



RESEARCH ARTICLE

**REVISED** **Representing antibiotic relationships using measurements of efficacy against clinical isolates [version 2; peer review: 1 approved, 1 approved with reservations]**

Liam Shaw

Nuffield Department of Medicine, University of Oxford, Oxford, UK

**V2** **First published:** 28 May 2019, 4:86  
<https://doi.org/10.12688/wellcomeopenres.15304.1>  
**Second version:** 04 Oct 2019, 4:86  
<https://doi.org/10.12688/wellcomeopenres.15304.2>  
**Latest published:** 17 Sep 2020, 4:86  
<https://doi.org/10.12688/wellcomeopenres.15304.3>

**Abstract**

**Introduction.** Antimicrobial resistance (AMR) is a worrying and confusing problem for both patients and medical professionals. Current antimicrobial nomenclature makes it difficult to understand how different antibiotics are related to one another, and their empirical uses and efficacy. Here, I attempt to address this problem by using experimental data from hundreds of thousands of clinical isolates to infer relationships between antibiotics and represent them with simple diagrams.

**Methods.** The minimum inhibitory concentration (MIC) of a bacterial isolate for a given antibiotic is defined as the lowest concentration that prevents visible growth. Measuring MICs for multiple antibiotics using the same isolate implicitly records the relationships of the antibiotics for a given species. The basic principle is that antibiotics with similar mechanisms of action should give rise to similar mechanisms of resistance, so should have correlated MICs across large numbers of isolates. This information can then be used to calculate distances between antibiotics based on pairwise correlations of their rank-ordered MICs. I apply this approach to a large historical AMR surveillance dataset (the Pfizer ATLAS surveillance dataset, 2004-2017).

**Results.** I demonstrate that clustering antibiotics in this way allows a simple visual comparison of how similar antibiotics are to each other based on their efficacy within a species. The resulting visualizations broadly recapitulate antibiotic classes. They also clearly show the dramatic effects of combining beta-lactam antibiotics with beta-lactamase inhibitors, as well as highlighting antibiotics which have unexpected correlations in MICs that are not predicted from their chemical similarities alone.

**Conclusion.** Large AMR surveillance datasets can be used in a hypothesis-free manner to show relationships between antibiotics based on their real-world efficacy. The resulting 'data-driven'

**Open Peer Review****Reviewer Status**

Invited Reviewers

1

2

**version 3**

(revision)

17 Sep 2020

**version 2**

(revision)

04 Oct 2019



report



report

**version 1**

28 May 2019



report

1. **Mark H. Wilcox**, University of Leeds, Leeds, UK

2. **Benno H Ter Kuile**, University of Amsterdam, Amsterdam, The Netherlands

Any reports and responses or comments on the article can be found at the end of the article.

diagrams could be used as simple visual aids to explain the complexity of antibiotic relationships and AMR.

### Keywords

Antimicrobial resistance, AMR, data reuse, AMR surveillance, antibiotics, cross-resistance, chemical similarity

**Corresponding author:** Liam Shaw ([liam.philip.shaw@gmail.com](mailto:liam.philip.shaw@gmail.com))

**Author roles: Shaw L:** Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** Prior to the Wellcome Data Re-use Prize (AMR surveillance), I was employed as a consultant researcher by the Open Data Institute (ODI) funded by the Wellcome Trust to work on the 90-day pilot project 'Antimicrobial Resistance Surveillance: Sharing industry data'. This project assessed the current availability of data and subsequently led to the ATLAS dataset from Pfizer being released through the Wellcome Trust. However, the analysis presented here is my own independent work and does not represent the views of Pfizer, ODI, or the Wellcome Trust.

**Grant information:** This work was a winning entry of the 2019 Wellcome Data Re-use Prize in AMR surveillance and is eligible for publication on Wellcome Open Research.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2019 Shaw L. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Shaw L. **Representing antibiotic relationships using measurements of efficacy against clinical isolates [version 2; peer review: 1 approved, 1 approved with reservations]** Wellcome Open Research 2019, 4:86 <https://doi.org/10.12688/wellcomeopenres.15304.2>

**First published:** 28 May 2019, 4:86 <https://doi.org/10.12688/wellcomeopenres.15304.1>

**REVISED Amendments from Version 1**

No major changes to content apart from some modification of potentially misleading statements. The abstract has been rewritten to clarify the central premise and remove a potentially misleading claim about 'combination therapies' which arose from a lax use of the term. Similarly, the section of Results relating to species-specific effects for beta-lactam and beta-lactamase inhibitors has been tightened up to avoid confusion. Finally, a paragraph has been added in the Discussion cautioning against using the dendrograms for inference of the effects of dual-antibiotic combinations.

**Any further responses from the reviewers can be found at the end of the article**

**Introduction**

Antimicrobial resistance (AMR) is a set of relationships. A bacterium is not simply 'resistant'; it is resistant to an antibiotic. The problem of AMR is therefore an umbrella term for these many different pairwise relationships, each relationship specific to particular bacteria and particular antibiotics. We need approaches to represent the considerable complexity of these relationships in an understandable way.

It is undeniable that we need new antibiotics to treat patients with multi-drug-resistant infections<sup>1</sup>; we have arguably also failed to keep track of or use existing antibiotics properly. A major driver of AMR is the inappropriate use of antibiotics. One meta-analysis estimated the pooled rate of inappropriate empirical treatment across available studies at 25.5–31.8%<sup>2</sup>. Even recommended treatment is remarkably varied. For example, one survey found a diversity of recommendations for common conditions such as pneumonia across 105 NHS Trusts in England, in both the antimicrobial agent and the length of the treatment<sup>3</sup>. Surprisingly little evidence is also available for the claim that failing to complete the full length of a physician-recommended course of antibiotics leads to the development of resistance<sup>4</sup>.

These examples demonstrate that despite being conceptually simple ('antibiotic use leads to resistance'), we currently lack a great deal of knowledge about AMR. Unfortunately, it seems that even having an awareness of existing knowledge about how different antibiotics relate to each other, and of the bacterial mechanisms which confer resistance, is challenging. A survey of Master of Pharmacy students at the University of Brighton, UK, found that around a third of third- and fourth-year students incorrectly agreed with a statement that 'bacterial beta-lactamase enzymes inactivate aminoglycoside antibiotics'<sup>5</sup>. If future pharmacists can be so poorly taught about the mechanisms of AMR, what hope is there for the rest of us?

It is not an original observation that communicating AMR is challenging. Suggestions for improving this tend to focus around discourse, recommending caution about the use of catastrophist terms like 'post-antibiotic apocalypse'<sup>6</sup>, moving away from the metaphor of a 'war' on 'superbugs'<sup>7</sup>, or introducing the concept of an 'antibiotic footprint' analogous to a carbon footprint<sup>8</sup>. The problem I seek to address here is not as general. I am specifically interested in the problem of imparting knowledge: how to convey the relationships between different antibiotics to non-specialists.

In order to do this, I use a large open dataset of experimental measurements on bacterial isolates taken from clinical settings. This dataset offers the opportunity to derive relationships between antibiotics that do not rely on knowledge of the underlying pharmacology, but are based on their actual efficacy. Simple visual representations of these show the similarities of antibiotics, giving an intuitive handle on the complexity of the relationships between them.

**Dataset**

The Antimicrobial Testing Leadership and Surveillance (ATLAS) dataset was made publicly available as part of the Wellcome Data Re-use Prize<sup>9</sup>. The ATLAS dataset comes from Pfizer's overall programme to fulfill regulatory requirements and support appropriate use measures for both marketed antibiotics and antibiotics in development. It contains 633,820 isolates from 73 countries between 2004 and 2017. Each isolate has minimum inhibitory concentrations measured against multiple antibiotics (median 11, range: 5–23). For further details on the dataset and previous publications, see the ATLAS website<sup>10</sup>.

Here, I report results for only the five most represented bacterial species in the dataset: *Staphylococcus aureus* ( $N = 113,693$ ); *Escherichia coli* ( $N = 80,500$ ); *Klebsiella pneumoniae* ( $N = 64,296$ ); *Pseudomonas aeruginosa* ( $N = 61,799$ ); and *Enterobacter cloacae* ( $N = 39,391$ ). Analysis includes all years and countries.

**Methods**

This article developed from a submission to the Wellcome Data Re-use Prize (AMR surveillance). The original submission contains further code and analyses<sup>11</sup>. Supplementary material for this article is available as *Extended data* on Figshare<sup>12</sup>.

**Understanding minimum inhibitory concentrations (MICs)**

The ATLAS dataset contains information on the MIC of isolates tested against specific antibiotics. The MIC is the lowest concentration of an antibiotic that prevents visible growth. Consider the following table of isolate information:

Isolate	Species	$A_1$	$A_2$	$A_3$
X	<i>E. coli</i>	1	<0.5	4
Y	<i>E. coli</i>	2	<0.5	8

This information tells us that isolate X is an *E. coli* strain which cannot grow in concentrations of  $A_1$  of 1 mg/L, concentrations of  $A_2$  of 0.5 mg/L or more, and concentrations of  $A_3$  of 4 mg/L. Here, we are interested in the correlation between antibiotic MICs. For example,  $A_1$  and  $A_3$  appear to be correlated (on the basis of two isolates), because when the MIC for  $A_1$  halves so does the MIC of  $A_3$ .

MICs are most commonly used to determine whether a bacteria is to be considered as 'resistant' to an antibiotic or not; an MIC is chosen based on clinical, experimental, and/or epidemiological data to mark the boundary between 'susceptible' and 'resistant' bacteria. Most publications on AMR report only resistance rates for single antibiotics. Sometimes, cross-resistance rates are considered (when a bacteria is resistant to two antibiotics at the

same time). However, this approach still loses information from the full MIC data. Using the correlation of actual MIC values between pairs of antibiotics, and performing the analysis on thousands of isolates for all available antibiotics, should give us a definition of how similar antibiotics are based on their measured activities on real-world strains.

MICs are slightly unwieldy and complex variables to analyse statistically because they have unusual properties. For example, consider the MIC scale:

$$<1, 1, 2, 4, 8, >8$$

These values are ‘interval-censored’: the continuum of underlying true MICs is restricted by measurement into a scale of intervals, which contain ranges of values.

**Example:** if a concentration of 2 did not inhibit growth but a concentration of 4 did, the isolate has a ‘true’ MIC of  $x$  where  $2 \leq x < 4$ , but will have a recorded MIC of 4.

Also, the maximum value of the scale is ‘left-censored’: an MIC of  $>8$  gives no information about the ‘true’ MIC apart from that it is greater than 8. Correspondingly, the minimum value is ‘right-censored’.

Approaches such as regression models that account for interval-censored data<sup>13</sup> have been applied to model MICs (e.g. 14). However, here I will use a naive and simple ‘solution’: I simply ignore the censoring. While not optimal, this is adequate for looking at large-scale correlations of MICs at a global level.

### MIC-based similarity scores

First, I order all possible values of MICs across the entire dataset. This step requires making an arbitrary choice about how to treat left/right-censored intervals i.e. those which feature an inequality.

Here, I make a particular ordering choice which is best explained with an example of ordered MIC values:

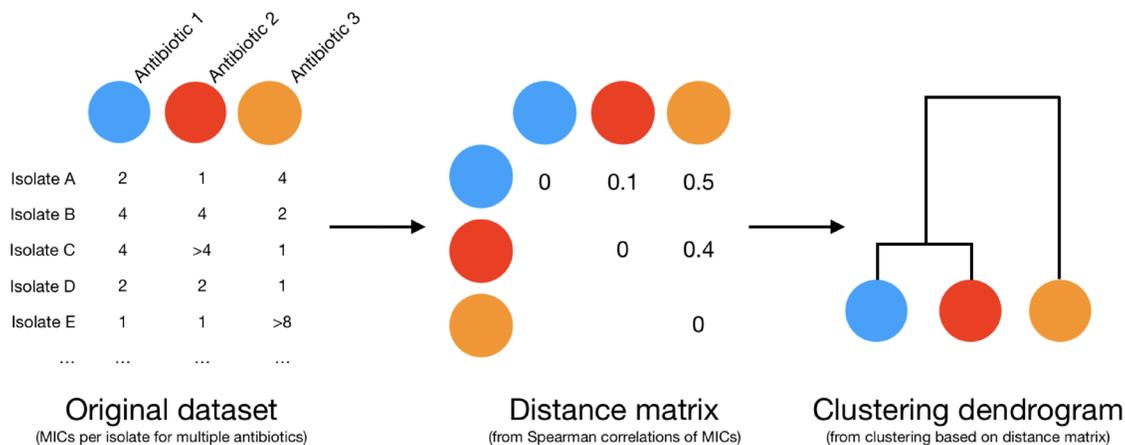
$$<0.5, <1, 1, >1, 2, \dots, etc.$$

The reason that this is not ideal is that it assumes that ‘ $>1$ ’ is less than ‘2’, but clearly a value of 2.1 could be ranked as ‘ $>1$ ’ in one scale and ‘2’ in another. Thus, it is possible that identical isolates could be tested in different scales and ranked differently. Fortunately, the standardized methodology used in the ATLAS dataset means that for a given antibiotic, the scales are usually consistent, i.e. they will not contain potentially contradictory values. Even the subset which do, this should not qualitatively affect the rank correlations as it amounts to swapping (or removing) particular ranks.

Once MICs have been ranked, these ranked MIC values can then be used to calculate a Spearman rank correlation between antibiotics for a species. Figure 1 shows an overview of the procedure. Not all isolates are tested against all antibiotics, so the Spearman rank correlation  $\rho(A_1, A_2)$  between the MICs for antibiotics  $A_1$  and  $A_2$  is calculated only on the subset of isolates with measured MICs for both  $A_1$  and  $A_2$ . Missing entries in the correlation matrix (where the antibiotics have never been tested simultaneously on any isolates of the species) are set to  $\rho = 0$ . These pairwise correlations are then used to create a distance matrix  $d(A_1, A_2) = 1 - \rho(A_1, A_2)^2$  (see below). The distance matrix is then used to cluster antibiotics into a dendrogram with `hclust` in R (v3.5.1)<sup>15</sup> Ward’s hierarchical grouping algorithm<sup>16</sup>. Dendrograms are plotted with `ggdendro` (v0.1-20)<sup>17</sup> and `ggplot2` (v3.1.0)<sup>18</sup>.

### Antibiotic classes

I obtained classification codes for each antibiotic by searching in the WHO ATC/DDD index 2019<sup>19</sup>. The Anatomical Therapeutic Chemical (ATC) Classification System gives a



**Figure 1. Schematic for antibiotic MIC correlation method.** All possible pairwise correlations of MICs for pairs of antibiotics are calculated, and these are used to build a distance matrix and then a clustering dendrogram. Antibiotic 1 and Antibiotic 2 are strongly correlated (see dataset on the left), so have a smaller distance between them (0.1) than Antibiotic 1 and Antibiotic 3 (0.5).

hierarchical classification of drugs according to their therapeutic properties. For example, ciprofloxacin has the code J01MA02, which should be read as:

- J Anti-infectives for systemic use
  - J01 Antibacterials for systemic use
    - J01M Quinolone antibacterials
      - J01MA Fluoroquinolones
        - J01MA02 Ciprofloxacin

Some antibiotics in the ATLAS dataset are not yet classified by ATC. I classified combinations of J01D beta-lactams and avibactam as J01DI ('Other beta-lactams, Other').

### Chemical similarity scores

In order to compare the similarity scores generated from MICs in the ATLAS dataset, I also collected simplified molecular-input line-entry system (SMILES) representations for all single-molecule antibiotics (excluding combination antibiotics e.g. piperacillin/tazobactam) from their respective pages on Wikipedia (2nd May 2019). I then used the R package `rdck` (v3.4.7.1)<sup>20</sup> and followed the instructional vignette<sup>21</sup> to parse these with `parse.smiles`, obtain the chemical fingerprint with `get.fingerprint` and then construct a distance matrix with `fp.sim.matrix` using the Tanimoto similarity metric (in a quirk of disciplines, this is the same as the – to me – more familiar Jaccard coefficient<sup>22</sup>). This gives a score for each pair of antibiotics based on how many chemical features they share.

For ten single-molecule antibiotics that had >5,000 measurements for each of the top five bacterial species (ampicillin, ceftaroline, ceftazidime, ceftriaxone, doripenem, imipenem, levofloxacin, meropenem, minocycline, tigecycline) I calculated the correlation between the chemical similarity matrix and the MIC-based similarity matrix using `mantel.randtest` from the R package `ade4` (v1.7-13)<sup>23</sup>. Table 1 shows that the strength of the

correlation was increased for all species when the MIC-based Spearman correlation was squared, so I used this to produce the MIC-based distance matrix between antibiotics.

### Results

Clustering antibiotics based on pairwise distances calculated from correlations of their MICs broadly captures known antibiotic classes (Figure 2). The layout of a dendrogram gives a clear visualisation of when antibiotics are very similar in terms of their efficacy on clinical isolates of a bacterial species. They also allow a quick comparison of antibiotics in terms of their action across different species.

Although the whole point of these visualisations is to convey a general overview of the relationships from the dataset alone, independent of any previous pharmacological knowledge, it is perhaps worth pointing out a few examples of how they connect to (and convey) what is known about AMR.

**Similarities within antibiotic classes.** Tigecycline and minocycline (both members of the tetracycline class) cluster together in all five species. This is unsurprising, as tetracyclines all have the same mechanism of action: they bind to 16S ribosomal RNA (rRNA) in the ribosome and thus inhibit protein synthesis<sup>24</sup>. Mutations in the 16S rRNA gene which confer resistance to e.g. tigecycline typically also do so for all tetracyclines.

**Similarities across antibiotic classes.** Ampicillin and amoxicillin clavulanate are clearly very similar in *E. coli* (Figure 2b). However, the community usage of the J01CA subclass of antibiotics (which includes ampicillin) across 12 European countries does not correlate well with the prevalence of resistance to ampicillin<sup>11</sup>. This might seem surprising, but even without the pharmacological knowledge that ampicillin and amoxicillin are very similar in activity we could use this clustering of antibiotics to verify that amoxicillin clavulanate usage (in the J01CR class) correlates with ampicillin resistance<sup>11</sup>. This underlines the importance of being aware of the mechanisms of AMR rather than just fitting statistical models to data.

**Species-specific effects.** The beta-lactam antibiotic ceftazidime has limited activity against *S. aureus*. Therefore, combining it with avibactam (a non-beta-lactam beta-lactamase inhibitor with little to no antibiotic activity of its own) has little additional effect; the two treatments cluster close together (Figure 2a). However, the dramatic effect of using avibactam in combination with ceftazidime in Gram-negative species such as *E. coli* and *K. pneumoniae* is readily apparent; ceftazidime/avibactam is shown far apart from ceftazidime in these species (Figure 2b–2e).

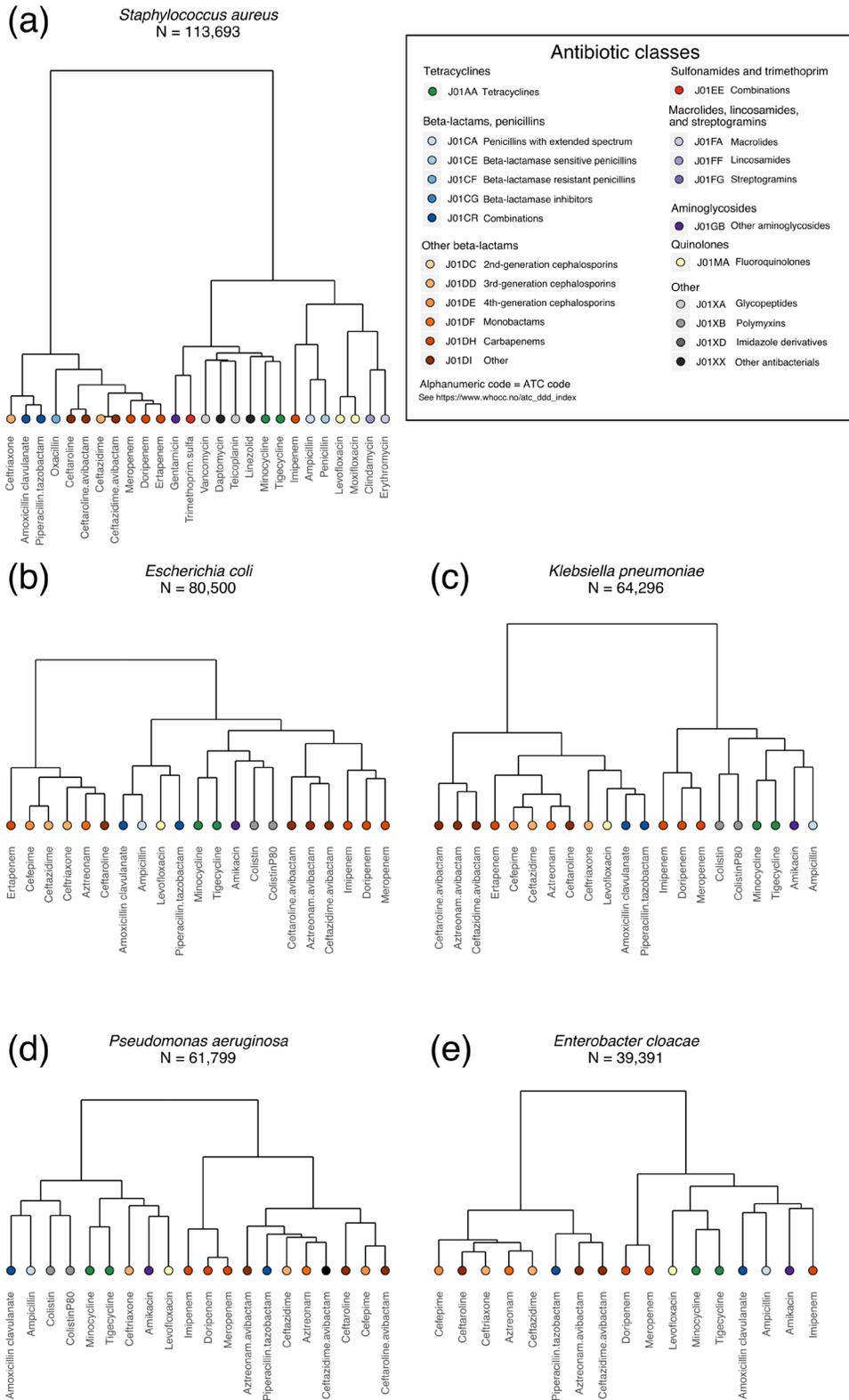
### Discussion

I have outlined how a large AMR surveillance dataset can be used to characterise relationships between antibiotics with no prior knowledge of pharmacology. I have shown that computing distances between antibiotics based on their correlations across clinical isolates shows a good correlation with chemical

**Table 1. Mantel correlations between chemical similarities and MIC-based similarities using linear ( $\rho$ ) or squared ( $\rho^2$ ) Spearman correlations, for ten antibiotics.**

There is a significant strong association between the similarities. Squaring the MIC-based correlations increases the strength of the correlation with chemical similarities.  $M$  is the observed correlation between distance matrices using the Mantel test,  $p$  is the p-value.

Species	$\rho$		$\rho^2$	
	$M$	$p$	$M$	$p$
<i>S. aureus</i>	0.319	0.019	0.396	0.006
<i>E. coli</i>	0.516	0.001	0.541	0.001
<i>K. pneumoniae</i>	0.517	0.002	0.566	0.001
<i>P. aeruginosa</i>	0.602	0.001	0.729	0.001
<i>E. cloacae</i>	0.585	0.001	0.672	0.001



**Figure 2. Clustering of antibiotics according to their MIC correlations within the top five species in the ATLAS dataset.** Similar antibiotics generally cluster together across all five species (a)–(e). However, the precise pattern of this clustering varies by species. Colour of points indicates the class of antibiotic, as judged using the WHO ATC code for the drug (see legend, top right). Greater vertical distance represents greater separation (vertical axis conserved across subfigures). Only antibiotics with >5,000 test results for a species are included. Total number of isolates is shown in title of each subfigure.

similarity scores (Table 1) and can be used to produce clear visualisations of relationships in AMR (Figure 2).

These visualisations make clear one of the challenges of AMR: use of one antibiotic can lead to increasing resistance to other antibiotics as well due to their chemical similarity. By representing antibiotics closer together when their MICs are correlated, the diagrams convey the fact that bacteria increasing their MIC to one specific antibiotic can be associated with an increased MIC for other similar antibiotics. Whole branches of the tree may become 'off-limits' once resistance develops, rather than antibiotics being picked off one by one. Conceivably, simple displays based on these visualisations could be used in multiple settings: when explaining the importance of antibiotic stewardship in hospitals; when recommending courses of antibiotics for patients; or even included on antibiotic packaging to underline the difference between antibiotics and other classes of drugs.

### Limitations

As noted, I have not attempted to perform any serious statistical treatment of the MICs and have simply sorted them into rank-order, acknowledging that this can lead to inconsistent measured MICs from identical true MICs. It is possible that improving this step could improve the quality of the inferred relationships. Furthermore, aside from a crude attempt to check broad consistency with chemical similarity scores, I have not attempted to prove that these results have strong meaning beyond simply representing the data. Because of the fact that not all antibiotics are tested against all others, specific pairwise comparisons may have little meaning, and the broad structure of the dendrogram may be driven by a biased pattern of testing.

This method implicitly assumes a 'one size fits all' approach for a given species. In reality, there are typically multiple mechanisms of resistance to antibiotics. For the beta-lactams, over 2,000 separate resistance genes are known. Furthermore, high correlations between antibiotic MICs for a particular bacterial species could be driven by multiple factors. Possible scenarios include but are not limited to: the antibiotics have similar activity and therefore similar resistance mechanisms; the antibiotics are commonly used together, so because resistances are selected at the same time they are commonly associated despite different resistance mechanisms; certain strains possess multiple mechanisms of drug-resistance and drive the high correlation. I have not attempted to disentangle these causes here.

I caution against inferring potential treatment interactions between antibiotics based on these dendrograms. In general, two antibiotics used in combination can be additive, synergistic (i.e. greater than the sum of their individual effects), or antagonistic (i.e. less than the sum of their individual effects). Phenotypic testing of combinations is necessary to understand these interactions.

Even leaving this point aside, the clinical relevance of what I have done is unclear. The term 'resistance' is tied to a specific clinical meaning: that standard treatment at recommended doses will likely not be effective in a patient infected with

such a strain. The MIC offers a proxy for determining when this is true, but is not in itself a clinical parameter. However, I would argue that using the whole range of the MIC distribution to produce the visualisations is more statistically justifiable to get an overview of the range of relationships between antibiotics. Correlations could be calculated based on binary cross-resistance proportions and equivalent diagrams plotted, but proportions of cross-resistant isolates may be very small.

Finally, the underlying data has its own limitations. Despite being of high quality and standardisation, MIC phenotypes can vary between experiments. For example, colistin adheres to plastics, so sometimes MICs are tested after adding the surfactant polysorbate 80 (P80), which can alter the MIC of Gram-negative species<sup>25</sup>, as seen in the data here (Figure 2b–d).

### Future work

This work is intended only as a sketch of a possible way to represent relationships between antibiotics using large datasets from clinical settings. I am sure the elementary approach I have outlined here could be improved upon; I offer some suggestions here.

As other large AMR historical surveillance datasets become available, as endorsed in principle by major pharmaceutical companies<sup>26</sup>, correlations of quantitative MICs can also be investigated in these datasets. Furthermore, as sequencing becomes cheaper, it will become increasingly common to also perform whole genome sequencing of banks of isolates or do other molecular characterization. Linking MIC-based relationships to the genetic mechanisms of AMR would be an extension of this work.

During the course of preparing this article for publication, I have come across other authors noting the complementary information provided by 'semantic similarity' and 'chemical similarity'. Notably, one such article's illustrative example of two drugs that are chemically very different but semantically similar, is two beta-lactamase inhibitors<sup>27</sup>. It is also true that compounds with a high chemical similarity score do not necessarily have the same biological activity<sup>22</sup>. Combining the complementary information of chemical similarity scores with semantic 'efficacy' scores based on MICs could lead to a better representational hierarchy of existing antibiotics. Unfortunately, clinical efficacy scores by definition do not exist for novel compounds, so this would not help with drug development (a common use of chemical similarity scores).

At present, cross-resistance rates between major classes of antibiotics are reported in surveillance. It would in theory be possible to model antibiotic cross-correlation rather than simply cross-resistance. Comparing the cross-correlation distributions between different patient groups, geographical areas, or over time, would reveal additional insights. The problem I see here is defining what is 'of interest' as a substantially different correlation between two groups. Confidence intervals can be calculated on Spearman rank correlations (using the Fisher z-transformation), but because of the large numbers of isolates, even in subgroups, these would rarely overlap. Finding a

‘statistically significant’ difference between subgroups does not necessarily mean clinical or biological significance.

I have argued for the importance of a general overview of relationships in AMR, but the specifics are vital. There is still a great deal we still do not understand about the mechanisms of action and resistance for antibiotics. Novel cross-resistance phenotypes continue to challenge previously held ideas about the relationship of use of one antibiotic with resistance to another. For example, in a recent report of an unexpected mechanism conferring resistance to ceftazidime in *Stenotrophomonas maltophilia*, the authors conclude that using ceftazidime may already be ‘eroding the future potential’ of experimental siderophore-conjugated antibiotics against this species<sup>28</sup>. With mechanisms such as this in mind, one option would be to use MIC-based correlations to quantify the strength of possible interactions between antibiotics in terms of shared resistance mechanisms. These values could then be used in subsequent modelling attempts to link antibiotic consumption rates to resistance rates.

## Conclusion

AMR surveillance data is a valuable source of information – not only for historical resistance rates for single antibiotics, but also because it contains information about the relationships between antibiotics. Visualising similarities between antibiotics based on their efficacy against real clinical isolates is a useful way to conceptualise their relationships.

## Data availability

### Underlying data

Access to the data was through the Wellcome Data Re-use Prize – Surveillance<sup>9</sup>, Synapse ID [syn17009517](https://synapse.org/syn17009517).

Data underlying the results of this study can be obtained from the ATLAS Programme, which was run by Pfizer. Data are

available from <https://atlas-surveillance.com>, following free registration.

## Extended data

Figshare: Representing antibiotic relationships using measurements of efficacy against clinical isolates. <https://doi.org/10.6084/m9.figshare.8159525.v1><sup>12</sup>.

The following extended data are available:

- **Supplementary File 1.** The ATLAS dataset after data cleaning and ordering of MICs.
- **Supplementary File 2.** Metadata on antibiotics (classes, ATC code, SMILES).
- **Supplementary File 3.** R code necessary to reproduce [Table 1](#) and [Figure 2](#).

Extended data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

## Acknowledgements

I acknowledge Pfizer for releasing their historic AMR surveillance dataset (ATLAS) through the Wellcome Data Re-Use Prize and encourage other holders of AMR surveillance data to do similarly. I also acknowledge the Wellcome Trust for administering the Wellcome Data Re-Use Prize.

I thank Nicole Stoesser and Sarah Walker for their extremely helpful comments on a draft, and Phil Fowler for a useful discussion on computing chemical similarities of antibiotics.

This work was unfunded but resulted in a prize award of £5,000 from the Wellcome Trust, which has been donated to Medical Aid for Palestinians ([map.org.uk](http://map.org.uk)).

## References

1. O'Neill J: **Tackling drug-resistant infections globally: final report and recommendations.** The Review on Antimicrobial Resistance, London, 2016. [Reference Source](#)
2. Kariv G, Paul M, Shani V, et al.: **Benchmarking inappropriate empirical antibiotic treatment.** *Clin Microbiol Infect.* 2013; **19**(7): 629–633. [PubMed Abstract](#) | [Publisher Full Text](#)
3. Llewelyn MJ, Hand K, Hopkins S, et al.: **Antibiotic policies in acute English NHS trusts: implementation of ‘Start Smart-Then Focus’ and relationship with *Clostridium difficile* infection rates.** *J Antimicrob Chemother.* 2015; **70**(4): 1230–1235. [PubMed Abstract](#) | [Publisher Full Text](#)
4. Llewelyn MJ, Fitzpatrick JM, Darwin E, et al.: **The antibiotic course has had its day.** *BMJ.* 2017; **358**: j3418. [PubMed Abstract](#) | [Publisher Full Text](#)
5. Inácio J, Barnes LM, Jeffs S, et al.: **Master of Pharmacy students' knowledge and awareness of antibiotic use, resistance and stewardship.** *Curr Pharm Teach Learn.* 2017; **9**(4): 551–559. [PubMed Abstract](#) | [Publisher Full Text](#)
6. Nerlich B, James R: **“The post-antibiotic apocalypse” and the “war on superbugs”: catastrophe discourse in microbiology, its rhetorical form and political function.** *Public Underst Sci.* 2009; **18**(5): 574–590. [PubMed Abstract](#) | [Publisher Full Text](#)
7. Mendelson M, Balasegaram M, Jinks T, et al.: **Antibiotic resistance has a language problem.** *Nature.* 2017; **545**(7652): 23–25. [PubMed Abstract](#) | [Publisher Full Text](#)
8. Limmathurotsakul D, Sandoe JAT, Barrett DC, et al.: **‘Antibiotic footprint’ as a communication tool to aid reduction of antibiotic consumption.** *J Antimicrob Chemother.* 2019; **74**(8): 2122–2127. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Wellcome Trust: **Wellcome Data Re-use Prize – Surveillance.** 2019. [Reference Source](#)
10. Pfizer: **Antimicrobial testing leadership and surveillance (ATLAS).** 2019. [Reference Source](#)
11. Shaw L: **Wellcome AMR Data Reuse Prize Submission.** 2019. [Publisher Full Text](#)
12. Shaw L: **Representing antibiotic relationships using measurements of efficacy**

- against clinical isolates. 2019.  
<http://www.doi.org/10.6084/m9.figshare.8159525.v1>
13. Buckley J, James I: **Linear regression with censored data.** *Biometrika.* 1979; **66**(3): 429–436.  
[Publisher Full Text](#)
  14. van de Kasstele J, van Santen-Verheuevel MG, Koedijk FD, *et al.*: **New statistical technique for analyzing MIC-based susceptibility data.** *Antimicrob Agents Chemother.* 2012; **56**(3): 1557–1563.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  15. R Core Team: **R: A Language and Environment for Statistical Computing.** R Foundation for Statistical Computing, Vienna, Austria, 2018.  
[Reference Source](#)
  16. Ward JH Jr: **Hierarchical grouping to optimize an objective function.** *J Am Stat Assoc.* 1963; **58**(301): 236–244.  
[Publisher Full Text](#)
  17. de Vries A, Ripley BD: **ggdendro: Create Dendrograms and Tree Diagrams Using 'ggplot2'.** R package version 0.1-20. 2016.  
[Reference Source](#)
  18. Wickham H: **ggplot2: Elegant Graphics for Data Analysis.** Springer-Verlag New York, 2016.  
[Publisher Full Text](#)
  19. WHO Collaborating Centre for Drug Statistics Methodology: **ATC/DDD Index 2019.**  
[Reference Source](#)
  20. Guha R: **Chemical informatics functionality in r.** *J Stat Softw.* 2007; **18**(5).  
[Publisher Full Text](#)
  21. Guha R: **Using the CDK from R.** 2018.  
[Reference Source](#)
  22. Maggiora G, Vogt M, Stumpfe D, *et al.*: **Molecular similarity in medicinal chemistry.** *J Med Chem.* 2014; **57**(8): 3186–3204.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  23. Dray S, Dufour AB: **The ade4 package: Implementing the duality diagram for ecologists.** *J Stat Softw.* 2007; **22**(4): 1–20.  
[Publisher Full Text](#)
  24. Chukwudi CU: **rRNA Binding Sites and the Molecular Mechanism of Action of the Tetracyclines.** *Antimicrob Agents Chemother.* 2016; **60**(8): 4433–4441.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  25. Hindler JA, Humphries RM: **Colistin MIC variability by method for contemporary clinical isolates of multidrug-resistant Gram-negative bacilli.** *J Clin Microbiol.* 2013; **51**(6): 1678–1684.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  26. Wellcome Trust: **Antimicrobial Resistance Surveillance: Sharing industry data.** 2018.  
[Reference Source](#)
  27. Ferreira JD, Couto FM: **Semantic similarity for automatic classification of chemical compounds.** *PLoS Comput Biol.* 2010; **6**(9): pii: e1000937.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  28. Calvopiña K, Dulyayangkul P, Heesom KJ, *et al.*: **TonB dependent uptake of  $\beta$ -lactam antibiotics in the opportunistic human pathogen *Stenotrophomonas maltophilia*.** *bioRxiv.* 2019; 615369.  
[Publisher Full Text](#)

# Open Peer Review

Current Peer Review Status:  

---

## Version 2

Reviewer Report 27 August 2020

<https://doi.org/10.21956/wellcomeopenres.16965.r39923>

© 2020 Ter Kuile B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



### Benno H Ter Kuile

Laboratory for Molecular Biology and Microbial Food Safety, University of Amsterdam, Amsterdam, The Netherlands

This manuscript presents an innovative approach to the analysis of cross resistances for 5 bacterial species and 22 antibiotics. As such it makes a valuable contribution. I have, however, some reservations of a general nature that I hope can be addressed.

The author oversells the insights considerably. Statements such as: "Current antimicrobial nomenclature makes it difficult to understand how different antibiotics are related to one another, and their empirical uses and efficacy", suggest more than is delivered. The results are five relationship trees comparable to evolutionary trees that show the probability that resistance for one antibiotic in one of the five species correlates to resistance for another. Again, the approach is valuable, but it does not solve the problems the author mentions, such as misconceptions of third and fourth years pharmacy students. If the clinical relevance is unclear (see page 7), then one should avoid sweeping statements.

The interpretation of Table 1 does not square very well with Figure 2. While it is claimed that chemical similarity and resistance correlate, this is not obvious from Figure 2. For the analysis of Table 1, ten antibiotics were selected. Was this selection random, or were these the ten that gave the anticipated outcome?

All in all, I am not convinced that the last sentence of the conclusion is justified by the preceding text. I suggest the author takes a close look at it before taking responsibility for it.

#### Is the work clearly and accurately presented and does it cite the current literature?

Yes

#### Is the study design appropriate and is the work technically sound?

Partly

#### Are sufficient details of methods and analysis provided to allow replication by others?

Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**

Not applicable

**Are all the source data underlying the results available to ensure full reproducibility?**

Yes

**Are the conclusions drawn adequately supported by the results?**

Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Antimicrobial resistance, specifically mechanism of de novo development.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 09 Sep 2020

**Liam Shaw**, University of Oxford, Oxford, UK

*This manuscript presents an innovative approach to the analysis of cross resistances for 5 bacterial species and 22 antibiotics. As such it makes a valuable contribution. I have, however, some reservations of a general nature that I hope can be addressed.*

Thanks for the fair and critical assessment.

*The author oversells the insights considerably. Statements such as: "Current antimicrobial nomenclature makes it difficult to understand how different antibiotics are related to one another, and their empirical uses and efficacy", suggest more than is delivered. The results are five relationship trees comparable to evolutionary trees that show the probability that resistance for one antibiotic in one of the five species correlates to resistance for another. Again, the approach is valuable, but it does not solve the problems the author mentions, such as misconceptions of third and fourth years pharmacy students. If the clinical relevance is unclear (see page 7), then one should avoid sweeping statements.*

I have made a few alterations to try to avoid "overselling". For example, in the discussion I have made clear that the possible future uses of similar diagrams I suggest are very much speculation, and that I don't think the figures in this paper are currently fit for these purposes. I hope these changes are satisfactory.

*The interpretation of Table 1 does not square very well with Figure 2. While it is claimed that chemical similarity and resistance correlate, this is not obvious from Figure 2. For the analysis of Table 1, ten antibiotics were selected. Was this selection random, or were these the ten that gave the anticipated outcome?*

Apologies if this was unclear. These were the ten single-molecule antibiotics with the most data: they all had >5,000 measurements for each of the top five bacterial species. This information was in the text but was missing from the caption for Table 1. It has now been added.

*All in all, I am not convinced that the last sentence of the conclusion is justified by the preceding text. I suggest the author takes a close look at it before taking responsibility for it.*

I acknowledge this criticism. I have deleted the last sentence of the conclusion in the abstract, and altered the last sentence of the conclusion: "Visualising similarities between antibiotics based on their efficacy against real clinical isolates is one way to use this information. "

**Competing Interests:** No competing interests were disclosed.

Reviewer Report 18 October 2019

<https://doi.org/10.21956/wellcomeopenres.16965.r36666>

© 2019 Wilcox M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Mark H. Wilcox**

University of Leeds, Leeds, UK

I am happy with the revisions.

**Is the work clearly and accurately presented and does it cite the current literature?**

Partly

**Is the study design appropriate and is the work technically sound?**

Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**

Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**

Partly

**Are all the source data underlying the results available to ensure full reproducibility?**

Partly

**Are the conclusions drawn adequately supported by the results?**

Partly

**Competing Interests:** No competing interests were disclosed.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

## Version 1

Reviewer Report 23 September 2019

<https://doi.org/10.21956/wellcomeopenres.16704.r36348>

© 2019 Wilcox M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Mark H. Wilcox**

University of Leeds, Leeds, UK

This is an interesting manuscript. The premise (that 'antibiotics with similar activities have similar mechanisms of action') is only partly correct. Actually, two antibiotics with identical MICs for a given strain could act by completely different mechanisms.

That the analysis shows 'The resulting visualizations broadly recapitulate antibiotic classes' is not surprising. The next claim that 'They also clearly show the dramatic effects of combination therapies ...' is an important potential leap of faith. The author describes a theoretical effect of combining antibiotics that may not actually be true in reality, as the effects of two antibiotics may be additive, synergistic or antagonistic. The only way to determine which of these is true for any given combination of antibiotics and for a specific strain (or possibly species) is to carry out phenotypic testing (typically using a checkerboard techniques). The combinations of a  $\beta$ -lactam and a  $\beta$ -lactamase inhibitor (e.g. piperacillin-tazobactam) is a very specific scenario, which technically is not usually referred to a combining antibiotics; rather, it is using an antibiotic (e.g. piperacillin) plus an enzyme inhibitor (e.g. tazobactam). The enzyme inhibitor usually has no (or very little) antibiotic activity itself.

As such, the manuscript/abstract is currently potentially misleading and requires modification.

**Is the work clearly and accurately presented and does it cite the current literature?**

Partly

**Is the study design appropriate and is the work technically sound?**

Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**

Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**

Partly

**Are all the source data underlying the results available to ensure full reproducibility?**

No source data required

**Are the conclusions drawn adequately supported by the results?**

No

**Competing Interests:** No competing interests were disclosed.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 02 Oct 2019

**Liam Shaw**, University of Oxford, Oxford, UK

*This is an interesting manuscript. The premise (that 'antibiotics with similar activities have similar mechanisms of action') is only partly correct. Actually, two antibiotics with identical MICs for a given strain could act by completely different mechanisms.*

I agree with the last statement. The fuzzy premise of 'similar activities' against a given strain implying 'similar mechanisms of action' is indeed not necessarily true, and in fact is not necessary for the correlation approach – the idea is to look at the patterns across isolates, where covariation would be suggestive of similar (or linked) mechanisms of resistance. I have rewritten this statement in the abstract, which I hope better summarises the central idea:

"The basic principle is that antibiotics with similar mechanisms of action should give rise to similar (or linked) mechanisms of resistance, so should have correlated MICs across large numbers of isolates."

*That the analysis shows 'The resulting visualizations broadly recapitulate antibiotic classes' is not surprising. The next claim that 'They also clearly show the dramatic effects of combination therapies ...' is an important potential leap of faith. The author describes a theoretical effect of combining antibiotics that may not actually be true in reality, as the effects of two antibiotics may be additive, synergistic or antagonistic. The only way to determine which of these is true for any given combination of antibiotics and for a specific strain (or possibly species) is to carry out phenotypic testing (typically using a checkerboard techniques).*

I agree the 'leap of faith' interpretation would be stretching it! I wasn't intending to imply that the correlation dendrograms could be used to derive effects of combining antibiotics – which of course should be tested phenotypically. It arose from a simple misunderstanding

on my part about what the term 'combination therapy' includes (see next comment). I have also added a paragraph to the discussion cautioning against any inference of potential effects of combining antibiotics from these dendrograms:

"I caution against inferring potential treatment interactions between antibiotics based on these dendrograms. In general, two antibiotics used in combination can be additive, synergistic (i.e. greater than the sum of their individual effects), or antagonistic (i.e. less than the sum of their individual effects). Phenotypic testing of combinations is necessary to understand these interactions."

*The combinations of a  $\beta$ -lactam and a  $\beta$ -lactamase inhibitor (e.g. piperacillin-tazobactam) is a very specific scenario, which technically is not usually referred to as combining antibiotics; rather, it is using an antibiotic (e.g. piperacillin) plus an enzyme inhibitor (e.g. tazobactam). The enzyme inhibitor usually has no (or very little) antibiotic activity itself.*

I was aware that in combinations such as piperacillin-tazobactam the enzyme inhibitor (tazobactam) has very little to no antibiotic activity itself. However, I wasn't aware that referring to this as a 'combination therapy' would be misleading and invite extension of the statement to dual-antibiotic combinations. I have rewritten the abstract accordingly:

"They also clearly show the dramatic effects of combining beta-lactam antibiotics with beta-lactamase inhibitors, as well as highlighting antibiotics which have unexpected correlations in MICs that are not predicted from their chemical similarities alone."

I have also rewritten the appropriate section of the results to make clear that the combination beta-lactam and beta-lactamase inhibitor is a very specific scenario:

"The beta-lactam antibiotic ceftazidime has limited activity against *S. aureus*. Therefore, combining it with avibactam (a non-beta-lactam beta-lactamase inhibitor with little to no antibiotic activity of its own) has little additional effect; the two treatments cluster close together (Figure 2a). However, the dramatic effect of using avibactam in combination with ceftazidime in Gram-negative species such as *E. coli* and *K. pneumoniae* is readily apparent; ceftazidime/avibactam is shown far apart from ceftazidime in these species (Figure 2b-2e)."

*As such, the manuscript/abstract is currently potentially misleading and requires modification*

Thank you for highlighting these points. I hope the modifications are satisfactory and clear up any potential confusion or misunderstanding.

**Competing Interests:** No competing interests were disclosed.