Establishing statistical equivalence of data from different sampling approaches for assessment of bacterial phenotypic antimicrobial resistance

Heman Shakeri\textsuperscript{a,b,§} and Victoriya Volkova\textsuperscript{c,d,§,#}, Xuesong Wen\textsuperscript{a,b}, Andrea Deters\textsuperscript{a,c}, Charley Cull\textsuperscript{e}, James Drouillard\textsuperscript{f}, Christian Müller\textsuperscript{f}, Behnaz Moradjamei\textsuperscript{g}, and Majid Jaberi-Douraki\textsuperscript{a,h,#}

\textsuperscript{a} Institute of Computational Comparative Medicine, 
\textsuperscript{b} Department of Anatomy and Physiology, 
\textsuperscript{c} Department of Diagnostic Medicine/Pathobiology, 
\textsuperscript{d} Center for Outcomes Research and Epidemiology, 
\textsuperscript{f} Department of Animal Sciences and Industry, 
\textsuperscript{g} Department of Statistics, 
\textsuperscript{h} Department of Mathematics, 
Kansas State University, Manhattan, Kansas, USA
\textsuperscript{e} Veterinary and Biomedical Research Center, Inc., Manhattan, Kansas, USA

\textsuperscript{§} - H.S. and V.V. contributed equally to this work.

\textsuperscript{#} Corresponding authors: Majid Jaberi-Douraki, jaberi@vet.k-state.edu 
Victoriya Volkova, vv88@vet.k-state.edu

Running title: Equivalence tests for MIC distributions
Abstract

To assess phenotypic bacterial antimicrobial resistance (AMR) in different strata (e.g., host populations, environmental areas, manure, or sewage-effluents) for epidemiological purposes, isolates of target bacteria can be obtained from a stratum using various sample-types. Also, different sample-processing methods can be applied. Minimum inhibitory concentration (MIC) of each target antimicrobial drug for each isolate is measured. Statistical-equivalence testing of the MIC data for the isolates allows evaluating whether different sample-types or sample-processing methods yield equivalent estimates of the bacterial antimicrobial-susceptibility in the stratum. We demonstrate this approach on the antimicrobial-susceptibility estimates for (1) nontyphoidal Salmonella spp. from ground or trimmed meat vs. cecal-contents samples of cattle in processing plants in 2013-2014 and (2) nontyphoidal Salmonella spp. from urine, fecal, and blood human samples in 2015 (U.S. National Antimicrobial Resistance Monitoring System data). We find the sample-types for cattle yield non-equivalent susceptibility estimates for several antimicrobial drug classes and thus may gauge distinct Salmonellae sub-populations. The quinolone and fluoroquinolone susceptibility estimates for nontyphoidal Salmonellae from human blood are non-equivalent to those from urine or feces, conjecturally due to the fluoroquinolone (ciprofloxacin) use to treat nontyphoidal Salmonellae infections. We also demonstrate statistical-equivalence testing for comparing sample-processing methods for fecal samples (culturing one vs. multiple aliquots per sample) to assess AMR in fecal Escherichia coli. These methods yield equivalent results, except for tetracyclines. Importantly, statistical-equivalence testing provides the MIC difference at which the data from two sample-types or sample-processing methods differ statistically. Data-users (e.g., microbiologists, epidemiologists) may then interpret practical relevance of the difference.
Bacterial antimicrobial resistance (AMR) needs to be assessed in different populations or strata for the purposes of surveillance and determining the efficacy of interventions to halt AMR dissemination. To assess phenotypic antimicrobial-susceptibility, isolates of target bacteria can be obtained from a stratum using different sample-types or employing different sample-processing methods in the laboratory. Minimum inhibitory concentration (MIC) of each target antimicrobial drug for each of the isolates is measured, yielding the MIC distribution across the isolates from each sample-type or sample-processing method. We describe statistical-equivalence testing for the MIC data for evaluating whether two sample-types or sample-processing methods yield equivalent estimates of the bacterial phenotypic antimicrobial-susceptibility in the stratum. This includes estimating the MIC difference at which the data from the two approaches differ statistically. Data-users (e.g., microbiologists, epidemiologists, public health professionals) can then interpret whether that present difference is practically relevant.
Introduction

Informative assessments of bacterial antimicrobial resistance (AMR) within and among strata is the basic block in any investigation of AMR epidemiology or control approaches (1, 2). Such assessments are critical for identifying influential factors and mitigation strategies for AMR (1, 2). Examples of the strata are animal or human populations, food products, environmental areas, manure-effluents from food-animal farms, and human sewage-effluents. To assess AMR of a target bacterial species in a stratum, isolates of the bacteria are obtained from the sampling units (e.g., animal hosts or environmental area’s segments) in the stratum. Each isolate’s phenotypic susceptibility to each target antimicrobial drug is measured as the drug’s minimum inhibitory concentration (MIC) inhibiting visible overnight growth of the isolate culture (3). The data for all the obtained isolates provide the distribution of the tested antimicrobial’s MIC as an estimate for the target bacteria in the stratum. Descriptive statistics (e.g., elemental features of the data such as means, percentiles, or ranges) have been used extensively for the MIC-distributions due to the ease of interpretation (the statistics have been reviewed in detail in Wagner et al. 2003 (4)). Such distributions, however, could be subjected to statistical analyses to identify patterns and dynamics of the bacterial antimicrobial susceptibility in the stratum, compare sampling approaches or microbiological sample-processing methods for the susceptibility assessment, or contrast the susceptibilities between strata. Analyzing the MIC-distributions bears numerous challenges because the distributions tend to have complex shapes (e.g., do not follow the probability distributions commonly assumed for parametric statistical tests) and are inherently censored (i.e., all the isolates with MIC ≤ smallest drug concentration tested are in one category placed in the beginning and all the isolates with MIC > largest drug concentration tested are in one category located in the end of the distribution) (4-7). Thus far, the
analytical approaches have included comparing the histograms of relative frequency of the isolates with the specific MICs of the antimicrobial (7) and the cumulative frequency of the isolates over the increasing MIC-value of the antimicrobial (8, 9) in a stratum over time and between strata. The cumulative frequency distributions have been also used for comparing the antimicrobial susceptibility estimates between the isolate-sets from different sampling approaches in a stratum (10). Survival analysis has been adapted to compare the probabilities of isolates with the specific MICs of the antimicrobial (the time-to-event is replaced by the concentration-to-bacterial growth inhibition, MIC) in a stratum over time and between strata defined by experimental factors (6, 11). Linear regression on the $\log_2$ (MIC) has been used to compare the susceptibility to the antimicrobial in a stratum over time and between strata in the probabilistic framework (12), and to compare the MIC measurements obtained for the same strain-set by different microbiological laboratories in the Bayesian framework (5). It has been suggested that a power analysis should be included for the statistical tests of tendencies in the MIC/$\log_2$ (MIC) distributions, to support interpretation of the results (12).

Different sampling approaches can be used to assess AMR in a target bacterial species in a stratum. For example, different sample-types can be collected, from which the bacteria are then isolated. In other situations, once the samples have been collected, those can be subjected to different sample-processing methods for the bacterial isolation. For example, an aliquot of the sample can be plated on a bacteriological agar and a different number of the bacterial colonies tested for susceptibility to antimicrobials, or multiple aliquots of the sample can be plated and the bacterial colonies from each aliquot tested. The same analytical need arises in both these scenarios: sampling a stratum by different sample-types and applying different sample-processing methods to the samples of one type. The need is to determine whether the sampling or
the sample-processing approaches yield similar estimates of phenotypic antimicrobial-
susceptibility in the target bacteria in the stratum. This question can be formulated as to whether
the approaches yield equivalent estimates of the antimicrobial’s MIC distribution for the bacteria
in the stratum. This can be addressed by the statistical equivalence testing (13, 14). This
technique also provides a flexibility for the data-users to interpret whether the existing
differences between the bacterial susceptibility estimates for the stratum between the sampling or
sample-processing approaches are practically relevant (as shown below). The objective of this
study was to demonstrate the utility of the statistical equivalence testing as a method to compare
the bacterial antimicrobial susceptibility estimates for a stratum between sampling approaches
(e.g., different sample-types or sampling schemes) or sample-processing methods.

Results

Interpretation of the statistical equivalence testing for MIC data from different
sampling or sample-processing approaches. The most commonly used measurement of
susceptibility of a bacterial isolate to an antimicrobial is the drug’s MIC. When the MIC is
measured using the broth microdilution assay based on serial 2-fold dilutions of the drug, the
measurement is transformed to $\log_2(\text{MIC})$ for statistical analyses (12, 15). The measurements for
all the target bacterial species’ isolates obtained via a given sampling or sample-processing
approach from the target stratum yield the antimicrobial’s MIC distribution for the species in the
stratum. Such distributions from two sampling or sample-processing approaches can be
compared, and the minimum difference between the average $\log_2(\text{MIC})$ estimates from the two
approaches at which the estimates are still statistically equivalent can be determined. We denote
that difference $\Delta_{min}$. This threshold difference value can be found by performing the statistical
equivalence testing on the log$_2$(MIC) data from the two approaches starting from a large value of the difference $\Delta \gg 1 \geq \Delta_{\text{min}}$ and then reducing it until finding $\Delta \rightarrow \Delta_{\text{min}}$ below which the hypothesis of a statistically significant difference between the average log$_2$(MIC) cannot be rejected. This leads to the estimate of $\Delta_{\text{min}}$ obtained from the confidence interval of the difference between the average log$_2$(MIC) estimates from the two sampling or sampling-processing approaches (see Materials and Methods for details). The estimate of $\Delta_{\text{min}}$ can be interpreted by data-users as illustrated in Fig. 1. If a practically relevant difference $\Delta_1$ is outside $\Delta_{\text{min}}$, i.e., $\Delta_1 \geq \Delta_{\text{min}}$, the two approaches yield statistically equivalent data. On the contrary, if a practically relevant difference $\Delta_2$ is inside $\Delta_{\text{min}}$, i.e., $\Delta_2 < \Delta_{\text{min}}$, the two approaches yield statistically non-equivalent data. Thus, data-users could apply their perspectives of which difference between the average log$_2$(MIC) estimates from the two sampling or sample-processing approaches is practically relevant, and compare that to the existing statistically significant difference, $\Delta_{\text{min}}$. Summarizing the results as illustrated in Figs. 2-4 enables evaluating the differences in the average log$_2$(MIC) estimates between the two sampling or sample-processing approaches for individual antimicrobials tested (within and between the drug classes).

Here we provide a suggestion on how $\Delta_{\text{min}}$ could be interpreted systematically. When bioequivalence of two drug preparations is investigated based on a biological drug-response variable for which logarithmic transformations are appropriate, the preparations are considered equivalent if the difference $\Delta$ in the variable values is such that $2^\Delta \leq 1.25$ (16). This corresponds to $\leq$0.32 on the log$_2$(MIC) scale. If none of the two sampling or sample-processing approaches compared is a reference for the bacterial antimicrobial susceptibility assessment, data-users can consider the absolute value of the difference between the average log$_2$(MIC)
estimates. They could interpret that the two approaches yield non-equivalent estimates of the average log₂(MIC) if Δᵦ > 0.32. Such values of Δᵦ suggest the estimates differ beyond expected biological variation if the two approaches gather the isolates from the same sub-population of the target bacteria in the sampled stratum. Note that the statistical determination of Δᵦ accounts for variability in the data from the two approaches (see Materials and Methods for details).

Case study 1: Ground or trimmed meat vs. cecal-contents samples from cattle in processing plants for assessing antimicrobial susceptibility of nontyphoidal S. enterica subsp. enterica in cattle. Monitoring of AMR in the U.S. food chain is conducted by the National Antimicrobial Resistance Monitoring System (NARMS) (17). In cattle processing plants, both samples of ground or trimmed meat and of cecal-contents of cattle carcasses were collected in 2013-2014 (17, 18). Nontyphoidal Salmonella enterica subsp. enterica of diverse serovars were isolated from both these sample-types (17, 18). Phenotypic susceptibility of the isolates to antimicrobials representing major antimicrobial drug classes was tested (Table 1) (17, 18) (the data can be found here: https://www.fda.gov/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/ucm416741.htm). As the first case study, we investigated statistical equivalence of the average log₂(MIC) estimates of each tested antimicrobial (Table 1) for S. enterica yielded by the ground or trimmed meat samples (n=310 for 2013 and n=344 for 2014) vs. cecal-contents samples (n=435 for 2013 and n=318 for 2014). In each 2013 and 2014, the equivalence testing was used to determine for each antimicrobial Δᵦ – the difference between the average log₂(MIC) from the two sampling approaches at which the approaches still
yielded statistically equivalent data. The distributions of the log₂(MIC) for all the tested antimicrobials in each 2013 and 2014 did not follow a Normal distribution (Wilk-Shapiro test, p-value <0.05 for each of the two distributions from the sampling approaches). Because of this, the robust t-test was used for the equivalence testing (see Materials and Methods for details). The values of Δₘᵢₙ tended to be larger for phenicols, sulfonamides, and tetracyclines, in both 2013 and 2014 (Fig. 2A-B). Relatively large Δₘᵢₙ were also estimated for β-lactams, both aminopenicillins and cephems, in 2013 (Fig. 2A). The statistical non-equivalence of the log₂(MIC) data for these drug classes highlighted that the cecal-contents sampling gauged *S. enterica* populations that differ in their phenotypic AMR from *S. enterica* populations gauged via the ground or trimmed meat sampling in the cattle processing plants. This interpretation was based on considering the data for each antimicrobial from the two sampling approaches non-equivalent if Δₘᵢₙ > 0.32.

**Case study 2: Antimicrobial susceptibility of nontyphoidal *S. enterica* subsp. *enterica* from urine vs. fecal vs. blood samples of humans.** Monitoring of AMR in enteric pathogens of humans in the U.S. is also a part of the NARMS activities (17). Nontyphoidal *Salmonella enterica* subsp. *enterica* of diverse serovars were isolated from urine, fecal, and blood samples of humans in the U.S. in 2015 (the analyzed dataset is limited to those states that have permitted the U.S. Centers for Disease Control and Prevention to share the data with the public; the dataset can be found here: https://wwwn.cdc.gov/narmsnow/). Phenotypic susceptibility of the isolates to antimicrobials representing major antimicrobial drug classes was tested (Table 2). As the second case study, we applied the statistical equivalence testing to compare the estimates of antimicrobial susceptibility of nontyphoidal *S. enterica* isolates from urine (n=144), fecal
(n=1,495), and blood (n=181) samples of humans in the U.S. in 2015. The distributions of the 
\( \log_2(\text{MIC}) \) for all the tested antimicrobials for the \( S. \text{enterica} \) isolates from each urine, fecal, or 
blood samples did not follow a Normal distribution (Wilk-Shapiro test, \( p \)-value <0.05 for each 
distribution). Thus, the robust \( t \)-test was used for the equivalence testing. The results 
demonstrated that susceptibilities of the \( S. \text{enterica} \) isolates from human urine vs. from blood to 
\( \beta \)-lactams cephems, macrolides, phenicols, and quinolones were statistically non-equivalent (Fig. 
3A). Further, susceptibilities of the \( S. \text{enterica} \) isolates from human feces vs. from blood to 
aminoglycosides, \( \beta \)-lactams cephems, quinolones, and tetracyclines were non-equivalent (Fig. 
3B). While susceptibilities of the \( S. \text{enterica} \) isolates from human urine vs. from feces differed to 
a lesser extent, but still were non-equivalent for aminoglycosides and tetracyclines (Fig. 3C). 
These interpretations for each antimicrobial in a pair-wise comparison of the isolate sources were 
based on considering the data non-equivalent if \( \Delta_{\text{min}} > 0.32 \). The largest differences were found 
for the fluoroquinolone ciprofloxacin and older quinolone nalidixic acid (Fig. 3A-B), to which 
the isolates from human blood had lower susceptibilities (higher ciprofloxacin and nalidixic acid 
MIC values) than the isolates from either urine or feces. The difference between the average 
\( \log_2(\text{MIC}) \) of ciprofloxacin for the isolates from blood vs. urine was 1.08 (95% CI 0.74, 1.42), 
and for the isolates from blood vs. feces it was 1.20 (95% CI 0.86, 1.54). The difference between 
the average \( \log_2(\text{MIC}) \) of nalidixic acid for the isolates from blood vs. urine was 0.71 (95% CI 
0.46, 0.95), and for the isolates from blood vs. feces it was 0.80 (95% CI 0.57, 0.96).

Case study 3: Processing a fecal sample for assessing antimicrobial susceptibility of 
fecal \( E. \text{coli} \): multiple bacterial isolates from one aliquot vs. one isolate from each of several 
aliquots of the sample. Assessment of AMR in a target culturable bacterial species in a fecal
sample customary involves using a single aliquot from the sample (19, 20). The aliquot is diluted (the dilution is chosen based on the expected bacterial density) and the dilution(s) are plated on a bacteriological agar (19, 20). One or more of the bacterial colonies with typical morphology for the species on the agar are selected, each of the colonies is re-plated for isolation, and the isolate’s phenotypic susceptibility to antimicrobials is tested (19, 20). However, the population of commensal bacteria such as *Escherichia coli* in feces of an animal or human could consist of genetically diverse sub-populations (21-25). We have chosen to investigate for cattle fecal pads if testing susceptibility of *E. coli* obtained from a single aliquot of the pad (conventional approach) vs. from multiple aliquots taken over the longest axis of the pad yield equivalent estimates of phenotypic antimicrobial susceptibility of fecal *E. coli*. Fresh fecal pads (*n*=32) were collected from beef and dairy cattle at research facilities at Kansas State University during May to June 2016. The animals were sampled that have not received antimicrobial drugs in the preceding week, but 1 to 3 weeks prior received either antimicrobials or feed supplemented with copper (Cu) or zinc (Zn) (which can co-select AMR in the animal fecal bacteria (11, 26-29)). This animal selection was done to ensure the cattle fecal *E. coli* will have detectable but not uniformly high levels of phenotypic AMR, to facilitate statistical analyses of the data. From each pad, four aliquots were taken equidistantly along the longest axis of the pad. One of the four aliquots was randomly selected and four *E. coli* isolates were obtained from that aliquot. From each of the other three aliquots from the pad, one *E. coli* isolate was obtained. All the isolates were tested for phenotypic susceptibility to antimicrobials, which were chosen to represent most of the existing antimicrobial drug classes (Table 3). As the third case study, we investigated for each of these antimicrobials the statistical equivalence of the average log₂(MIC) estimates for fecal *E. coli* yielded by testing four bacterial isolates from a single aliquot of the fecal pad vs.
testing one bacterial isolate from each of four aliquots taken from locations spread over the longest axis of the fecal pad. The results demonstrated detectable differences between the two sampling-processing methods in the estimates of fecal *E. coli* susceptibility to β-lactams aminopenicillins, aminoglycosides, macrolides, and tetracyclines (Fig. 4). Only for tetracyclines, the two methods yielded statistically non-equivalent estimates of fecal *E. coli* susceptibility. The interpretation for each tested antimicrobial was based on considering the data non-equivalent if $\Delta_{\text{min}} > 0.32$. The observed variability in the log$_{2}$(MIC) for antimicrobials of newer classes (Table 3) was insufficient for the testing; the *E. coli* isolates were predominantly susceptible to these antimicrobials.

**Discussion**

The three case studies provided above illustrate the utility of statistical-equivalence testing (13, 14) for establishing equivalence of data yielded by two sampling or sample-processing approaches for assessing phenotypic antimicrobial susceptibility in a target bacterial species in a stratum. Advantages of the proposed method include the estimation difference in the average log$_{2}$(MIC) estimates, $\Delta_{\text{min}}$, below which the data from the two approaches are statistically significantly different. This provides data-users with flexibility to investigate where practically relevant differences fall relative to the existing differences in the data, as illustrated in Fig. 1. Note that in the proposed method, we determine $\Delta_{\text{min}}$ between the average log$_{2}$(MIC) estimates from the two sampling or sample-processing approaches using a sequential algorithm that re-tests the statistical non-equivalence hypothesis over a range of the average log$_{2}$(MIC) difference values based on the data from the two approaches. For this, we use statistical tests that accommodate censored data with unequal variances and different shapes of the log$_{2}$(MIC)
distributions from the two approaches. This is because the data on an antimicrobial’s MIC for a set of bacterial isolates from a stratum are inherently censored (all the isolates with MIC ≤ smallest drug concentration tested are in one category in the beginning and all the isolates with MIC > largest drug concentration tested are in one category in the end of the MIC distribution). The resulting log$_2$(MIC) distributions for commonly tested antimicrobial drugs have various shapes that often do not fit to a Normal distribution (7). In the proposed method for the equivalence testing, the Welch’s $t$-test (30) relaxes the assumption of equal variances in the two compared log$_2$(MIC) distributions and the robust $t$-test (31) further improves handling of over-dispersion (e.g., presence of long tails in the distributions). We keep the results obtained using both of these tests in Fig. 2-4 for illustrative purposes. In further applications, the robust $t$-test could be recommended for the log$_2$(MIC) distributions that do not follow a Normal distribution and demonstrate over-dispersion.

In the first case study, we evaluated whether sampling the cecal-contents and sampling of ground or trimmed meat in the U.S. cattle processing plants yield equivalent data on antimicrobial susceptibility of nontyphoidal S. enterica subsp. enterica in the processed cattle. The data were collected by the NARMS in 2013-2014 to monitor AMR in the U.S. cattle production chain and therefore we interpret the results at the same population level. The results showed that S. enterica sub-populations in the cecal-contents samples may be statistically non-equivalent in their phenotypic antimicrobial susceptibility to S. enterica sub-populations in the ground or trimmed meat samples (Fig. 2). Thus, the two sampling approaches yield non-equivalent data for monitoring phenotypic antimicrobial susceptibility in S. enterica in cattle in the processing plants. Possible explanations include de-contamination and cross-contamination of cattle carcasses and products within the plant, as well as mixing of different carcass parts for
the ground meat (32-34). These processes can reduce the role of the cattle intestinal contents as a source of *S. enterica* in the meat or ground meat products.

In the second case study, we compared statistical equivalence of the antimicrobial susceptibility estimates for nontyphoidal *S. enterica* isolates from urine, fecal, and blood samples of humans in the U.S. in 2015. The data were collected by the NARMS to monitor AMR in human enteric pathogens in the U.S. and thus again we interpret the results at the same population level. The results demonstrated that susceptibilities of the *S. enterica* isolates from human blood are non-equivalent to those from urine or feces for several major antimicrobial drug classes, such as β-lactams cephems, macrolides, phenicols, quinolones, and tetracyclines (Fig. 3A-B). Lesser differences were observed between the isolates from urine vs. feces (Fig. 3C). The differences of largest magnitude were found for quinolones, with the isolates from human blood being less susceptible than those from urine or feces to the fluoroquinolone ciprofloxacin and the older quinolone nalidixic acid (see Results for more details). Conjunctively, this could be due to the common use of fluoroquinolones, *e.g.* ciprofloxacin, as one of the first-line treatment choices for treating serious infections by nontyphoidal *Salmonella* in human adults (35-38). Another common treatment choice is β-lactams cephalosporins, *e.g.* ceftriaxone (the other choices include combinatory formulations containing β-lactams and β-lactamase inhibitors, aminoglycosides, and as the last resort polymyxins and β-lactams carbapenems) (35, 38). Considering the data for 2015, susceptibilities to individual cephems of the *S. enterica* isolates from human blood were less different from (although statistically non-equivalent to) the isolates from urine or feces, compared to the differences observed for quinolones (Fig. 3A-B). Notably, across the human nontyphoidal *Salmonella* isolates, the frequency of those with reduced ciprofloxacin...
susceptibility has been continuously raising and the frequency of those with reduced ceftriaxone susceptibility has overall increased in the U.S. since 1996 (38, 39).

In the third case study, we evaluated whether testing four bacterial isolates from a single aliquot of the cattle fecal pad vs. testing one bacterial isolate from each of four aliquots taken from locations spread over the longest axis of the pad yield equivalent data on phenotypic antimicrobial susceptibility of fecal E. coli at the population level (n=32 pads were tested). The results showed that the two sample-processing methods yield statistically non-equivalent estimates of E. coli susceptibility to tetracyclines, with smaller but detectable differences for β-lactams aminopenicillins, aminoglycosides, and macrolides (Fig. 4). These antimicrobial drug classes have been used in food-animals in the U.S. for longest periods of time (40). Tetracyclines, penicillins, and aminoglycosides were introduced in the 1940s and macrolides in the 1970s (40). Consequently, multiple genes encoding various degree of susceptibility to these drug classes have been observed in fecal E. coli and S. enterica of farm animals (41-45). The tetracycline resistance gene pool is especially diverse, with several tens of tet-genes described to date in different animal and human hosts (41, 46). Testing E. coli throughout the fecal pad may capture more of the present diversity in the susceptibility to tetracyclines compared to testing E. coli at a single location in the pad. Also, statistical power of the equivalence testing depends not only on the sample size but also on variability in the log$_2$(MIC) data. Strongly bimodal (less variable) log$_2$ (MIC) distributions for newer antimicrobial drug classes, due to the high frequencies of the highly susceptible bacterial isolates, impede the testing (see Tables 1-3 for examples).

Diagnostic microbiologists consider one 2-fold dilution of the antimicrobial drug to be an acceptable variation in the MIC measurement for an individual bacterial isolate in the broth.
microdilution assay (47). If such variation occurs randomly among the isolates in the two sampling or sample-processing approaches, it is accounted for in the variance component of a statistical test of the data (e.g., see Materials and Methods). Such random variation does not bias comparisons of the data between the approaches. However, if the variation is non-random and has a systematic source, it could bias the comparisons. Consider an extreme case of the MIC being skewed by one 2-fold drug dilution for every isolate obtained from one sampling or sample-processing approach but not from the other approach, e.g., if the samples from one approach were examined in one laboratory and the samples from the other approach in another laboratory. The data-user believes the antimicrobial’s MIC measurements in one laboratory are consistently one 2-fold dilution higher or lower than the MIC measurements in the second laboratory. In this case, there would be a difference $\Delta \geq 1$ for the antimicrobial between the average $\log_2$(MIC) estimates from the two laboratories. A statistically significant difference beyond that would be manifested as $\Delta_{min} > 1$.

We have included a suggestion for interpreting a $\Delta_{min} > 0.32$ as evidence of statistical non-equivalence of the $\log_2$(MIC) data for the antimicrobial for the bacterial species in the stratum between the two sampling or sample-processing approaches. This interpretation illustrated in Fig. 2-4 is an adaptation of a method for establishing bioequivalence of two drug preparations based on values of a biological drug-response variable (16). Other standardized interpretations may be proposed in the future for decision-making on whether the two sampling or sample-processing approaches are interchangeable or yield equivalent data on phenotypic antimicrobial susceptibility of the target bacteria (i.e., assess susceptibility in the same bacterial sub-population) in the stratum.
Materials and methods

Statistical equivalence testing

Rationale for testing the statistical equivalence hypothesis. Let $M_1$ and $M_2$ be the sampling or sample-processing approaches that yield samples $Y_1$ and $Y_2$, respectively, of isolates of the target bacterial species from the stratum. The samples $Y_1$ and $Y_2$ represent sub-populations $P_1$ and $P_2$ of the species in the stratum. The isolate susceptibility to a target antimicrobial is measured (e.g., in our three case studies the susceptibility was measured using the broth microdilution assay and the obtained MIC values were log$_2$-transformed for the analysis). The statistics $\mu_1$ and $\mu_2$ represent the central tendencies of the susceptibility of the unknown source sub-populations $P_1$ and $P_2$. “Conventional” hypothesis testing focuses on rejecting $H_0$ of no statistically significant difference between the central tendencies ($H_0$: $\mu_1 - \mu_2 = 0$; $H_a$: $\mu_1 - \mu_2 \neq 0$). However, even if $H_0$ is rejected, this provides no proof in favor of $H_a$. Importantly, testing the “conventional” $H_0$ delivers no information for what $\mu_1 - \mu_2$ difference signals that the central tendencies of the samples $Y_1$ and $Y_2$ are statistically significantly different.

Equivalence hypothesis. The equivalence hypothesis testing can provide the sought information on a statistically significant $\mu_1 - \mu_2$. The null and alternative hypotheses are defined as follows (13, 14):

$$H_0: |\mu_1 - \mu_2| > \Delta$$

$$H_a: |\mu_1 - \mu_2| \leq \Delta \quad [1]$$

The equivalence hypothesis testing in Equation [1] indicates that the samples $Y_1$ and $Y_2$ obtained by the approaches $M_1$ and $M_2$ have equal means up to an acceptable tolerance $\Delta$ with a predefined confidence interval $1 - 2\alpha$ (where $\alpha$ is probability of the type I error). This null
hypothesis is rejected if the data provide evidence of the equivalence of the means. Otherwise, the null hypothesis of a statistically significant $\mu_1 - \mu_2$ difference is accepted.

**Student’s t-test of the equivalence hypothesis.** The samples $Y_1$ and $Y_2$ are obtained by the sampling or sample-processing approaches $M_1$ and $M_2$. The sample means $\bar{Y}_1$ and $\bar{Y}_2$ are employed as the point estimators of $\mu_1$ and $\mu_2$, with standard errors $se_1$ and $se_2$. Therefore, the difference $\mu_1 - \mu_2$ can be estimated by $\bar{Y}_1 - \bar{Y}_2$ with a standard error $\sqrt{se_1^2 + se_2^2}$, which is equal to:

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

[1]

where $\sigma_1$ and $\sigma_2$ are the estimates of standard deviations of $P_1$ and $P_2$, and $n_1$ and $n_2$ are the samples sizes of $Y_1$ and $Y_2$. With large sample sizes and the fact that $P_1$ and $P_2$ are known, the sampling distribution of $\bar{Y}_1 - \bar{Y}_2$ could be estimated through a Normal distribution centered at $\mu_1 - \mu_2$ with the standard error given by Equation [2]. Instead, to avoid introducing extra variability from estimating $\sigma_1$ and $\sigma_2$ using sample variances $s_1$ and $s_2$, and to be able to handle the data with different sample sizes, the Student’s t-test can be used. The Student’s t-test assumes both the samples $Y_1$ and $Y_2$ are drawn from variables that follow a Normal distribution and have equal variances, which can be estimated by a pooled variance combining the sample variances $s_1$ and $s_2$. Applying Student’s t-test for samples with different variances or sample sizes can lead to unreliable conclusions with large type I and type II error probabilities (48).

**Welch’s t-test of the equivalence hypothesis.** The Welch’s t-test can handle unequal variances or sample sizes in the data from the two sampling or sample-processing approaches (30). The statistic for the Welch’s t-test of the hypothesis defined in Equation [1] is:
\[ t_1 = \frac{\bar{Y}_1 - \bar{Y}_2 - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad t_2 = \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \]

and \( H_0 \) is rejected if \( t_1 < -t_{a, df} \) and \( t_2 > t_{a, df} \), where \( df \) denotes degrees of freedom.

However, the censored nature of the MIC data results in presence of aggregated observations in the regions \( MIC \leq L \) and \( MIC > U \) where \( L \) and \( U \) are the smallest and largest drug concentrations tested, respectively. A long-tail in the left-hand end, \( MIC \leq L \) and a long-tail in the right-hand end, \( MIC > U \), of the distribution are common (7). Such long-tailed shapes are common in the distributions even after the log2(MIC) transformation (12).

Robust t-test of the equivalence hypothesis. As noted above, the log2(MIC) distributions most often do not follow a Normal distribution. To improve robustness of the equivalence hypothesis testing and handle the long tails in the log2(MIC) distributions, we built upon the Welch’s t-test to use the trimmed data and Winsorized variance. This is known as the robust t-test (31). Let \( Y_1 \) and \( Y_2 \) be the ordered sample data, e.g., \( Y_1 = [y_{1,1}, y_{1,2}, \ldots, y_{1,n_1}] \), then under \( H_0 \):

\[ t_{t_1} = \frac{\bar{Y}_{1,tg} - \bar{Y}_{2,tg} - \Delta}{\sqrt{\frac{s_{1,wg}^2}{n_1(n_1 - 1)} + \frac{s_{2,wg}^2}{n_2(n_2 - 1)}}} \]

\[ t_{t_2} = \frac{\bar{Y}_{1,tg} - \bar{Y}_{2,tg} + \Delta}{\sqrt{\frac{s_{1,wg}^2}{n_1(n_1 - 1)} + \frac{s_{2,wg}^2}{n_2(n_2 - 1)}}} \]

where \( \bar{Y}_{1,tg} \) and \( \bar{Y}_{2,tg} \) are the trimmed (denoted by \( t \)) means and \( g \) is the number of the trimmed data points from each of the sides of the ordered \( Y_1 \) and \( Y_2 \) defined as:

\[ \bar{Y}_{1,tg} = \frac{1}{n - 2g}(y_{1,g+1} + y_{1,g+1} + \cdots + y_{1,g+1}) \]
\[ Y_{2, tg} = \frac{1}{n-2g} \left( y_{2,g+1} + y_{2,g+1} + \cdots + y_{2,g+1} \right) \]

and the Winsorized (denoted by \( w \)) sum of squares \( s_{1, wg}^2 \) and \( s_{2, wg}^2 \) are defined as:

\[
s_{1, wg}^2 = \frac{1}{n_1} \left( (g + 1)y_{1,g+1} + y_{1,g+2} + \cdots + y_{1,n-g-1} + (g + 1)y_{1,n_1-g} \right)
\]

\[
s_{2, wg}^2 = \frac{1}{n_2} \left( (g + 1)y_{2,g+1} + y_{2,g+2} + \cdots + y_{2,n-g-1} + (g + 1)y_{2,n_2-g} \right)
\]

The trimmed \( t \)-statistics \( t_{t_1} \) and \( t_{t_2} \), each follows a \( t \)-distribution with the degrees of freedom \( d_f \) (49):

\[
\frac{1}{d_f} = \frac{c^2}{h_1 - 1} + \frac{(1 - c)^2}{h_2 - 1},
\]

where \( c \) is given by:

\[
c = \frac{s_{1wg}^2}{s_{1wg}^2 h_1(h_1 - 1) + s_{2wg}^2 h_2(h_2 - 1)}.
\]

thus \( H_0 \) is rejected if \( t_{t_1} < -t_{a,d_f} \) and \( t_{t_2} > t_{a,d_f} \).

**Power analysis of the statistical equivalence testing.** The power of testing statistical equivalence of the average log2 (MIC) estimates for each antimicrobial drug from the two sampling or sample-processing approaches for the target bacterial species in the stratum can be computed (using De Morgan’s law (50)) as follows:

\[
\text{Power} = 1 - P \left( -t_{a,d_f} \left| H_a \right. \right) - P \left( t_{t_2} \leq t_{a,d_f} \left| H_a \right. \right)
\]

To compute the right-hand side of Equation [4], we use the non-central \( t \)-distribution at \( t_{t_1} \) and its cumulative distribution \( T_{df}(t_{t_1}) \):

\[
\text{Power} = 1 - T_{df} \left( t_{a,d_f} \right) - T_{df} \left( t_{a,d_f} \left| t_{t_2} \right. \right)
\]
The test can be considered to have an acceptable power if the power is ≥0.80, following published guidelines (51, 52).

**Determining $\Delta_{\text{min}}$ from 95% confidence interval of the difference of the means.** The threshold difference $\Delta_{\text{min}}$ is the minimum $\Delta$ in Equation [1] at which the two sampling or sample-processing approaches still yield statistically non-equivalent estimates of susceptibility to the antimicrobial drug of the bacterial species in the sampled stratum. This threshold value for each antimicrobial and statistical test (Welch’s $t$-test or robust $t$-test) was found from the data via a sequential algorithm repeating the test over a range of the difference values, starting from a large value (e.g., a 4 log difference between the average $\log_2$ (MIC) estimates), and then shrinking $\Delta$ by a small step (0.01 log$_2$) and re-testing the $H_0$ of a statistically significant difference between the average $\log_2$ (MIC) estimates from the two approaches, until reaching $\Delta_{\text{min}}$ value below which the $H_0$ could no longer be rejected.

**Microbiological procedures**

**Case studies 1 and 2.** Microbiological procedures used by the NARMS are described in the program’s Manual of Laboratory Methods (53). Phenotypic susceptibility of the *S. enterica* isolates to antimicrobials is determined in the broth microdilution assay using the Sensititre™ system (TREK Diagnostic Systems Inc., Cleveland, OH), in accordance with the manufacturer recommendations and the Clinical and Laboratory Standards Institute (CLSI) guidelines (47, 53). The assays for the *S. enterica* isolates from cattle processing plants in 2013-2014 were performed using the Sensititre™ plate format CMV3AGNF as of those years, which included antimicrobials listed in Table 1. The assays for the nontyphoidal *S. enterica* isolates from humans in 2015 were performed using the Sensititre™ CMV3AGNF plate format as of that year.
and an additional plate format containing broad-spectrum β-lactams; the tested antimicrobials are listed in Table 2.

**Case study 3. Sampling.** Fresh fecal pads ($n = 32$) were collected (the entire pad was lifted from the ground without mixing, placed into a sterile plastic bag, and transported while being kept horizontally) from different beef and dairy cattle at research facilities at Kansas State University during May to June 2016. The animals were sampled that have not received antimicrobial drugs in the preceding week, but 1 to 3 weeks prior received either antimicrobials (macrolides or tetracyclines to treat limited bovine respiratory disease or as a part of a research study) or feed supplemented with copper or zinc (these feed additives can co-select AMR in the animal fecal bacteria (11, 26-29)). This animal selection was done to ensure the cattle fecal *E. coli* will have detectable but not uniformly high levels of AMR, to facilitate statistical analyses of the data. A collected fecal pad weighted 1.2 kg on average with (5%, 95%) of (0.4 kg, 2.0 kg). The average (5%, 95%) of the pad dimensions were: length 24 (18, 31) cm; width 18 (12, 25); and height 3 (5, 9) cm.

**Sample processing.** On each fecal pad, four locations along the longest axis of the pad – its length – were marked using a sterile plastic loop. The locations were spread along the pad length (depending on the length) equidistantly ~3-5 cm (~1.5-2 inches) apart. Feces at the four locations were opened to the depth of ~1 cm using sterile tools (to avoid a possibility of culturing *E. coli* that may have occurred from accidental contamination of the pad exterior). One fecal aliquot of ~1 g was aseptically collected from the bottom of the opening at each of the four locations. The locations were counted left to right. A random number was generated from a Uniform (1, 4). When the aliquot from the location with the number corresponding to the generated random number was plated on a MacConkey agar plate, 4 *E. coli* colonies were
obtained from the plate for isolation. When each of the aliquots from the other three locations on
the pad was plated on a MacConkey agar plate, one *E. coli* isolate was obtained from the plate
for isolation.

**Microbiological procedures.** Each fecal aliquot of ~1 g was diluted in 10 mL of buffered
peptone water (PBS) and vortexed gently until fully mixed. Of the supernatant, 100 µL were
diluted 1:10 in sterile PBS and 100 µL were diluted 1:100 in sterile PBS. Of each of the
dilutions, 100 µL were plated on a MacConkey agar plate and incubated at 37.5°C for 24 hours.
For the randomly selected aliquot on the pad, from the MacConkey plate with well-separated
colonies, 4 typical coliform colonies chosen from different parts of the plate (convenience
randomization) were each streaked on a tryptic soy broth supplemented with 5% sheep blood
agar plate (BAP) and incubated at 37.5°C for 24 hours. For each of the other three aliquots from
the pad, from the MacConkey plate with well-separated colonies, one typical coliform colony
(chosen via convenience randomization) was streaked on a BAP plate and incubated at 37.5°C
for 24 hours. Presumptive *E. coli* colonies from each BAP plate were subjected to the indole test
and the indole-producing ones were identified as *E. coli*. When needed, additional coliform
colonies from the MacConkey plate were re-plated for isolation and subjected to the indole test
to obtain the sought number of *E. coli* isolates from the fecal aliquot (*i.e.*, one or four isolates).
Phenotypic susceptibility to antimicrobials of each *E. coli* isolate was determined in the broth
microdilution assay following the Sensititre™ plate manufacturer instructions and in accordance
with the CLSI recommendations (47, 53). The Sensititre™ plate formats GN4F and BOPO6F as
of 2016 were used. The strain *E. coli* ATCC® 25922™ was used for the quality control of the
assays, along with the positive and negative control wells. The assay results on the plates were
read on the Sensititre™ ARIS™ automated reading instrument (TREK Diagnostic Systems Inc., Cleveland, OH).

Software

The data from the NARMS 2013-2015 are publicly available in and the data for the fecal pads were gathered in Microsoft Office Excel® (Microsoft, Inc., Redmond, WA) formats. The data were imported into R 3.4, in which the statistical analyses were performed. All the figures were made in Python 3 (Python Software Foundation).

Acknowledgments. This work was supported by the Kansas Bioscience Authority via their support for the Institute of Computational Comparative Medicine at Kansas State University. As well, contributions of VVV were supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R15GM126503. The manuscript content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
References

1. Johnson AP. 2015. Surveillance of antibiotic resistance. Philosophical Transactions of the Royal Society B-Biological Sciences 370:20140080.

2. WHO. 2015. Global Antimicrobial Resistance Surveillance System: Manual for Early Implementation. http://apps.who.int/iris/bitstream/10665/188783/1/9789241549400_eng.pdf?ua=1 (Last assessed on Jan 6, 2018), p 1-2. World Health Organization, Geneva, Switzerland.

3. CLSI. 2013. Performance standards for antimicrobial disk and dilution susceptibility tests for bacteria isolated from animals; approved standard, 4th ed. CLSI document VET01-A4. Wayne, PA, USA.

4. Wagner BA, Dargatz DA, Morley PS, Keefe TJ, Salman MD. 2003. Analysis methods for evaluating bacterial antimicrobial resistance outcomes. American Journal of Veterinary Research 64:1570-1579.

5. van de Kassteele J, van Santen-Verheuvel MG, Koedijk FD, van Dam AP, van der Sande MA, de Neeling AJ. 2012. New statistical technique for analyzing MIC-based susceptibility data. Antimicrob Agents and Chemotherapy 56:1557-63.

6. Stegeman JA, Vernooij JC, Khalifa OA, Van den Broek J, Mevius DJ. 2006. Establishing the change in antibiotic resistance of Enterococcus faecium strains isolated from Dutch broilers by logistic regression and survival analysis. Preventive Veterinary Medicine 74:56-66.

7. Mazloom R, Jaberi- Douraki M, Comer J, V. VV. 2017. Potential information loss due to categorization of MIC frequency distributions. Foodborne Pathogens and Disease 15:44-54.
8. Thornsberry C, Jones ME, Hickey ML, Mauriz Y, Kahn J, Sahm DF. 1999. Resistance surveillance of *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Moraxella catarrhalis* isolated in the United States, 1997-1998. Journal of Antimicrobial Chemotherapy 44:749-759.

9. Sahm DF, Jones ME, Hickey ML, Diakun DR, Mani SV, Thornsberry C. 2000. Resistance surveillance of *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Moraxella catarrhalis* isolated in Asia and Europe, 1997-1998. Journal of Antimicrobial Chemotherapy 45:457-466.

10. Wagner BA, Dargatz DA, Salman, Morley PS, Wittum TE, Keefe TJ. 2002. Comparison of sampling techniques for measuring the antimicrobial susceptibility of enteric *Escherichia coli* recovered from feedlot cattle. American Journal of Veterinary Research 63:1662-70.

11. Agga GE, Scott HM, Amachawadi RG, Nagaraja TG, Vinasco J, Bai J, Norby B, Renter DG, Dritz SS, Nelssen JL, Tokach MD. 2014. Effects of chlortetracycline and copper supplementation on antimicrobial resistance of fecal *Escherichia coli* from weaned pigs. Preventive Veterinary Medicine 114:231-246.

12. Zawack K, Li M, Booth JG, Love W, Lanzas C, Grohn YT. 2016. Monitoring antimicrobial resistance in the food supply chain and its implications for FDA policy initiatives. Antimicrobial Agents and Chemotherapy 60:5302-11.

13. Wellek S. 2010. Testing statistical hypotheses of equivalence and noninferiority, 2nd ed. Chapman and Hall/CRC Press; Taylor and Francis Group, LLC, Baco Raton, FL, USA.

14. Anderson SF, Maxwell SE. 2016. There's more than one way to conduct a replication study: Beyond statistical significance. Psychological Methods 21:1-12.
15. Love WJ, Zawack KA, Booth JG, Grhn YT, Lanzas C. 2016. Markov Networks of collateral resistance: National Antimicrobial Resistance Monitoring System surveillance results from *Escherichia coli* isolates, 2004-2012. PLoS Computational Biology 12:e1005160.

16. FDA CDER. 2001. Guidance for Industry. Statistical procedures for bioequivalence studies using a standard two treatment crossover design. 
[https://www.fda.gov/downloads/drugs/guidances/ucm070244.pdf](https://www.fda.gov/downloads/drugs/guidances/ucm070244.pdf) (Last accessed on Jul 1, 2017).

17. Karp B, Tate H, Plumblee JR, Dessai U, Whichard JM, Thacker EL, Hale KR, Wilson W, Friedman CR, Griffin PM, McDermott P. 2017. National Antimicrobial Resistance Monitoring System: two decades of advancing public health through integrated surveillance of antimicrobial resistance. 40:545-557.

18. FDA CVM. 2016. FDA Center for Veterinary Medicine. National Antimicrobial Resistance Monitoring System (NARMS) 2014 Integrated Report. 
[https://www.fda.gov/downloads/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/UCM528861.pdf](https://www.fda.gov/downloads/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/UCM528861.pdf). (Last accessed on Oct 1, 2017).

19. Garcia LS. 2010. Clinical microbiology procedures handbook. 3rd ed. Edited by L. S. Garcia. ASM Press. Washington, DC, USA.

20. WHO. 2003. Manual for identification and antimicrobial susceptibility testing of bacterial pathogens of public health importance in the developing world. 
[http://apps.who.int/iris/bitstream/10665/68554/1/WHO_CDS_CSR_RMD_2003.6.pdf](http://apps.who.int/iris/bitstream/10665/68554/1/WHO_CDS_CSR_RMD_2003.6.pdf) (Last assessed on Oct 1, 2017). World Health Organization, Geneva, Switzerland.
21. Ahmed S, Olsen JE, Herrero-Fresno A. 2017. The genetic diversity of commensal Escherichia coli strains isolated from non-antimicrobial treated pigs varies according to age group. PLoS One 12:e0178623.

22. Anderson MA, Whitlock JE, Harwood VJ. 2006. Diversity and distribution of Escherichia coli genotypes and antibiotic resistance phenotypes in feces of humans, cattle, and horses. Applied and Environmental Microbiology 72:6914-22.

23. Caugant DA, Levin BR, Selander RK. 1981. Genetic diversity and temporal variation in the E. coli population of a human host. Genetics 98:467-90.

24. Aslam M, Nattress F, Greer G, Yost C, Gill C, McMullen L. 2003. Origin of contamination and genetic diversity of Escherichia coli in beef cattle. Applied and Environmental Microbiology 69:2794-9.

25. Aslam M, Greer GG, Nattress FM, Gill CO, McMullen LM. 2004. Genetic diversity of Escherichia coli recovered from the oral cavity of beef cattle and their relatedness to faecal E. coli. Letters in Applied Microbiology 39:523-7.

26. Agga GE, Scott HM, Vinasco J, Nagaraja TG, Amachawadi RG, Bai J, Norby B, Renter DG, Dritz SS, Nelssen JL, Tokach MD. 2015. Effects of chlortetracycline and copper supplementation on the prevalence, distribution, and quantity of antimicrobial resistance genes in the fecal metagenome of weaned pigs. Preventive Veterinary Medicine 119:179-189.

27. Amachawadi RG, Scott HM, Vinasco J, Tokach MD, Dritz SS, Nelssen JL, Nagaraja TG. 2015. Effects of in-feed copper, chlortetracycline, and tylosin on the prevalence of transferable copper resistance gene, tcrB, among fecal Enterococci of weaned piglets. Foodborne Pathogens and Disease 12:670-678.
28. Amachawadi RG, Scott HM, Aperce C, Vinasco J, Drouillard JS, Nagaraja TG. 2015. Effects of in-feed copper and tylosin supplementations on copper and antimicrobial resistance in faecal enterococci of feedlot cattle. Journal of Applied Microbiology 118:1287-1297.

29. Feldpausch JA, Amachawadi R, Tokach MD, Scott HM, Dritz SS, Nagaraja TG, Goodband RD, Woodworth JC, DeRouchey JM. 2016. Effects of dietary Cu, Zn, and ractopamine-HCl on finishing pig growth performance, carcass characteristics, and antimicrobial susceptibility of enteric bacteria. Journal of Animal Science 94:73-74.

30. Welch BL. 1947. The generalisation of student's problems when several different population variances are involved. Biometrika 34:28-35.

31. Yuen KK. 1974. The 2-sample trimmed t for unequal population variances. Biometrika 61:165-170.

32. Koohmaraie M, Scanga JA, De La Zerda MJ, Koohmaraie B, Tapay L, Beskhlebnaya V, Mai T, Greeson K, Samadpour M. 2012. Tracking the sources of Salmonella in ground beef produced from nonfed cattle. Journal of Food Protection 75:1464-8.

33. Koohmaraie M, Arthur TM, Bosilevac JM, Guerini M, Shackelford SD, Wheeler TL. 2005. Post-harvest interventions to reduce/eliminate pathogens in beef. Meat Science 71:79-91.

34. Sofos JN, Kochevar SL, Bellinger GR, Buege DR, Hancock DD, Ingham SC, Morgan JB, Reagan JO, Smith GC. 1999. Sources and extent of microbiological contamination of beef carcasses in seven United States slaughtering plants. Journal of Food Protection 62:140-5.
35. Centers for Disease Control and Prevention. 2013. Antibiotic resistance threats in the United States, 2013. [http://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf](http://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf) (Last assessed on Feb 1, 2017).

36. Hohmann EL. 2001. Nontyphoidal salmonellosis. Clinical Infectious Diseases 32:263-269.

37. Onwuezobe IA, Oshun PO, Odigwe CC. 2012. Antimicrobials for treating symptomatic non-typhoidal *Salmonella* infection. Cochrane Database of Systematic Reviews:CD001167.

38. Centers for Disease Control and Prevention, National Antimicrobial Resistance Monitoring System (NARMS). 2016. 2014 Human Isolates Surveillance Report [https://www.cdc.gov/narms/pdf/2014-annual-report-narms-508c.pdf](https://www.cdc.gov/narms/pdf/2014-annual-report-narms-508c.pdf) (Last assessed on Feb 18, 2018).

39. FDA CVM. 2017. FDA Center for Veterinary Medicine. National Antimicrobial Resistance Monitoring System (NARMS) 2015 Integrated Report. [https://www.fda.gov/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/ucm059103.htm](https://www.fda.gov/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/ucm059103.htm). (Last accessed on Feb 18, 2018).

40. Volkova VV, DeMars Z. 2016. Short history of regulations and approved indications of antimicrobial drugs for food animals in the USA. Journal of Veterinary Pharmacology and Therapeutics 40:211-217.

41. Bryan A, Shapir N, Sadowsky MJ. 2004. Frequency and distribution of tetracycline resistance genes in genetically diverse, nonselected, and nonclinical *Escherichia coli*
strains isolated from diverse human and animal sources. Applied and Environmental Microbiology 70:2503-2507.

42. McDermott PF, Tyson GH, Kabera C, Chen Y, Li C, Folster JP, Ayers SL, Lam C, Tate HP, Zhao S. 2016. Whole-genome sequencing for detecting antimicrobial resistance in nontyphoidal Salmonella. Antimicrobial Agents and Chemotherapy 60:5515-20.

43. Tyson GH, McDermott PF, Li C, Chen Y, Tadesse DA, Mukherjee S, Bodeis-Jones S, Kabera C, Gaines SA, Loneragan GH, Edrington TS, Torrence M, Harhay DM, Zhao S. 2015. WGS accurately predicts antimicrobial resistance in Escherichia coli. Antimicrobial Agents and Chemotherapy 70:2763-9.

44. Santamaria J, Lopez L, Soto CY. 2011. Detection and diversity evaluation of tetracycline resistance genes in grassland-based production systems in Colombia, South america. Frontiers in Microbiology 2:252.

45. Cormier AC, Chalmers G, McAllister TA, Cook S, Zaheer R, Scott HM, Booker C, Read R, Boerlin P. 2016. Extended-spectrum-cephalosporin resistance genes in Escherichia coli from beef cattle. Antimicrobrial Agents and Chemotherapy 60:1162-1163.

46. Roberts MC. 2005. Update on acquired tetracycline resistance genes. FEMS Microbiology Letters 245:195-203.

47. CLSI. 2015. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard - 10th ed. (CLSI document M07-A10) Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, PA, USA.

48. Ruxton GD. 2006. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. Behavioral Ecology 17:688-690.
49. Yuen KK, Dixon WJ. 1973. Approximate behavior and performance of 2-sample trimmed \( t \) Biometrika 60:369-374.

50. Hurley PJ. 2014. A concise introduction to logic. 12th ed. Cengage Learning, Stamford, CT, USA.

51. Cohen J. 1992. A power primer. Psychological Bulletin 112:155-159.

52. Cohen J. 1988. Statistical power analysis for the behavioral sciences. 2nd ed. Lawrence Erlbaum Associates, New York, NY, USA.

53. National Antimicrobial Resistance Monitoring System (NARMS). 2016. Manual of Laboratory Methods, 3rd ed. https://www.fda.gov/downloads/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/UCM528831.pdf (Last assessed on Oct 1, 2017).
Figure legends.

Figure 1. Schematic representation of testing statistical equivalence of the bacterial antimicrobial susceptibility estimates from two sampling or sample-processing approaches used in a stratum. If a practically relevant difference between the average log$_2$(MIC) estimates from the two approaches is equal to or larger than $\Delta_{min}$ (e.g. $\Delta_1$), the hypothesis of statistical non-equivalence of the estimates will be rejected, signaling equivalence of the MIC-data from the two approaches. If a practically relevant difference is smaller than $\Delta_{min}$ (e.g. $\Delta_2$), the hypothesis of statistical non-equivalence of the estimates will be accepted. The maximum possible difference between the log$_2$(MIC) values from the two approaches is $\Delta_{max}$ ($\Delta_{max} = \max_{ MIC_1 \in Y_1 \text{ and } MIC_2 \in Y_2} | \log_2 MIC_1 - \log_2 MIC_2 |$) for the isolate-sets $Y_1$ and $Y_2$.

Figure 2. Testing statistical equivalence of the estimates of phenotypic antimicrobial susceptibility of nontyphoidal *Salmonella enterica* subsp. *enterica* isolates from the ground or trimmed meat samples ($n=310$ for 2013 and $n=344$ for 2014) vs. cecal-contents samples ($n=435$ for 2013 and $n=318$ for 2014) from cattle in the processing plants in the U.S. in 2013 (A) and 2014 (B). The data were collected by the NARMS.

Aminoglycosides: GEN – gentamicin, STR – streptomycin. $\beta$-lactams aminopenicillins: AMC - amoxicillin-clavulanic acid, AMP - ampicillin. $\beta$-lactams cephems: AXO - ceftriaxone, FOX – cefoxitin, TIO - ceftiofur. Macrolides: AZI - azithromycin. Phenolics: CHL - chloramphenicol.

Quinolones: CIP - ciprofloxacin, NAL- nalidixic acid. Sulfonamides: COT - trimethoprim-sulfamethoxazole, FIS- sulfisoxazole. Tetracyclines: TET - tetracycline.

Figure 3. Testing statistical equivalence of the estimates of phenotypic antimicrobial susceptibility of nontyphoidal *Salmonella enterica* subsp. *enterica* isolates from urine ($n=144$), fecal ($n=1,495$), and blood ($n=181$) samples of humans in the U.S. in 2015. (A) Urine vs. blood...
isolates. (B) Fecal vs. blood isolates. (C) Urine vs. fecal isolates. The data were collected by the NARMS.

Aminoglycosides: GEN - gentamicin, STR - streptomycin. β-lactams aminopenicillins: AMC - amoxicillin-clavulanic acid, AMP - ampicillin. β-lactams cephems: FOX - cefoxitin, TIO - ceftiofur. Macrolides: AZI - azithromycin. Phenics: CHL - chloramphenicol. Quinolones: CIP - ciprofloxacin, NAL - nalidixic acid. Sulfonamides: FIS - sulfisoxazole. Tetracyclines: TET - tetracycline.

**Figure 4.** Testing statistical equivalence of the estimates of phenotypic antimicrobial susceptibility of *E. coli* in cattle fecal pads (*n*=32). The sample-processing approaches compared were testing susceptibility of 4 *E. coli* isolates obtained from one aliquot of the pad vs. testing susceptibility of 4 *E. coli* isolates each obtained from a different aliquot of the pad, with the aliquots collected from locations spread along the longest axis of the pad. Aminoglycosides: TOB - tobramycin. β-lactams aminopenicillins: AMP - ampicillin, A/S2 - ampicillin/sulbactam 2:1 ratio. β-lactams cephems: FAZ - cefazolin, TIO - ceftiofur. Macrolides: TIL - tilmicosin, TUL - tulathromycin. Phenics: FFN - florfenicol. Tetracyclines: CTET - chlortetracycline, MIN - minocycline, OXY - oxytetracycline, TET - tetracycline.
Table captions.

**Table 1.** Antimicrobial drugs susceptibility to which was tested in nontyphoidal *Salmonella enterica* subsp. *enterica* isolates from ground or trimmed meat samples and from cecal-contents samples collected from cattle in the processing plants in the U.S. in 2013 and 2014. The data were collected by the U.S. National Antimicrobial Resistance Monitoring System.

**Table 2.** Antimicrobial drugs susceptibility to which was tested in nontyphoidal *Salmonella enterica* subsp. *enterica* isolates from urine, fecal, and blood samples of humans in the U.S. in 2015. The data were collected by the U.S. National Antimicrobial Resistance Monitoring System.

**Table 3.** Antimicrobial drugs susceptibility to which was tested in *Escherichia coli* isolates from cattle fecal pads.
Table 1. Antimicrobial drugs susceptibility to which was tested in nontyphoidal *Salmonella enterica* subsp. *enterica* isolates from ground or trimmed meat samples and from cecal-contents samples collected from cattle in the processing plants in the U.S. in 2013 and 2014. The data were collected by the U.S. National Antimicrobial Resistance Monitoring System.

| Antimicrobial drug class | Sub-class | Combinatory formulation (Yes/No) | Drug |
|--------------------------|-----------|-----------------------------------|------|
| Aminoglycosides          | -micins   | No                                | Gentamicin |
|                          | -mycins   | No                                | Kanamycin* (tested in 2013 only) |
| β-lactams                | Aminopenicillins with β-lactamase inhibitors | Yes | Amoxicillin with clavulanic acid |
|                          | Aminopenicillins | No | Amoxicillin |
|                          | Cephs     | No | Ceftriaxone* |
|                          |           | No | Cefoxitine |
|                          |           | No | Cefidoxur |
| Macrolides               | Azalides  | No | Azithromycin |
| Phenicol                | N/A       | No | Chloramphenicol |
| Quinolones              | Fluoroquinolones | No | Ciprofloxacin |
|                          | Quinolones | No | Nalidixic acid |
| Sulfonamides            | N/A       | Yes | Sulfamethoxazole with trimethoprim |
| Tetracyclines           | Tetracyclines | No | Sulfisoxazole |

* Due to a low variability in the data for kanamycin in 2013 and ceftriaxone in 2014, the equivalence testing could not be performed.
Table 2. Antimicrobial drugs susceptibility to which was tested in nontyphoidal *Salmonella enterica* subsp. *enterica* isolates from urine, fecal, and blood samples of humans in the U.S. in 2015. The data were collected by the U.S. National Antimicrobial Resistance Monitoring System.

| Antimicrobial drug class | Sub-class | Combinatory formulation (Yes/No) | Drug                          |
|-------------------------|-----------|----------------------------------|-------------------------------|
|                         | Aminoglycosides |                                   |                               |
|                         | -micins     | No                               | Gentamicin                    |
|                         | -mycins     | No                               | Streptomycin                  |
| β-lactams               | Penicillins including amino-, carboxy-, and ureido-, with β-lactamase inhibitors | Yes                            | Amoxicillin with clavulanic acid |
|                         | -Cephalosporins |                                   |                               |
|                         | Penicillins including amino-, carboxy-, and ureido-Cephalosporins | No                            | Ampicillin                    |
|                         | Penicillins including amino-, carboxy-, and ureido-Cephalosporins | No                            | Cefoxitin                     |
|                         | Penicillins including amino-, carboxy-, and ureido-Cephalosporins | No                            | Cefotaxime                    |
| Macrolides              | Azalides    | No                               | Azithromycin                  |
| Phenicols               | N/A         | Yes                              | Chloramphenicol               |
| Quinolones              | Fluoroquinolones |                                   | Ciprofloxacin                 |
|                         | Quinolones  | No                               | Nalidixic acid                |
| Sulfonamides            | N/A         | Yes                              | Sulfisoxazole                 |
| Tetracyclines           | Tetracyclines |                                   | Tetracycline                  |

Variability in the log$_2$(MIC) data from urine, fecal, and blood samples was sufficient to perform the statistical equivalence testing.

β-lactams | Penicillins including amino-, carboxy-, | Yes | Piperacillin with tazobactam |

The statistical equivalence testing could not be performed because of a low number of each urine and blood samples (n<10) tested with the antimicrobial.

β-lactams | Penicillins including amino-, carboxy-, | Yes | Piperacillin with tazobactam |
| and ureido-, with β-lactamase inhibitors | constant |
|-----------------------------------------|----------|
| Carbapenems                             | No       |
| Cephems with β-lactamase inhibitors     | Yes      |
| Cephems                                 | No       |
| Monobactam                              | No       |
|                                         | Imipenem |
|                                         | Cefotaxime with clavulanic acid |
|                                         | Cefazidime with clavulanic acid |
|                                         | Cefepime |
|                                         | Cefotaxime |
|                                         | Cefquinome |
|                                         | Cefazidime |
|                                         | Aztreonam |
Table 3. Antimicrobial drugs susceptibility to which was tested in *Escherichia coli* isolates from cattle fecal pads.

| Antimicrobial drug class | Sub-class | Combinatory formulation (Yes/No) | Drug |
|--------------------------|-----------|-----------------------------------|------|
| **Aminoglycosides**      | -myncins  | No                                | Tobramycin |
| **β-lactams**             | Penicillins including amino-, carboxy-, and ureido- | No | Ampicillin |
| Penicillins including amino-, carboxy-, and ureido-, with β-lactamase inhibitors | Yes | Ampicillin with sulbactam 2:1 ratio |
| Cephalosporins           | No        | Cefazolin                          |
|                          | No        | Ceftiofur                          |
| **Macrolides**           | Macrolides | No                                | Tilmicosin |
|                          | No        | Tulathromycin                      |
| **Phenicols**            | N/A       | No                                | Florfenicol |
| **Tetracyclines**        | Tetracyclines | No                                | Chlorotetracycline |
|                          | No        | Minocycline                        |
|                          | No        | Oxytetracycline                    |
|                          | No        | Tetracycline                       |

**Variability in the log$_2$(MIC) data from the two sample-processing approaches was sufficient to perform the statistical equivalence testing.**

| Aminoglycosides      | -myncins  | -myncins | -myncins | Amikacin |
|----------------------|-----------|----------|----------|----------|
|                      | No        | No       | No       | Gentamicin |
|                      |           |          |          | Neomycin  |
|                      |           |          |          | Spectinomycin |
| **β-lactams**        | Penicillins including amino-, carboxy-, and ureido- | No | Penicillin |
| Penicillins including amino-, carboxy-, and ureido-, with β-lactamase inhibitors | Yes | Piperacillin |
| Carbapenems          | Yes       | Yes      | No       | Doripenem |

**Variability in the log$_2$(MIC) data from the two sample-processing approaches was insufficient to perform the statistical equivalence testing.**

|  |  |  |  |  |
| Class          | Species       | Ertapenem | Imipenem | Meropenem | Aztreonam | Cefepime | Ceftazidime | Ceftriaxone |
|----------------|---------------|-----------|----------|-----------|-----------|----------|-------------|-------------|
| Monobactams    | No            | No        | No       | No        | No        | No       | No          | No          |
| Cepham         | No            | No        | No       | No        | No        | No       | No          | No          |
| Macrolides     | No            | No        | No       | No        | No        | No       | No          | No          |
| Nitrofurans    | No            | No        | No       | No        | No        | No       | No          | No          |
| Pleuromutilins | No            | No        | No       | No        | No        | No       | No          | No          |
| Lincosanides   | No            | No        | No       | No        | No        | No       | No          | No          |
| Quinolones     | No            | No        | No       | No        | No        | No       | No          | No          |
| Fluoroquinolones | No         | No        | No       | No        | No        | No       | No          | No          |
| Sulfonamides   | N/A           | No        | No       | No        | No        | No       | No          | No          |
| Tetracyclines  | Glycylcyclines| No        | No       | No        | No        | No       | No          | No          |
| Macrolides     | N/A           | No        | Tylosin  | No        | Nitrofurantoin | | |
| Nitrofurans    | N/A           | No        | No       | No        | No        | No       | No          | No          |
| Pleuromutilins | N/A           | No        | Pleuromutilin | No       | Tiamulin | No       | No          | No          |
| Lincosanides   | N/A           | No        | Clindamycin | No       | No       | No       | No          | No          |
| Quinolones     | Quinolones    | No        | Nalidixic acid | No       | No       | No       | No          | No          |
| Fluoroquinolones | No          | No        | Ciprofloxacin | No       | No       | No       | No          | No          |
| Sulphadimethoxine | No       | Yes       | Danofloxacin | No       | No       | No       | No          | No          |
| Sulfamethoxazole with trimethoprim | No       | Yes       | Enrofloxacin | No       | No       | No       | No          | No          |
| Tigecycline    | No            | No        | No        | No        | No        | No       | No          | No          |