antimicrob Chemother 48, 493-500.

Shaban, R., Simon, G., Trott, D., Turnidge, J., Jordan, D. 2014. Surveillance and reporting of antimicrobial resistance and antibitoic usage in animals and agriculture in Australia (Canberra).

Silley, P., 2012. Susceptibility testing methods, resistance and breakpoints: what do these terms really mean? Revue Scientifique Et Technique-Office International Des Epizooties 31, 33-41.

Turnidge, J., Paterson, D.L., 2007. Setting and revising antibacterial susceptibility breakpoints. Clin Microbiol Rev 20, 391-408.

United Nations 2016. At UN, global leaders commit to act on antimicrobial resistance. Collective effort to address a challenge to health, food security and development.

Joint News Release, 21 September 2016, OPGA/WHO/FAO/OIE, ed. (New York, http://www.who.int/mediacentre/news/releases/2016/commitment-antimicrobial-resistance/en/).

WHO 2015a. AGISAR 5-Year Strategic Framework to Support Implementation of the Global Action Plan on Antimicrobial Resistance (2015-2019). Report of the 6th Meeting., Advisory Group on Integrated Surveillance of Antimicrobial Resistance, ed. (Seoul, http://apps.who.int/iris/bitstream/10665/190954/1/9789241509534_eng.pdf?ua=1).

WHO 2015b. Global Action Plan on Antimicrobial Resistance. In 68th World Health Assembly (Geneva, http://www.wpro.who.int/entity/drug_resistance/resources/global_action_plan_eng.pdf).

**Table 1** Disc diffusion and broth-microdilution interpretative criteria for twelve antimicrobials evaluated in this study and applied to 994 clinical *Escherichia coli* isolates derived from animals.

| | | Susceptible Breakpoints | | Resistant Breakpoints | | |
|---|---|---|---|---|---|---|
| **Antimicrobial** | **Abbreviation** | **Disc diffusion zone diameter (mm)** | **Broth-microdilution MIC (µg/ml)** | **Disc diffusion zone diameter (mm)** | **Broth-microdilution MIC (µg/ml)** | **MIC range (µg/ml)** |
| Amoxicillin-clavulanic acid | AMC | ≥18* | ≤8* | ≤13* | ≥32* | 1.0 – 64 |
| Amikacin | AMK | ≥17* | ≤16* | ≤14* | ≥64* | 0.5 – 64 |
| Ampicillin | AMP | ≥17* | ≤8* | ≤13* | ≥32* | 1.0 – 128 |
| Cephalothin | CEF | ≥18* | ≤8* | ≤14* | ≥32* | 2.0 – 128 |
| Ceftiofur | CFT | ≥21* | ≤2* | ≤17* | ≥8* | 0.06 – 64 |
| Ciprofloxacin | CIP | ≥21† | ≤1† | ≤15† | ≥4† | 0.008 – 8 |
| Cefovecin | CVN | ≥23^ | ≤2^ | ≤19^ | ≥8^ | 0.12 – 128 |
| Cefoxitin | FOX | ≥18† | ≤8† | ≤14† | ≥32† | 1.0 – 128 |
| Gentamicin | GEN | ≥16* | ≤2* | ≤12* | ≥8* | 0.12 – 64 |
| Imipenem | IPM | ≥23* | ≤1* | ≤19* | ≥4* | 0.06 – 4 |
| Trimethoprim-sulfamethoxazole | SXT | ≥16* | ≤2* | ≤10* | ≥4* | 0.12 – 16 |
| Tetracycline | TET | ≥19* | ≤4* | ≤14* | ≥16* | 0.12 – 128 |

* Derived from CLSI VET01-S3.

† Derived from CLSI M100-S25.

^ Cefovecin breakpoints based on manufacturer's recommendation.

**Table 2** Diagnostic performance estimates of disc diffusion relative to broth-microdilution for 994 clinical *Escherichia coli* isolates from animals using CLSI susceptible and resistant breakpoints. DSe, diagnostic sensitivity; DSp diagnostic specificity; AUC, area under the curve. Exact 95% confidence intervals are given in supplementary materials.

| | Susceptible Breakpoint Estimates | | | Resistant Breakpoint Estimates | | |
|---|---|---|---|---|---|---|
| **Antimicrobial** | **Relative DSe** | **Relative DSp** | **AUC[a]** | **Relative DSe** | **Relative DSp** | **AUC** |
| Amoxicillin-clavulanic acid | 0.23 | 0.99 | 0.82 | 0.79 | 0.99 | 0.98 |
| Amikacin | NA | 0.99 | NA | NA | 1.0 | NA |
| Ampicillin | 0.93 | 0.81 | 0.96 | 0.97 | 0.95 | 0.98 |
| Cephalothin | 0.70 | 0.81 | 0.82 | 0.75 | 0.98 | 0.92 |
| Ceftiofur | 0.84 | 0.99 | 0.94 | 0.94 | 0.99 | 0.98 |
| Ciprofloxacin | 0.96 | 1.0 | 0.99 | 0.99 | 1.0 | 1.0 |
| Cefovecin | 0.67 | 0.96 | 0.87 | 0.88 | 0.99 | 0.97 |
| Cefoxitin | 0.33 | 1.0 | 0.78 | 0.83 | 0.99 | 0.97 |
| Gentamicin | 0.50 | 0.99 | 0.82 | 0.92 | 1.0 | 0.97 |
| Imipenem | NA | 0.99 | NA | NA | 1.0 | NA |
| Trimethoprim-sulfamethoxazole | 0.70 | 0.99 | 0.93 | 0.72 | 0.99 | 0.94 |
| Tetracycline | 0.93 | 0.98 | 0.97 | 0.95 | 0.99 | 0.98 |

NA, not available due to insufficient data for the analysis.

**Table 3** Estimates of likelihood ratios of disc diffusion relative to broth-microdilution for 994 clinical *Escherichia coli* isolates using CLSI susceptible and resistant breakpoints. $LR^+$, likelihood ratio of a positive test result; $LR^-$, likelihood ratio of a negative result. Exact 95% confidence intervals are given in the supplementary materials.

| | Susceptible breakpoint estimates | Resistant breakpoint estimates |
|---|---|---|

| Antimicrobial | $LR^+$ | $LR^-$ | $LR^+$ | $LR^-$ |
|---|---|---|---|---|
| Amoxicillin-clavulanic acid | 15.8 | 0.79 | 118.1 | 0.21 |
| Amikacin | NA | NA | NA | NA |
| Ampicillin | 4.8 | 0.09 | 21.0 | 0.03 |
| Cephalothin | 3.7 | 0.37 | 35.4 | 0.25 |
| Ceftiofur | 67.3 | 0.16 | 168.4 | 0.06 |
| Ciprofloxacin | 220.6 | 0.04 | 454.6 | 0.01 |
| Cefovecin | 17.2 | 0.34 | 131.2 | 0.12 |
| Cefoxitin | 61.8 | 0.67 | 124.9 | 0.18 |
| Gentamicin | 63.3 | 0.51 | 289.3 | 0.08 |
| Imipenem | NA | NA | NA | NA |
| Trimethoprim-sulfamethoxazole | 68.8 | 0.31 | 72.9 | 0.28 |
| Tetracycline | 53.5 | 0.07 | 154.4 | 0.05 |

NA, not available due to insufficient data for the analysis.

**Table 4** Agreement estimates between broth-microdilution and disc diffusion for 994 clinical *Escherichia coli* isolates from animals using CLSI susceptible and resistant breakpoints. Exact 95% confidence intervals for estimates are in supplementary materials. BMD, broth-microdilution; DD, disc diffusion.

| | Susceptible breakpoint estimates | | | | | Resistant breakpoint estimates | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Antimicrobial** | **BMD resistant** | **DD resistant** | **McNemars p-value** | **Observed agreement** | **PABAK** | **BMD resistant** | **DD resistant** | **McNemars p-value** | **Observed agreement** | **PABAK** |
| Amoxicillin-clavulanic acid | 0.79 | 0.18 | <0.001* | 0.39 | NA | 0.10 | 0.09 | <0.001* | 0.97 | 0.95 |
| Amikacin | 0.02 | 0.01 | 0.02* | 0.97 | 0.94 | 0.02 | 0.02 | 0.63 | 1.0 | NA |
| Ampicillin | 0.35 | 0.45 | <0.001* | 0.85 | 0.70 | 0.28 | 0.30 | <0.001* | 0.96 | 0.92 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cephalothin | 0.92 | 0.66 | <0.001* | 0.71 | 0.41 | 0.20 | 0.17 | <0.001* | 0.94 | 0.87 |
| Ceftiofur | 0.11 | 0.11 | 0.20 | 0.97 | 0.94 | 0.10 | 0.10 | 0.77 | 0.99 | 0.98 |
| Ciprofloxacin | 0.08 | 0.08 | 0.73 | 0.99 | 0.99 | 0.07 | 0.07 | 0.63 | 1.0 | 0.99 |
| Cefovecin | 0.15 | 0.14 | 0.08 | 0.92 | 0.84 | 0.10 | 0.10 | 0.17 | 0.98 | 0.97 |
| Cefoxitin | 0.25 | 0.09 | <0.001* | 0.83 | 0.65 | 0.09 | 0.08 | 0.05* | 0.98 | 0.96 |
| Gentamicin | 0.10 | 0.06 | <0.001* | 0.94 | 0.89 | 0.05 | 0.05 | 0.73 | 0.99 | 0.99 |
| Imipenem | 0.04 | 0.02 | <0.001* | 0.95 | 0.89 | 0 | 0 | 0.2 | 1.0 | 0.99 |
| Trimethoprim-sulfamethoxazole | 0.21 | 0.15 | <0.001* | 0.93 | 0.86 | 0.19 | 0.15 | <0.001* | 0.94 | 0.88 |
| Tetracycline | 0.19 | 0.19 | 0.85 | 0.97 | 0.95 | 0.18 | 0.18 | 0.30 | 0.99 | 0.97 |

NA, not available due to insufficient data for analysis.

* Significant mid-*p* McNemar's chi-square test ($p<0.05$).

**Table 5** Estimates of accuracy of disc diffusion relative to broth-microdilution for 994 clinical *Escherichia coli* isolates from animals using zone diameter interpretative criteria produced from the dBETS program. DSe, diagnostic sensitivity; DSp diagnostic specificity; ZD, zone diameter. Exact 95% confidence intervals for estimates provided in supplementary materials.

| | **dBETS Susceptible Breakpoint Estimates** | | | | **dBETS Resistant Breakpoint Estimates** | | | |
|---|---|---|---|---|---|---|---|---|
| **Antimicrobial** | **ZD susceptible breakpoint (mm)** | **Relative DSe** | **Relative DSp** | **Observed agreement[a]** | **ZD resistant breakpoint (mm)** | **Relative DSe** | **Relative DSp** | **Observed agreement** |
| Amoxicillin-clavulanic acid | 21 | 0.70 | 0.87 | 0.66 | 15 | 0.92 | 0.98 | 0.98 |
| Amikacin | 16 | NA | 1.0 | 0.97 | 12 | NA | 1.0 | 1.0 |
| Ampicillin | 11 | 0.80 | 0.98 | 0.92 | 7 | 0.96 | 0.98 | 0.97 |
| Cephalothin | 18 | 0.70 | 0.81 | 0.71 | 13 | 0.68 | 0.99 | 0.93 |
| Ceftiofur | 22 | 0.86 | 0.98 | 0.97 | 18 | 0.96 | 0.99 | 0.99 |
| Ciprofloxacin | 18 | 0.96 | 1.0 | 1.0 | 11 | 0.90 | 1.0 | 0.99 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Cefovecin | 23 | 0.67 | 0.96 | 0.92 | 19 | 0.88 | 0.99 | 0.98 |
| Cefoxitin | 22 | 0.43 | 0.97 | 0.83 | 18 | 0.91 | 0.99 | 0.98 |
| Gentamicin | 16 | 0.50 | 0.99 | 0.94 | 12 | 0.92 | 1.0 | 0.99 |
| Imipenem | 23 | NA | 0.99 | 0.95 | 15 | NA | 1.0 | 1.0 |
| Trimethoprim-sulfamethoxazole | 25 | 0.87 | 0.86 | 0.86 | 21 | 0.79 | 0.98 | 0.94 |
| Tetracycline | 18 | 0.93 | 0.99 | 0.97 | 13 | 0.95 | 0.99 | 0.98 |

NA, not available due to insufficient data for the analysis.

**Figure captions**

**Fig 1. ROC plots demonstrating overall performance of disc diffusion relative to broth-microdilution assays in clinical *Escherichia coli* isolates from animals for six antimicrobials.** The black-closed-dot curve and the open-diamond-dash curve represent the dichotomisation at resistant and susceptible breakpoints respectively. CIP, ciprofloxacin; CVN, cefovecin; CFT, ceftiofur; CEF, cephalothin; TET, tetracycline; FOX, cefoxitin.

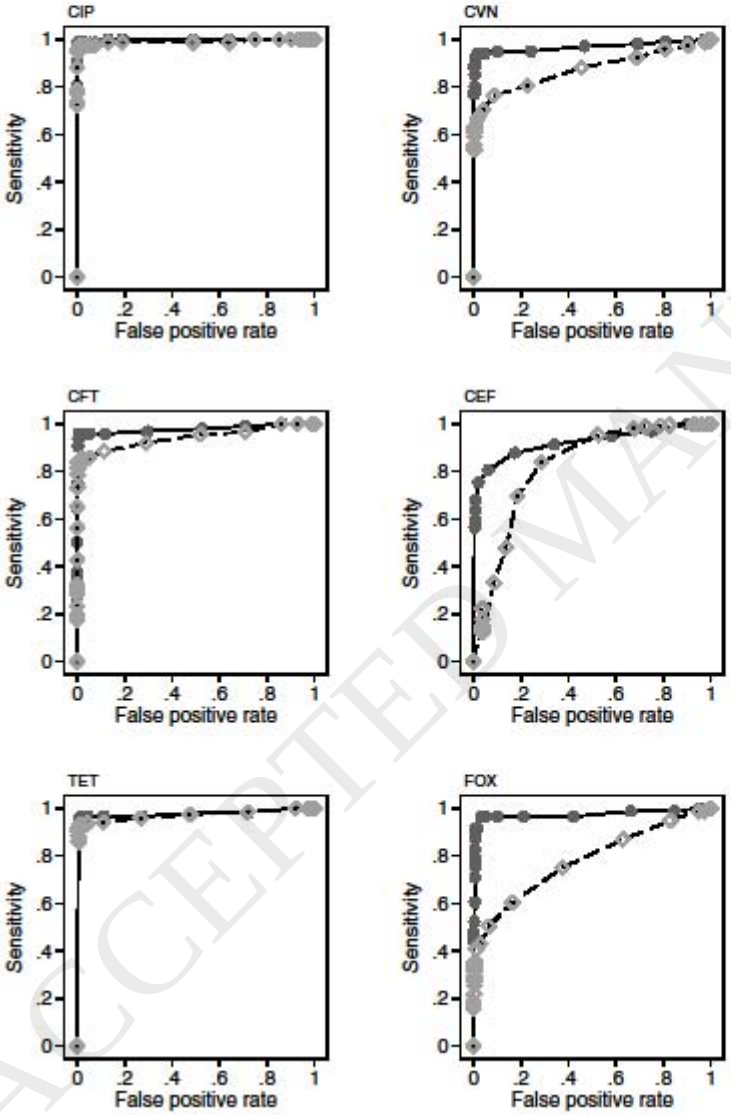


**Fig 2. Two-graph ROC (TG-ROC) plots of disc diffusion performance relative to broth-microdilution for ciprofloxacin (CIP), cefovecin (CVN), and trimethoprim-**

**sulfamethoxazole (SXT).** The TG-ROC curves for (a) susceptible and (b) resistant breakpoints are represented in the left and right column (a) and (b) respectively. Relative sensitivity (blue solid line), relative specificity (red dash line), CLSI breakpoint (black solid line), and dBETS breakpoint (green-dash line) are plotted on each graph.

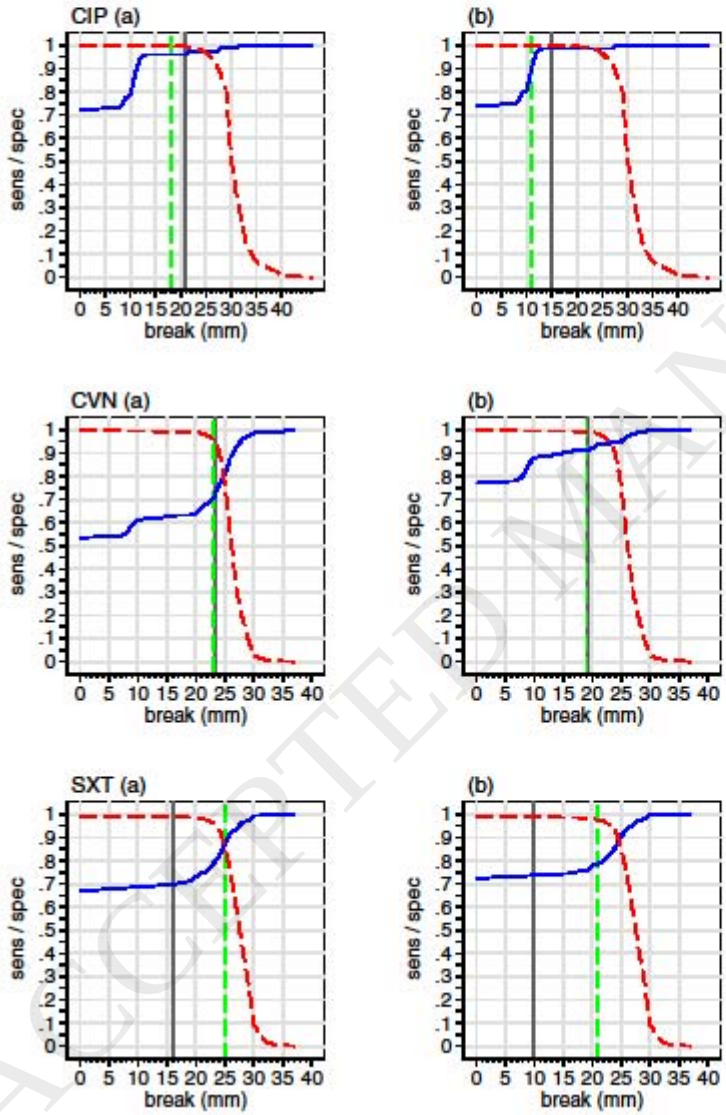


**Fig 3. Distribution of zone diameter results for clinical *E. coli* isolates derived from animals (n=994) for six antimicrobials.** CLSI resistant breakpoint (red short-dash line) and

susceptible breakpoint (blue long-dash line) is plotted over each distribution. AMC, amoxicillin-clavulanic acid; CIP, ciprofloxacin; CEF, cephalothin; TET, tetracycline; FOX, cefoxitin; CFT, ceftiofur.