Reconciling the potentially irreconcilable? Genotypic and phenotypic amoxicillin-clavulanate resistance in *Escherichia coli*

Authors: Timothy J. Davies[a,b](#), Nicole Stoesser[a,b,c], Anna E Sheppard[a,b], Manal Abuoun[d], Philip Fowler[a], Jeremy Swann[a,b], T. Phuong Quan[a,b,e], David Griffiths[a,e], Alison Vaughan[a,e], Marcus Morgan[c], Hang TT Phan[a]*, Katie J Jeffery[c], Monique Andersson[c], Matt J Ellington[f], Oskar Ekelund[g], Neil Woodford[b,f], Amy J. Mathers[h], Robert A. Bonomo[i], Derrick W. Crook[a,b,c,e], Tim E.A. Peto[a,b,c,e], Muna F Anjum[d], A. Sarah Walker[a,b,e]

Affiliations:

a) Nuffield Department of Medicine, Oxford University, Oxford, United Kingdom

b) National Institute for Health Research (NIHR) Health Protection Research Unit on Healthcare Associated Infections and Antimicrobial Resistance at University of Oxford, UK

c) Oxford University Hospitals NHS Foundation Trust, Oxford, UK

d) Bacteriology, Animal and Plant Health Agency, Surrey UK

e) NIHR Biomedical Research Centre, Oxford, UK

f) Antimicrobial Resistance and Healthcare Associated Infections (AMRHAI) Reference Unit, National Infection Service, Public Health England, London UK

g) Department of Clinical Microbiology and the EUCAST Development Laboratory, Kronoberg Region, Central Hospital, Växjö, Sweden

h) Division of Infectious Diseases and International Health, Department of Medicine, University of Virginia Health System, Charlottesville, Virginia, USA
Clinical Microbiology, Department of Pathology, University of Virginia Health System, Charlottesville, Virginia, USA

i) Louis Stokes Cleveland Veterans Affairs Medical Centre, Research Service, Cleveland, OH; Case Western Reserve University, Departments of Medicine, Biochemistry, Molecular Biology and Microbiology, Pharmacology, and Proteomics and Bioinformatics; CWRU-Cleveland VAMC Centre for Antimicrobial Resistance and Epidemiology (Case VA CARES); and Geriatric Research Education and Clinical Centers (GRECC), Louis Stokes Cleveland Department of Veterans Affairs, Cleveland, OH

Running Title: Amoxicillin-clavulanate Resistance Genetics in *E. coli*

#Address correspondence to Timothy J Davies, timothy.davies@ndm.ox.ac.uk

Abstract:

Resistance to amoxicillin-clavulanate, a widely used beta-lactam/beta-lactamase inhibitor combination antibiotic, is rising globally, yet susceptibility testing remains challenging. To test whether whole-genome sequencing (WGS) could provide a more reliable assessment of susceptibility than traditional methods, we predicted resistance from WGS for 976 *E. coli* bloodstream infection isolates from Oxfordshire, UK, comparing against phenotypes from the BD Phoenix (calibrated against EUCAST guidelines). 339/976 (35%) isolates were amoxicillin-clavulanate resistant. Predictions based solely on beta-lactamase presence/absence performed poorly (sensitivity 23% (78/339)) but improved when genetic features associated with penicillinase hyper-production (e.g. promoter mutations, copy number estimates) were considered (sensitivity 82% (277/339); $p<0.0001$). Most discrepancies occurred in isolates with peri-breakpoint MICs. We investigated two potential causes; the phenotypic reference and the binary resistant/susceptible classification. We performed reference standard, replicated phenotyping in a random stratified subsample of 261/976 (27%) isolates using agar dilution, following both EUCAST and CLSI guidelines, which use different clavulanate concentrations. As well as disagreeing with each other, neither agar dilution phenotype aligned perfectly with genetic features. A random-effects model investigating associations between genetic features and MICs showed that some genetic features had small, variable and additive effects, resulting in variable resistance classification. Using model fixed-effects to predict MICs for the non-agar dilution isolates, predicted MICs were in essential agreement (±1 doubling dilution) with observed (BD Phoenix) MICs for 691/715 (97%) isolates. This suggests amoxicillin-clavulanate resistance in *E. coli* is quantitative, rather than qualitative, explaining the poorly reproducible

58 binary (resistant/susceptible) phenotypes and suboptimal concordance between different

59 phenotypic methods and with WGS-based predictions.

60

Introduction

Rising amoxicillin-clavulanate resistance in *E. coli* is a major healthcare challenge, with increasing incidence of resistant bloodstream infections (BSI)(1) threatening its utility as the most commonly used antibiotic in Europe.(2) Consequently, many hospitals are considering broadening their first-line empiric antibiotics for common infections. However, significant uncertainty is created by observed differences between the two main assays for amoxicillin-clavulanate susceptibility in the classification of clinical samples.(3) These differences are so large that increasing amoxicillin-clavulanate resistance was suggested to be primarily due to laboratories switching from US Clinical Laboratory Standards Institute (CLSI) to European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines.(4) Recent work,(5) however, suggests that changes in laboratory protocols are unlikely to account for the majority of the increase in resistance. Only one study has investigated whether there are underlying genetic causes for the ongoing rise in amoxicillin-clavulanate resistance,(6) but found no evidence of clonal expansion of any specific amoxicillin-clavulanate-resistant strains. However, the genetic epidemiology of amoxicillin-clavulanate resistance mechanisms was not investigated.

In addition to its widespread clinical use, amoxicillin-clavulanate is a model for beta-lactam/beta-lactamase inhibitor (BL/BLI) combinations, which are the focus of renewed attention(7) due to the development of novel BL/BLIs with activity against highly drug-resistant organisms.(8) EUCAST has recently published guidelines on setting breakpoints for BL/BLIs,(9) but the inconsistencies seen in testing and clinically interpreting amoxicillin-clavulanate resistance likely extend to novel BL/BLIs.(10)

One solution is to instead identify the genetic determinants characterizing resistance (resistance genotype) using whole-genome sequencing (WGS).(11) This approach may be particularly helpful for BL/BLI, as recent studies have suggested that traditional phenotyping is less accurate in isolates producing extended spectrum beta-lactamases.(12) Rather than resistance being associated with the simple presence/absence of specific genes, previous studies have found much amoxicillin-clavulanate resistance is likely attributable to mechanisms which increase the effective concentration of beta-lactamases (e.g. additive effects of multiple beta-lactamases,(13) increasing gene expression(14) or modifying cell permeability(15)). Given the added complexity of both phenotype and genotype, studies using WGS to predict phenotypic resistance have either not included amoxicillin-clavulanate,(16, 17) compared against only one set of breakpoints,(18) or only tested small sets of pre-selected samples.(19) Similar studies investigating other BL/BLIs, such as piperacillin-tazobactam, reported poor accuracy when predicting resistance from genotype.(20)

We therefore investigated concordance between WGS-derived genotypes and amoxicillin-clavulanate susceptibility phenotypes in a large, unselected set of Oxfordshire *E. coli* BSI isolates from 2013-2015. We assessed whether extending the usual presence/absence genetic approach to include features that might increase beta-lactamase expression (copy number, promoter type) would improve concordance, and quantified the impact of particular genetic variants and testing guidelines (EUCAST, CLSI) on minimum inhibitory concentrations (MICs).

Results

Routine laboratory phenotypes and amoxicillin-clavulanate resistance genotypes

Of the 1039 *E. coli* BSI occurring between January 2013-August 2015 in Oxfordshire, UK, 1000 had at least one isolate stored by Oxford University Hospitals (OUH) NHS Foundation Trust microbiology laboratory. In most (992/1000 (99%)) infections, only a single *E. coli* was isolated; however, two different *E. coli* were grown from culture in 8 cases, giving a total of 1008 distinct *E. coli* isolates. Each of these isolates had linked antimicrobial susceptibility test (AST) data from the Oxford University Hospitals (OUH) NHS Foundation Trust microbiology laboratory using the BD Phoenix (Beckton, Dickinson and Company). All obtained isolates were sequenced, with 976/1008 (97%) having WGS data meeting pre-determined quality controls designed to identify mixtures and poor-quality sequences. Overall, these 976 isolates represented 968/1039 [93%] *E. coli* BSI (Supplementary Figure S1). 339/976 (36%) had amoxicillin-clavulanate MIC > 8/2 mg/L by EUCAST breakpoints (Supplementary Table S1).

The collection was highly diverse, representing 152 different sequence types (STs). The most common was ST73 (161,17%) (Supplementary Figure S2), followed by ST131 (124,13%), which had the highest percentage of phenotypically-resistant isolates (N=74,60%) and was the only ST associated with amoxicillin-clavulanate resistance (chi-squared $p<0.0001$ compared with $p>0.16$ for all other STs).

The most common beta-lactam resistance mechanisms identified (using ARIBA(21) (default parameters) and tBLASTn/BLASTn (see Methods)) were acquired beta-lactamase genes, which were identified in 515/976 (53%) isolates. Most of these (448/515 (87%)) harbored only a single

transmissible beta-lactamase gene. Among the 67 isolates with more than one beta-lactamase gene, the most common combination was *bla*$_{\text{CTX-M-15}}$ and *bla*$_{\text{OXA-1}}$ (N=27, Supplementary Table S2B). Overall *bla*$_{\text{TEM}}$ was by far the most common mechanism identified (occurring in 427/976 (44%) isolates), followed by *bla*$_{\text{CTX-M}}$ (N=73 (7%)), *bla*$_{\text{OXA}}$ (N=62 (6%)) and *bla*$_{\text{SHV}}$ (N=23 (2%)) (Supplementary Figure S2, Supplementary Table S2). For the 594 transmissible beta-lactamases identified, median DNA copy number from mapping coverage was 2.23 (IQR 1.73,3.31); 227 (38%) had >2.5-fold coverage (the threshold to predict resistance derived from receiver operating characteristic (ROC) analysis of isolates with only one beta-lactamase identified, Supplementary Methods, Supplementary Figure S3). Variant *bla*$_{\text{TEM}}$ and *ampC* promoters considered to be associated with increased expression were identified in 49 (5%) and 20 (2%) isolates respectively (Supplementary Table 3A, 3C). 31 (3%) isolates potentially had one non-functional porin, of which 22 also contained a beta-lactamase gene; however, no isolate had "functionally lost" both *ompC* and *ompF* (Supplementary Table S4).

WGS-derived resistance prediction compared with routine phenotyping

We compared two genetic resistance prediction algorithms for amoxicillin and amoxicillin-clavulanate (see Methods). The first, denoted the "basic" prediction algorithm, was analogous to common WGS-based resistance methods and only predicting resistance for isolates containing inhibitor resistant beta-lactamase genes (e.g. *bla*$_{\text{OXA-1}}$, *bla*$_{\text{TEM-30}}$). The second denoted the "extended" prediction algorithm, additionally evaluated *bla*$_{\text{TEM}}$ and *ampC* promoter mutations, estimates of beta-lactamase gene DNA copy number and porin loss-of-function mutations. Including the additional features (i.e. the "extended" approach) had little impact on our ability to identify ampicillin resistance but significantly improved amoxicillin-clavulanate resistance

prediction (Table 1~~Table 1~~, amoxicillin sensitivity 98% basic vs 96% extended; amoxicillin-clavulanate sensitivity 23% basic vs 82% extended, McNemar's p<0.0001). However, the increased sensitivity also came at the cost of modestly reduced specificity (Table 1). Overall categorical agreement of WGS-derived with observed phenotype increased from 712 (73%) to 868 (89%) when these extended genetic features were included.

Investigating the cause of lower than optimal agreement, even using the extended algorithm, showed that most false positive predictions were made on the basis of increased beta-lactamase gene DNA copy-number (**Error! Reference source not found.**~~Table 2~~). Although there was a clear association between increasing copy-number and MIC (p<0.0001), resistance prediction based on increased DNA copy-number (>2.5) was less accurate than other extended algorithm components (positive predictive value (PPV) = 0.77 compared to >0.97 for all other algorithm components), with both resistant isolates with lower copy-number beta-lactamases and susceptible isolates with higher copy beta-lactamases (Figure 1).

The distribution of MICs in isolates with concordant vs discordant predictions suggested an alternative explanation (Figure 2), with the extended algorithm performing better at predicting susceptibility/resistance in non-peri-breakpoint isolates. Overall the algorithm correctly classified 463/469 (99%) isolates with MIC <=4/2 mg/L as susceptible and 230/250 (92%) isolates with MIC >=32/2 mg/L as resistant. Notably, of 79 discordant isolates containing only non-inhibitor-resistant beta-lactamases, 64 (81%) had peri-breakpoint (8/2-16/2mg/L) MICs.

Given these findings, we therefore investigated two other hypotheses that could explain the low agreement in peri-breakpoint isolates: (*i*) variable accuracy of the different phenotypic methods, and (*ii*) the binary resistant/susceptible classification being too simplistic.

Variability in reference standard agar dilution phenotypes (EUCAST and CLSI based)

291/976 (30%) isolates were selected for repeated agar dilution phenotyping using stratified random subsampling to enrich for resistant isolates both with and without beta-lactamase genes (see Methods, Supplementary Methods, Supplementary Figure S1). Of these 291 isolates, 261 (90%) passed the additional quality control steps designed to remove potential undetected mixtures, and were included in the agar-dilution subsample (details in Supplementary Methods; in brief, all colonies had to be of one morphology on blood-agar purity plates and MICs for each of amoxicillin and both amoxicillin-clavulanate fixed (EUCAST)/ratio (CLSI) tests had to be in essential agreement on 2 or more repeat tests). The stratified random sampling enriching for resistant phenotypes meant that 160/261 (61%) subsample isolates were amoxicillin-clavulanate-resistant by routine AST (Supplementary Table S1). All STs with >10 isolates in the main sample were represented, with 52 (20%), 43 (16%) and 29 (11%) isolates being ST131, ST73 and ST69, respectively, as were all resistance gene families in the main sample (Supplementary Figure S2).

As expected, phenotypes from different reference-standard AST methods were often discordant (Figure 3). EUCAST-based agar dilution (using the fixed 2mg/L clavulanate concentration) only

agreed with CLSI-based agar dilution (using the 2:1 ratio of amoxicillin:clavulanate) for 143/261 (55%) isolates (27 agreed resistant, 116 agreed susceptible). For the remaining 118 isolates, EUCAST-based agar dilution results were more conservative than CLSI-based agar dilution. Major discrepancies occurred for 39 isolates, being classed resistant by EUCAST-based agar dilution and susceptible by CLSI-based agar dilution. The remaining 79 isolates were EUCAST-resistant CLSI-intermediate. Excluding isolates classified as intermediate by CLSI, categorical agreement between the two reference-standard methods was 79%. Considering CLSI-intermediate as resistant had little impact on the overall categorical agreement (85%). Each of these test methods also often classified isolates differently to the BD Phoenix (Figure 3), but, as expected, given the BD Phoenix used in the OUH routine laboratory is calibrated against EUCAST guidelines, EUCAST-based agar dilution was in agreement more often. Of note, one isolate which only contained a partial bla_{TEM} gene was repeatedly resistant on both EUCAST and CLSI-based agar dilution testing but was identified susceptible by the BD Phoenix.

MIC results from both methods were variable on retesting (as part of triplicate repeats): more so for EUCAST-based agar dilution MICs (Supplementary Figure S4*),* which were not constant across repeats for 158/261 (61%) isolates versus only 73 (28%) for CLSI-based agar dilution MICs. While differences across repeats were in essential agreement with one another (i.e. less than (±1 doubling dilution) for all but 12 isolates for EUCAST-based agar dilution and all but 1 isolate for CLSI-based agar dilution, they did cause changes in resistance classification. For EUCAST-based agar dilution, 40/261 (15%) isolates were identified as both resistant and susceptible across repeats, suggesting that even within-method categorical agreement is far poorer than the standards required for regulatory approval. Likewise for CLSI-based agar

dilution, MIC differences across repeats resulted in variation in resistance classification for 31 (12%) isolates; however because of the CLSI-intermediate category, 28/31 (90%) of these would be classed as minor discrepancies.

WGS-derived resistance prediction compared with reference-standard agar dilution phenotypes

Overall, using the extended algorithm above, WGS classified as resistant 23/27 (85%) isolates agreed resistant by EUCAST and CLSI-based tests, 107/118 (91%) indeterminate isolates (76/79 EUCAST-resistant/CLSI-intermediate, 31/39 EUCAST-resistant/CLSI-susceptible) and 17/116 (15%) agreed susceptible isolates (Figure 4). Again predictions based on the presence of high copy number (>2.5x) non-inhibitor beta-lactamases alone were the least congruent with the reference-standard phenotypes. Specifically, 16/62 (26%) isolates with increased copy number beta-lactamase genes were agreed susceptible (accounting for 16 of the 17 resistance predictions in agreed susceptible isolates). Further, whilst 46/62 (74%) isolates with this mechanism were resistant on EUCAST-based agar dilution, similar to the PPV with BD Phoenix on the whole dataset, for CLSI-based agar dilution only 2/62 (3%) were CLSI resistant, 24/62 (39%) were CLSI-intermediate and 36/62 (58%) were CLSI susceptible, suggesting this threshold performs more poorly for predicting CLSI-based agar dilution phenotypes. However, as when selecting the initial threshold against BD Phoenix results, there was no threshold which perfectly predicted the CLSI-based phenotype (Supplementary Figure S2B). Only 8 (30%) of the 27 agreed resistant by EUCAST and CLSI-based tests contained inhibitor-resistant beta-lactamases. Conversely, 24/79 (30%) CLSI-intermediate and 10/39 (26%) CLSI-susceptible isolates contained *bla*$_{\text{OXA-1}}$, showing that identification of inhibitor-resistant beta-lactamases was neither necessary nor sufficient to predict resistance for the CLSI-based tests.

241 Similarly, assessment of the individual contribution of other genetic features to the phenotype
242 was challenging due to co-occurrence of features in the same isolate and the impact of some
243 features on susceptibility varying both between isolates and within isolate repeats
244 (Supplementary Figures S3, S4). For example, 4/9 isolates with ampC promoter mutations in the
245 agar dilution subsample were both resistant and intermediate on repeat testing using CLSI-based
246 agar dilution.

247

248 *WGS-derived resistance prediction in peri-breakpoint and non-peri-breakpoint isolates*

249 As with routine AST, WGS predictions of reference-standard phenotypes were more accurate for
250 non-peri-breakpoint MICs (EUCAST-based agar dilution: (≤4/2 mg/ml, ≥32/2 mg/ml), CLSI-
251 based agar dilution: (≤4/2 mg/ml, ≥32/16 mg/ml)). For EUCAST-based agar dilution, WGS
252 correctly identified resistance/susceptibility in 169/177 (95%) isolates with non-peri-breakpoint
253 MICs, versus only 60/84 (71%) with peri-breakpoint MICs. Similarly, for CLSI-based agar
254 dilution, excluding 79 intermediate isolates (16/8 mg/L), WGS correctly predicted 97/106 (92%)
255 non-peri-breakpoint isolates, but predicted 43/76 (57%) isolates with MIC 8/4 mg/L as resistant.

256 Interestingly, however, there were three consistently resistant (EUCAST-based MIC ≥
257 32/2mg/L, CLSI-based MIC ≥32/16 mg/L) and three consistently susceptible (EUCAST-based
258 MIC ≤4/2mg/L, CLSI-based MIC ≤4/2 mg/L) discrepants. All three resistant discrepants were
259 explained by complexities inferring phenotype from WGS. One had a novel *bla*$_{\text{CTX-M}}$ variant
260 (CTX-M-15-like, Ser130Gly mutation). Previous work on mechanisms of beta-lactamase
261 inhibition suggests mutations at Ambler position(22) 130 likely lead to inhibitor resistance(7),
262 and a similar mutation (Ser130Thr CTX-M-190) resulted in sulbactam and tazobactam
263 resistance.(23) The other two isolates had antibiograms consistent with *ampC* hyper-production

(cefoxitin resistant, ceftazidime resistant, cefepime susceptible), but we were unable to identify complete promoter sequences matching our reference (CP009072.1) in the region upstream of *ampC.* This may suggest insertion of alternative elements upstream of *ampC* could have led to both fragmented assemblies and have driven increased expression; however from WGS data alone it is difficult to distinguish if this has truly occurred or instead may be due to other undetected beta-lactamase resistance mechanisms. All three susceptible discrepants had beta-lactamases present at mildly elevated copy numbers (2.5-3.5x relative DNA coverage) leading to WGS prediction of resistance, which may be due to inherent unavoidable difficulty selecting cutoffs for predicting phenotype (Supplementary Figure S3B).

Impact of individual resistance features on a continuous measure of susceptibility

Random-effects models were used to investigate the impact of test method and WGS-identified genetic elements on agar dilution $\log_2$ MICs simultaneously, and to create a WGS-based resistance prediction for comparison with phenotype (Supplementary Methods). Elements were categorised depending on frequency (Supplementary Table S5). The most predictive aspect of each element (including presence/absence of genes and/or promoter mutations and/or gene dosage) was selected using the Akaike Information Criterion (AIC) (Supplementary Methods). Interaction terms between genetic elements (reflecting saturation effects) and with test methodology (reflecting differential impact of the same genetic mechanism depending on the amoxicillin:clavulanate ratio) were included where $p<0.05$.

All beta-lactamases were associated with increased MICs in univariable models (Supplementary Table S6) and these associations generally persisted in multivariable models, although their magnitude decreased markedly (Figure 5, Supplementary Table S7). The largest effects of beta-lactamase presence/absence were for *bla*$_{\text{OXA-1}}$ (a class 2d beta-lactamase, denoted blaOXA:2d) and members of the 'other' group of beta-lactamases, comprising either inhibitor-resistant beta-lactamases (N=10), or those with unknown impact on beta-lactam susceptibility (N=4) (Supplementary Table S5). These caused two-three fold and four-fold doubling dilution increases in EUCAST-based MIC respectively. The effects of non-inhibitor resistant *bla*$_{\text{TEM}}$ (denoted *bla*TEM:2b) and *bla*$_{\text{SHV}}$ (denoted blaSHV:2b) genes were more complex. For each, presence alone in an isolate was only associated with a small, often non-significant increase in MIC by either method (*bla*$_{\text{TEM}}$: impact on change in $\log_2$(MIC) CLSI-based=+0.36 [p=0.01], EUCAST-based=+0.14 [p=0.51]; *bla*$_{\text{SHV}}$: change in $\log_2$(MIC) CLSI-based=+0.03 [p=0.93], EUCAST-based=-0.61 [p=0.27]). However, for both, higher copy number (i.e. gene dosage) was associated with higher MIC. These effects were small but additive (e.g. EUCAST-based MIC change in $\log_2$(MIC) per doubling of copy number *bla*$_{\text{TEM}}$=+0.79 [p<0.0001], *bla*$_{\text{SHV}}$=+0.71 [p=0.004], detail in Supplementary Table S7A). Like beta-lactamases, all "significant" promoter mutations were associated with increased MICs (p<0.0001). In particular, "significant" *ampC* promoter mutations were independently associated with large increases in MIC (impact on change in $\log_2$(MIC) CLSI-based=+2.60 [p<0.0001], EUCAST-based=+4.25 [p<0.0001]). Interestingly, there was no clear change in MIC independently associated with suspected porin loss in our data (p≥0.06), despite porin loss being associated with a large effect in unadjusted analysis (Supplementary Table S6, change in $\log_2$(MIC) +2.28 (CLSI) and +4.17 (EUCAST))..

Of note, when increased copy number effects were included, EUCAST-based testing methodology accentuated increases in MIC caused by genetic resistance features other than for suspected porin loss and presence of *bla*$_{\text{CTX-M}}$ genes (the blaCTX-M:2be group)($p_{heterogeneity} \leq 0.05$). EUCAST-based methodology however was also associated with increased between and within sample standard deviation (Supplementary Table S7B).

Predictions of MIC in an independent validation set

Final EUCAST-based agar-dilution model estimates were then used to predict MICs for the 715/976 non-subsample isolates, which were then compared with BD Phoenix MICs. MIC predictions were in agreement for 557/715 (78%) isolates and in essential agreement (within ±1 doubling dilution) for 691/715 (97%). However, these 715 non-subsample isolates included 11isolates which contained resistance mechanisms not present among the agar-dilution subsample isolates from which the model was derived (e.g. different beta-lactamase variants). Excluding these, prediction performance was similar, with agreement for 554/704 (79%) isolates (Figure 6) and essential agreement for 683/704 (97%) isolates. Similarly to comparisons between the different antimicrobial susceptibility testing methods (Figure 3), agreement between predicted and observed resistant/susceptible classifications was lower (90%) despite having high essential agreement of MICs. While overall performance was good, three isolates had predicted MICs three doubling dilutions lower than observed. One had an unusual yet reproducible phenotype [ampicillin susceptible, amoxicillin-clavulanate resistant]. This rare phenotype has generally been found in non-*E. coli* Enterobacteriaceae, and is thought to be due to either mechanisms of *ampC* induction(24) or the differential activity of amoxicillin and ampicillin.(25) We were unable to identify a clear causative mechanism in this isolate; however, of note, it was

the only isolate to contain a -11 C>T *ampC* promoter mutation.. The other two both had observed MIC ≥32/2mg/L but only contained a low copy number $bla_{\text{TEM-1}}$ and had predicted MIC 8/2mg/L.

Discussion

Decisions about broadening recommended empiric antimicrobial regimens from amoxicillin-clavulanate are currently being made based on unclear AST data which appears poorly concordant with WGS-identified determinants of beta-lactamase resistance. Here, we have demonstrated that this lack of concordance is not due to unknown genetic features or inherent phenotyping problems as previously hypothesized.(26) Instead, it appears to arise from poor interpretation of how known genetic mechanisms of resistance impact phenotype. Contrary to the often assumed paradigm that beta-lactam resistance is generally due to the presence/absence of specific beta-lactamases alone, mechanisms of resistance to amoxicillin-clavulanate seen regularly in a large unselected clinical dataset were multifactorial, resulting from combinations of multi-copy beta-lactamase genes, mutations in resistance gene-associated promoters, and inhibitor resistance mechanisms. The individual effects of some of these features on MIC were small, variable and additive, resulting in only minor shifts around clinical breakpoints. This potentially explains inconsistencies on repeated phenotyping, and may be a consequence of the genetic basis of resistance rather than an inherent test weakness. A further corollary is that discrepancies between genotypic predictions and phenotype are inevitable when using susceptible/resistant binary classifications. Finally, the phenotypic testing methodology significantly affected the magnitude of the effect of these resistance features on the MIC. These issues, when combined, resulted in inconsistent binary phenotypes despite reliable MICs, and

354 consequently led to inevitable suboptimal concordance both between different phenotypic testing
355 methodologies and also with WGS-based susceptibility/resistance predictions An alternative
356 approach would be to use WGS to predict MICs directly. We demonstrated this was possible by
357 predicting the MIC to within one doubling dilution (essential agreement) of the observed MIC
358 for 97% of isolates from a population-representative set of *E. coli* BSI.

359

360 Our study highlights the importance of isolate sampling frame, phenotyping method and
361 breakpoint selection. A previous study of 76 *E. coli* isolated from cattle (19), which reported
362 high sensitivity and specificity of WGS to predict amoxicillin-clavulanate resistance, contained
363 highly-resistant isolates (30% containing bla_{CMY-2}), and only attempted to predict CLSI-defined
364 resistance (>32/16 mg/L). In contrast, in our study, similar to other population representative
365 studies of human isolates,(6, 14) only a small proportion of amoxicillin-clavulanate resistance
366 was due to inhibitor-resistant beta-lactamases, with most of the resistance being due to hyper-
367 production of beta-lactamases. Further, while EUCAST argue that pharmacodynamic data
368 support choice of breakpoint and clavulanate concentration (9, 27), there is no definitive
369 evidence as to which method has stronger associations with either clinical outcome or genotype.
370 We therefore assessed WGS against both commonly used methods (EUCAST, CLSI).

371

372 Compared with other studies of BL/BLIs and *E. coli* causing human infections, we found less
373 BL/BLI resistance was accounted for by inhibitor-resistant beta-lactamases.(20) To identify
374 resistance in our population-representative set of isolates, we found it critical to consider genetic
375 features that alter expression of beta-lactamases. Although the individual effects of some of these
376 features on MICs were small, they were important, because MICs for many isolates were close to

the breakpoint. Further, given the small size of these effects and effects of testing methodology, isolates could exhibit either susceptible or resistant phenotypes on repeat testing, supporting the concept of an "intermediate" phenotype, which is not accounted for in the EUCAST guidelines. The discrepancies between EUCAST and CLSI phenotypes we observed were similar to previous studies,(3) suggesting that phenotypic interpretation for one of our most commonly used clinical antibiotics remains open to question. The one mechanism for which we found no evidence of effect was porin loss: this may reflect difficulty detecting these effects from WGS alone or be a simple power issue given rarity of this mechanism in our population-based sample, since porin loss has been associated with raised MICs to other BL/BLI combination antibiotics in isolates containing *Klebsiella pneumoniae* carbapenemases.(28)

The main study strengths are the large, population-representative sampling frame; detailed, replicated, reference-grade phenotyping for a substantial subset of isolates on which prediction models were developed ("training set") and a large number of additional isolates with single phenotypes assessed by a commercial clinically accredited platform ("test set"); detailed and complete genotyping; and the statistical modelling. While model performance on other datasets with geographical differences in resistance mechanisms and prevalence is unknown, its good performance in unselected clinical isolates with resistance mechanisms commonly seen in practice (in contrast to many previous studies of WGS-based resistance prediction(29)) suggest it may be generalizable. Further, the consistency of individual findings with previous literature (including inconsistent phenotype(30), mechanisms of genetic resistance(31) and poor performance of beta-lactamase presence/absence-only prediction for BL/BLI resistance (20, 32)) provide confidence that the combined results may apply across different settings. A further study

strength was our use of more complete representations of the mechanisms of beta-lactam resistance. Compared to other studies of WGS-based resistance prediction which have either just used presence/absence of beta-lactamases or machine learning methods directly on sequence data, our models more easily align with traditional approaches of studying antimicrobial resistance and provide interpretable estimates of the direct effects of different mechanisms. A limitation is that we could only investigate proxies for some important genetic features, e.g. increased DNA copy number leading to increased expression. WGS is unable to directly quantify these effects, which thus require additional characterization by alternative methods, leading to concern that resistance prediction from WGS alone would be highly challenging. In practice however, the good predictive performance of our model using relatively simple proxies suggests many of these features can indeed be approximated from WGS data. Modelling associations between resistance features and MIC directly allowed us to avoid inferring the phenotype from the genotype using pre-specified rules and account for the effects of multiple features existing in individual isolates. The complexity of the underlying associations we discovered highlights the challenges facing standardized methods for predicting resistance across multiple drugs and species,(33) and the need for automated approaches based on machine learning to take into account proxies for increased expression.

The main limitations of this study relate to its size. While we determined repeat agar dilution phenotypes for a relatively large number of isolates (n=261) compared to other studies,(30) many resistance elements were still rare. This had three important consequences: some infrequent features had to be categorized together for modelling, interactions between all combinations features (e.g. combinations of beta-lactamases) could not be definitively assessed, and some

mechanisms present in our testing set were not present in the training set (e.g. some rare known beta-lactamase variants, many only in a single isolate, Supplementary Table S2) and so their effect could not be estimated. Model results however suggested these had limited consequences. Firstly, the features causing the greatest MIC increases were those traditionally associated with amoxicillin-clavulanate resistance,(7) their specific impact being modelled here for the first time. Secondly, only a small number of isolates had resistance mechanisms not seen in the training dataset (N=11), meaning impact on performance was minimal (essentially their effect was assumed to be 0). This issue is inevitable given the substantial diversity of *E. coli* and incomplete knowledge of resistance mechanisms, but the excellent performance in the remaining 704/715 isolates suggests the vast majority of clinical isolates could be amenable to WGS-based MIC prediction, leaving a much smaller, more tractable number of isolates needing additional phenotypic investigation. Another potential limitation was the use of agar dilution as our reference-standard phenotype, a method which, while previously endorsed by EUCAST,(34) is no longer recommended, with broth microdilution now recommended instead. By contrast CLSI still considers agar dilution as equivalent to broth microdilution.(35) Reassuringly, differences we found between the BD Phoenix and agar-dilution were similar to a previous study comparing BD Phoenix with reference-standard broth microdilution,(36) suggesting this would have relatively little impact on our overall results.

In summary, amoxicillin-clavulanate resistance in *E. coli* is quantitative, rather than qualitative; in reality, resistance is a continuum built up by many individual features inevitably resulting in poor reproducibility and suboptimal concordance with binary classifications. WGS can identify the causes of amoxicillin-clavulanate resistance in *E. coli* provided the approach is extended to

consider the complicated, polygenic, and expression-related nature of this resistance. This suggests a genetic approach could offer a less assay dependent way to assess amoxicillin-clavulanate resistance. With renewed interest in using BL/BLIs to treat highly drug-resistant infections, our study has implications for both clinical practice and research. Given susceptibility phenotypes are highly dependent on the phenotypic method used, they must be interpreted with caution. Genetic approaches have the potential to circumvent this issue. Importantly however, the assumption that BL/BLI resistance is binary (susceptible/resistant) may be unhelpful as the same underlying resistance feature can be associated with MICs just below or just above the breakpoint. Given the variability and complexity in both the underlying mechanisms and resulting phenotype, a more transparent approach considering background genetic features, expression levels of beta-lactamases, MIC values and clinical syndrome, is likely needed to guide management decisions.

Materials and Methods

Study population and routine microbiological processing

E. coli isolated from all monomicrobial or polymicrobial blood cultures at Oxford University Hospitals (OUH) NHS Foundation Trust between 01/Jan/2013-31/Aug/2015 were included, excluding repeat positive cultures within 90-days of an index positive. Automated AST was performed in the routine laboratory (BD Phoenix; Beckton, Dickinson and Company) and MICs interpreted using EUCAST breakpoints. Data were extracted from the Infectious Diseases in Oxfordshire Research Database (IORD)(37) which has Research Ethics Committee and Health Research Authority approvals (14/SC/1069, ECC5-017(A)/2009).

DNA extraction and sequencing

Isolates were re-cultured from frozen stocks stored in nutrient broth plus 10% glycerol at -80ºC. DNA was extracted using the QuickGene DNA Tissue Kit S (Kurabo Industries, Japan) as per manufacturer's instructions, with an additional mechanical lysis step (FastPrep, MP Biomedicals, USA) immediately following chemical lysis. A combination of standard Illumina and in-house protocols were used to produce multiplexed paired-end libraries which were sequenced on the Illumina HiSeq 2500, generating 151bp paired-end reads. High quality sequences (Supplementary Methods) were de-novo assembled using Velvet(38) as previously described.(39) *In silico* Achtman(40) multi-locus sequence types (MLST) types were defined using ARIBA.(21)

Evaluating the importance of genetic features that modify effective beta-lactamase concentration

We identified components of two genetic resistance prediction algorithms for amoxicillin and amoxicillin-clavulanate (Table 2, Supplementary Methods) using ARIBA(21) (default parameters) and tBLASTn/BLASTn.(41) The "basic" prediction used only presence/absence of relevant genes in the Resfinder(17) database, and the "extended" prediction additionally included bla_{TEM} and *ampC* promoter mutations, estimates of DNA copy number and predicted porin loss-of-function. High DNA copy number was used as an indicator of possible gene duplication or high plasmid copy number, which are both known to cause increased beta-lactam resistance.(42, 43) We made no attempt to distinguish between these two causes due to the limitations of short-read sequencing data. For bla_{TEM} and *ampC* promoters, sequences identified using

ARIBA/BLASTn were searched for variant sites and regions previously associated with significantly increased expression.(44–46) For transmissible resistance genes, we estimated DNA copy number by comparing mapping coverage with the mean coverage of MLST genes and defined a relative coverage of >2.5 as increased copy number (based on receiver-operator-curve (ROC) analysis, Supplementary Methods; Supplementary Figure S3A). Finally, sequences found by ARIBA using reference *ompC* and *ompF* sequences (RefSeq: NC_000913.3) were inspected for features such as indels and truncations suggesting functional porin loss.

Evaluating the impact of different phenotypic methods

A subset of 291 isolates were selected for replicate agar dilution phenotyping using random sampling within strata defined by phenotype-genotype combinations (see Supplementary Methods for full sub-sampling procedure, Supplementary Figure S1A,B). Replicate agar dilution phenotyping used clavulanate concentration and MIC interpretation according to both EUCAST and CLSI guidelines. The aim was to explore reasons for discordance between observed phenotype and predictions made on the basis of beta-lactamase gene presence/absence alone. The subsampling therefore aimed to enrich for several groups of isolates, including resistant (by BD Phoenix) isolates both with and without beta-lactamases identified from WGS, isolates with peri-breakpoint MICs and susceptible isolates containing beta-lactamases. For each method, sub-cultures (from frozen stocks) were tested in triplicate using ISO-Sensitest agar plates containing amoxicillin and clavulanate in a 2:1 ratio (CLSI) or a fixed concentration of clavulanate (2 mg/L) (EUCAST), with *E. coli* controls ATCC25922 (wild type) and ATCC35218 (TEM-1 beta-lactamase producer).(47) For additional quality control, bacterial isolates were plated on sheep blood agar and incubated overnight at 37°C to check purity, with isolates excluded if multiple

513 colonial morphologies were seen. Isolates were included in analyses if two or more MICs for
514 each of amoxicillin, EUCAST-based amoxicillin-clavulanate and CLSI-based amoxicillin-
515 clavulanate were in essential agreement (i.e. a minimum of 2/3 for each drug; see Supplementary
516 Methods). Isolates with less than two MICs for each of the tests passing quality control were
517 tested an additional time to reduce the risk of selection bias against isolates with genetic
518 mechanisms causing variable expression and underestimating natural phenotypic variability. For
519 each included isolate, susceptibility classification for that isolate was defined using the "upper
520 median" MIC (choosing the higher MIC when the median lay between two MIC readings) of the
521 test repeats.

522

523 *Modelling and predicting MICs*

524 Random-effects models (Stata 14.2; StataCorp LP, 2015) were used to investigate the impact of
525 test method and WGS-identified genetic elements on agar dilution $\log_2$ MICs simultaneously
526 (additional details in Supplementary Methods). Elements were categorised depending on
527 frequency (Supplementary Table S4). Models included method-specific random-effects for each
528 isolate and testing batch, and method-specific (heteroskedastic) errors. All genetic element
529 categories were included *a priori*, but the most predictive effects of each (including
530 presence/absence of genes and/or promoter mutations and/or gene dosage) were selected using
531 the Akaike Information Criterion (AIC) (Supplementary Methods). Lastly, interaction terms
532 between genetic elements (reflecting saturation effects) and with test methodology (reflecting
533 differential impact of the same genetic mechanism depending on the amoxicillin:clavulanate
534 ratio) were included where $p<0.05$. Final estimates were then used to predict MICs in all non-
535 subsample isolates and in non-subsample isolates which did not contain resistance features not

536 present in the agar dilution subsample. Predicted MICs were then compared to routine laboratory
537 phenotypes from the BD Phoenix.

538 **Acknowledgments:**

539 This work uses data provided by patients and collected by the NHS as part of their care and
540 support. We thank all the people of Oxfordshire who contribute to the Infections in Oxfordshire
541 Research Database. Research Database Team: R Alstead, C Bunch, DCW Crook, J Davies, J
542 Finney, J Gearing (community), L O'Connor, TEA Peto (PI), TP Quan, J Robinson
543 (community), B Shine, AS Walker, D Waller, D Wyllie. Patient and Public Panel: G Blower, C
544 Mancey, P McLoughlin, B Nichols. We would also like to thank the HPRU Steering Group (N
545 French, C Marwick, J Coia, M Sharland).

546 **Funding**:

547 The study was funded by the National Institute for Health Research Health Protection Research
548 Unit (NIHR HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at Oxford
549 University in partnership with Public Health England (PHE) [grant HPRU-2012-10041]. DWC,
550 TEAP and ASW are supported by the NIHR Oxford Biomedical Research Centre. The report
551 presents independent research funded by the National Institute for Health Research. The views
552 expressed in this publication are those of the authors and not necessarily those of the NHS, the
553 National Institute for Health Research, the Department of Health or Public Health England. NS is
554 funded by a PHE/University of Oxford Clinical Lectureship. DWC, TEAP and ASW are NIHR
555 Senior Investigators.

556 **Author contributions**:

TJD, NS, MJE, NW, DWC, TEAP, MFA and ASW designed the study. KJ, MA(OUH), MM, TPQ obtained the automated susceptibility phenotypes from archived BD phoenix records. TJD, and MA(APHA) performed agar dilution on samples. DG and AV sequenced isolates. TJD, HP and JS ran resistance genotype prediction on samples. TJD, AS, NS, OE, RB and AM interpreted the genetic results and established rules regarding the relationship with phenotype. TJD and ASW fitted random-effects models to the data. TJD, NS, AS, PF, ASW and MFA prepared the first draft. All authors commented on the data and its interpretation, revised the content critically and approved the final version.

Competing interests:

None to declare. However, for NW and MJE, PHE's AMRHAI Reference Unit has received financial support for conference attendance, lectures, research projects or contracted evaluations from numerous sources, including: Accelerate Diagnostics, Achaogen Inc, Allecra Therapeutics, Amplex, AstraZeneca UK Ltd, AusDiagnostics, Basilea Pharmaceutica, Becton Dickinson Diagnostics, bioMérieux, Bio-Rad Laboratories, The BSAC, Cepheid, Check-Points B.V., Cubist Pharmaceuticals, Department of Health, Enigma Diagnostics, European Centre for Disease Prevention and Control, Food Standards Agency, GlaxoSmithKline Services Ltd, Helperby Therapeutics, Henry Stewart Talks, IHMA Ltd, Innovate UK, Kalidex Pharmaceuticals, Melinta Therapeutics, Merck Sharpe & Dohme Corp, Meiji Seika Pharma Co., Ltd, Mobidiag, Momentum Biosciences Ltd, Neem Biotech, NIHR, Nordic Pharma Ltd, Norgine Pharmaceuticals, Rempex Pharmaceuticals Ltd, Roche, Rokitan Ltd, Smith & Nephew UK Ltd, Shionogi & Co. Ltd, Trius Therapeutics, VenatoRx Pharmaceuticals, Wockhardt Ltd., and the World Health Organization.

Data and materials availability:

580 Sequences used in the study are made available at PRJNA540750. MIC data and code used for
581 this analysis are available at
582 https://github.com/TimothyJDavies/reconciling_the_potentially_irreconcilable

583

584

585 References:

586 1. Public Health England. 2018. English Surveillance Programme for Antimicrobial
587 Utilisation and Resistance (ESPAUR).

588 2. Plachouras D, Kärki T, Hansen S, Hopkins S, Lyytikäinen O, Moro ML, Reilly J, Zarb P,
589 Zingg W, Kinross P, Weist K, Monnet DL, Suetens C, Group the PPSS. 2018.
590 Antimicrobial use in European acute care hospitals: results from the second point
591 prevalence survey (PPS) of healthcare-associated infections and antimicrobial use, 2016 to
592 2017. Eurosurveillance 23:1800393.

593 3. Diez-Aguilar M, Morosini M-I, Lopez-Cerero L, Pascual A, Calvo J, Martinez-Martinez
594 L, Marco F, Vila J, Ortega A, Oteo J, Canton R. 2015. Performance of EUCAST and
595 CLSI approaches for co-amoxiclav susceptibility testing conditions for clinical
596 categorization of a collection of Escherichia coli isolates with characterized resistance
597 phenotypes. J Antimicrob Chemother 70:2306–2310.

598 4. Mouton J. 2017. Antimicrobial susceptibility testing with EUCAST breakpoints and
599 methods, p. . *In* Mouton, J (ed.), ECCMID. Vienna, Austria.

600 5. Vihta K-D, Stoesser N, Llewelyn MJ, Quan TP, Davies T, Fawcett NJ, Dunn L, Jeffery K,
601 Butler CC, Hayward G, Andersson M, Morgan M, Oakley S, Mason A, Hopkins S, Wyllie
602 DH, Crook DW, Wilcox MH, Johnson AP, Peto TEA, Walker AS. 2018. Trends over time
603 in Escherichia coli bloodstream infections, urinary tract infections, and antibiotic
604 susceptibilities in Oxfordshire, UK, 1998-2016: a study of electronic health records.
605 Lancet Infect Dis 18:1138–1149.

606 6. Kallonen T, Brodrick HJ, Harris SR, Corander J, Brown NM, Martin V, Peacock SJ,

607 Parkhill J. 2017. Systematic longitudinal survey of invasive Escherichia coli in England
608 demonstrates a stable population structure only transiently disturbed by the emergence of
609 ST131. Genome Res.

610 7. Drawz SM, Bonomo RA. 2010. Three decades of beta-lactamase inhibitors. Clin
611 Microbiol Rev 23:160–201.

612 8. Bush K. 2015. A resurgence of beta-lactamase inhibitor combinations effective against
613 multidrug-resistant Gram-negative pathogens. Int J Antimicrob Agents 46:483–493.

614 9. European Committee on Antimicrobial Susceptibility Testing. 2017. Guidance document
615 on beta-lactamase inhibitor combinations.

616 10. Humphries R. 2018. Mechanisms of resistance to ceftazidime- avibactamECCMID.
617 Madrid, Spain.

618 11. Ellington MJ, Ekelund O, Aarestrup FM, Canton R, Doumith M, Giske C, Grundman H,
619 Hasman H, Holden MTG, Hopkins KL, Iredell J, Kahlmeter G, Koser CU, MacGowan A,
620 Mevius D, Mulvey M, Naas T, Peto T, Rolain J-M, Samuelsen O, Woodford N. 2017. The
621 role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report
622 from the EUCAST Subcommittee. Clin Microbiol Infect 23:2–22.

623 12. Harris PNA, Tambyah PA, Lye DC, Mo Y, Lee TH, Yilmaz M, Alenazi TH, Arabi Y,
624 Falcone M, Bassetti M, Righi E, Rogers BA, Kanj S, Bhally H, Iredell J, Mendelson M,
625 Boyles TH, Looke D, Miyakis S, Walls G, Al Khamis M, Zikri A, Crowe A, Ingram P,
626 Daneman N, Griffin P, Athan E, Lorenc P, Baker P, Roberts L, Beatson SA, Peleg AY,
627 Harris-Brown T, Paterson DL, MERINO Trial Investigators and the Australasian Society
628 for Infectious Disease Clinical Research Network (ASID-CRN). 2018. Effect of

629 Piperacillin-Tazobactam vs Meropenem on 30-Day Mortality for Patients With *E coli* or
630 *Klebsiella pneumoniae* Bloodstream Infection and Ceftriaxone Resistance. JAMA
631 320:984.

632 13. Livermore DM, Day M, Cleary P, Hopkins KL, Toleman MA, Wareham DW, Wiuff C,
633 Doumith M, Woodford N. 2018. OXA-1 β-lactamase and non-susceptibility to
634 penicillin/β-lactamase inhibitor combinations among ESBL-producing Escherichia coli. J
635 Antimicrob Chemother dky453-dky453.

636 14. Ortega A, Oteo J, Aranzamendi-Zaldumbide M, Bartolome RM, Bou G, Cercenado E,
637 Conejo MC, Gonzalez-Lopez JJ, Marin M, Martinez-Martinez L, Merino M, Navarro F,
638 Oliver A, Pascual A, Rivera A, Rodriguez-Bano J, Weber I, Aracil B, Campos J. 2012.
639 Spanish multicenter study of the epidemiology and mechanisms of amoxicillin-
640 clavulanate resistance in Escherichia coli. Antimicrob Agents Chemother 56:3576–3581.

641 15. Livermore DM. 1997. beta-Lactamases: quantity and resistance. Clin Microbiol Infect 3
642 Suppl 4:S10–S19.

643 16. Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Del Ojo Elias C, Johnson JR,
644 Walker AS, Peto TEA, Crook DW. 2013. Predicting antimicrobial susceptibilities for
645 Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. J
646 Antimicrob Chemother 68:2234–2244.

647 17. Zankari E, Hasman H, Kaas RS, Seyfarth AM, Agerso Y, Lund O, Larsen MV, Aarestrup
648 FM. 2013. Genotyping using whole-genome sequencing is a realistic alternative to
649 surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob
650 Chemother 68:771–777.

18. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu C-H, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. 2019. Validating the NCBI AMRFinder Tool and Resistance Gene Database Using Antimicrobial Resistance Genotype-Phenotype Correlations in a Collection of NARMS Isolates. Antimicrob Agents Chemother AAC.00483-19.

19. 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. J Antimicrob Chemother 70:2763–2769.

20. Shelburne SA, Kim J, Munita JM, Sahasrabhojane P, Shields RK, Press EG, Li X, Arias CA, Cantarel B, Jiang Y, Kim MS, Aitken SL, Greenberg DE. 2017. Whole-Genome Sequencing Accurately Identifies Resistance to Extended-Spectrum beta-Lactams for Major Gram-Negative Bacterial Pathogens. Clin Infect Dis 65:738–745.

21. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, Harris SR. 2017. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microb genomics 3:e000131.

22. Ambler RP, Coulson AF, Frère JM, Ghuysen JM, Joris B, Forsman M, Levesque RC, Tiraby G, Waley SG. 1991. A standard numbering scheme for the class A beta-lactamases. Biochem J 276 (Pt 1:269–70.

23. Shen Z, Ding B, Bi Y, Wu S, Xu S, Xu X, Guo Q, Wang M. 2017. CTX-M-190, a Novel beta-Lactamase Resistant to Tazobactam and Sulbactam, Identified in an Escherichia coli Clinical Isolate. Antimicrob Agents Chemother 61.

673 24. Brumfitt W, Hamilton Miller JMT, Dixson S, Gargan RA, Gooding A. 1983. Phenomenon
674 of resistance to Augmentin associated with sensitivity to ampicillin: Occurrence and
675 explanation. J Clin Pathol 36:670–673.

676 25. Brumfitt W, Hamilton-Miller JM. 1983. Differential sensitivities to
677 amoxycillin/clavulanate and amoxycillin in general practice urinary infections. Lancet
678 (London, England). England.

679 26. Su M, Satola SW, Read TD. 2019. Genome-based prediction of bacterial antibiotic
680 resistance. J Clin Microbiol.

681 27. 2018. EUCAST Frequently Asked Questions EUCAST Frequently Asked Questions 1.
682 EUCAST Disk Diffusion Test-Medium.

683 28. Haidar G, Clancy CJ, Chen L, Samanta P, Shields RK, Kreiswirth BN, Nguyen MH. 2017.
684 Identifying Spectra of Activity and Therapeutic Niches for Ceftazidime-Avibactam and
685 Imipenem-Relebactam against Carbapenem-Resistant Enterobacteriaceae. Antimicrob
686 Agents Chemother 61:e00642-17.

687 29. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S,
688 Hsu C-H, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J,
689 Folster JP, Klimke W. 2019. Using the NCBI AMRFinder Tool to Determine
690 Antimicrobial Resistance Genotype-Phenotype Correlations Within a Collection of
691 NARMS Isolates. bioRxiv 550707.

692 30. Díez-Aguilar M, Morosini M-I, López-Cerero L, Pascual Á, Calvo J, Martínez-Martínez
693 L, Marco F, Vila J, Ortega A, Oteo J, Cantón R. 2015. Performance of EUCAST and
694 CLSI approaches for co-amoxiclav susceptibility testing conditions for clinical

categorization of a collection of *Escherichia coli* isolates with characterized resistance phenotypes. J Antimicrob Chemother 70:2306–2310.

31. Ortega A, Oteo JJ, Aranzamendi-Zaldumbide M, Bartolome RM, Bou GG, Cercenado E, Conejo MC, Gonzalez-Lopez JJ, Marin M, Martinez-Martinez L, Merino MM, Navarro F, Oliver A, Pascual A, Rivera A, Rodriguez-Bano J, Weber I, Aracil BB, Campos JJ, Bartolomé RM, Bou GG, Cercenado E, Conejo MC, José González-López J, Marín M, Martínez-Martínez L, Merino MM, Navarro F, Oliver A, Pascual Á, Rivera A, Rodríguez-Baño J, Weber I, Aracil BB, Campos JJ. 2012. Spanish multicenter study of the epidemiology and mechanisms of amoxicillin-clavulanate resistance in Escherichia coli. Antimicrob Agents Chemother 56:3576–3581.

32. Spafford K, MacVane S, Humphries R. 2019. Evaluation of empiric beta-lactam susceptibility prediction among Enterobacteriaceae by molecular beta-lactamase gene testing. J Clin Microbiol JCM.00674-19.

33. Kim J, Greenberg DE, Pfifer R, Jiang S, Xiao G, Xie Y, Shelburne SA, Koh A, Zhan X. 2019. VAMPr: VAriant Mapping and Prediction of antibiotic resistance via explainable features and machine learning. bioRxiv 537381.

34. 2000. Determination of minimum inhibitory concentrations (MICs) of antibacterial agents by agar dilution. Clin Microbiol Infect 6:509–515.

35. CLSI. 2019. CLSI supplement M100,Performance Standards for Antimicrobial Susceptibility Testing, 29th ed. Wayne, PA.

36. María D-A, María-Isabel M, María-Carmen C, Álvaro P, Jorge C, Luis M-M, Francesc M, Jordi V, Adriana O, Jesús O, Rafael C. 2016. Establishing the validity of different

susceptibility testing methods to evaluate the in vitro activity of amoxicillin-clavulanate against Escherichia coli. Diagn Microbiol Infect Dis 84:334–336.

37. Finney JM, Walker AS, Peto TEA, Wyllie DH. 2011. An efficient record linkage scheme using graphical analysis for identifier error detection. BMC Med Inform Decis Mak 11:7.

38. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829.

39. Stoesser N, Sheppard AE, Peirano G, Anson LW, Pankhurst L, Sebra R, Phan HTT, Kasarskis A, Mathers AJ, Peto TEA, Bradford P, Motyl MR, Walker AS, Crook DW, Pitout JD. 2017. Genomic epidemiology of global Klebsiella pneumoniae carbapenemase (KPC)-producing Escherichia coli. Sci Rep 7:5917.

40. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MCJ, Ochman H, Achtman M. 2006. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol Microbiol 60:1136–1151.

41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

42. Wu P-J, Shannon K, Phillips I. 1995. Mechanisms of hyperproduction of TEM-1 β-lactamase by clinical isolates of *Escherichia coli*. J Antimicrob Chemother 36:927–939.

43. Shropshire WC, Aitken SL, Pifer R, Kim J, Bhatti MM, Li X, Kalia A, Galloway-Peña J, Sahasrabhojane P, Arias CA, Greenberg DE, Hanson BM, Shelburne SA. Title: β-lactamase Amplification and Porin Loss Drive Progressive β-lactam Resistance in Recurrent ESBL Enterobacteriaceae Bacteremia One Sentence Summary: Amplification of β-lactam genes with porin loss was associated with development of β-lactam resista.

739 44. Lartigue MF, Leflon-Guibout V, Poirel L, Nordmann P, Nicolas-Chanoine M-H. 2002.
740 Promoters P3, Pa/Pb, P4, and P5 upstream from bla(TEM) genes and their relationship to
741 beta-lactam resistance. Antimicrob Agents Chemother 46:4035–4037.

742 45. Peter-Getzlaff S, Polsfuss S, Poledica M, Hombach M, Giger J, Bottger EC, Zbinden R,
743 Bloemberg G V. 2011. Detection of AmpC beta-lactamase in Escherichia coli:
744 comparison of three phenotypic confirmation assays and genetic analysis. J Clin Microbiol
745 49:2924–2932.

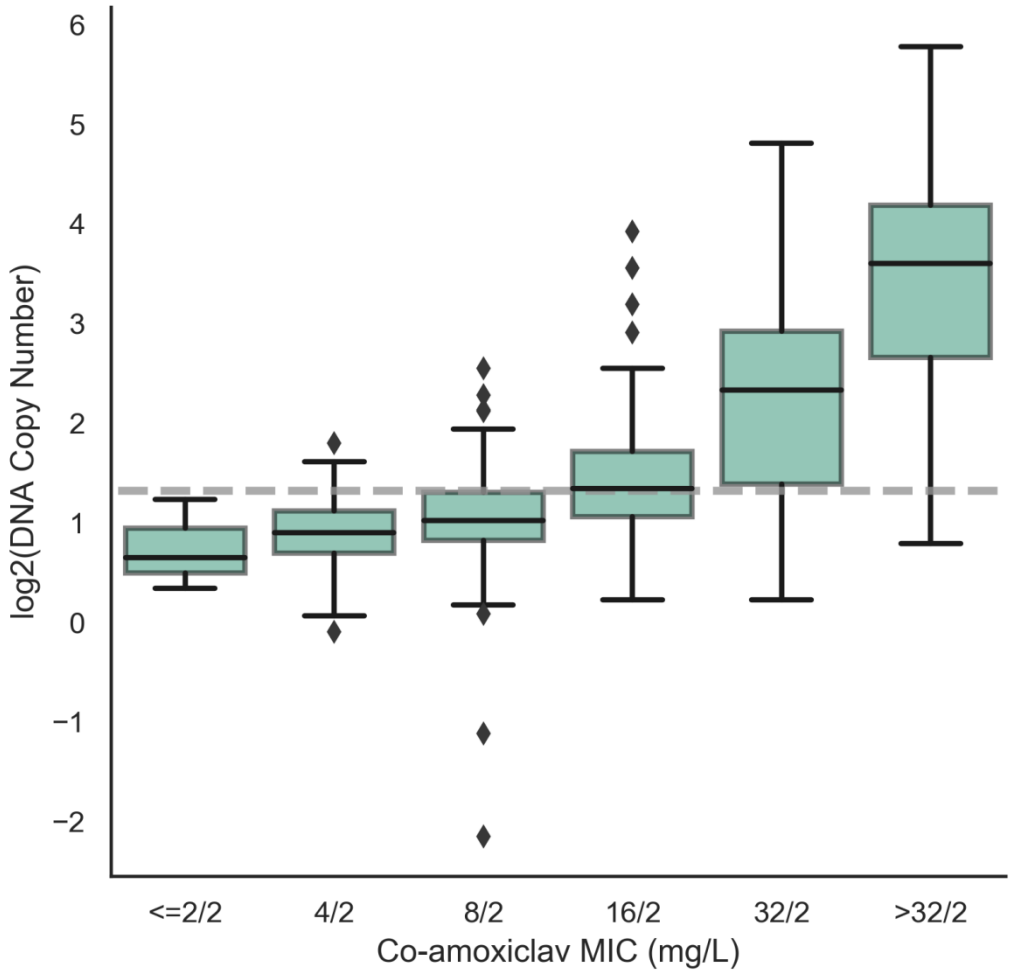
746 46. Caroff N, Espaze E, Bérard I, Richet H, Reynaud A. 1999. Mutations in the ampC
747 promoter of Escherichia coli isolates resistant to oxyiminocephalosporins without
748 extended spectrum β-lactamase production. FEMS Microbiol Lett 173:459–465.

749 47. Andrews JM. 2001. Determination of minimum inhibitory concentrations. J Antimicrob
750 Chemother 48:5–16.

751

752

Figure 1: Association between transmissible beta-lactamase gene DNA copy number and amoxicillin-clavulanate MIC in isolates with no alternative resistance features.



Note: Evidence for association between MIC and log2(DNA copy number) $p<0.0001$, estimated using quantile regression. Grey line indicates 2.5 threshold used to define resistance in the extended algorithm based on ROC analysis (Supplementary Figure S3A). Of these 328 isolates, 294 had bla_{TEM} genes (290 with $bla_{\text{TEM-1}}$, 4 with other non-inhibitor resistant bla_{TEM} genes), 19 had non-inhibitor resistant bla_{SHV} genes and 15 had $bla_{\text{CTX-M}}$ genes.

Figure 2: Proportion WGS predicted resistant (extended algorithm) by routine laboratory MIC

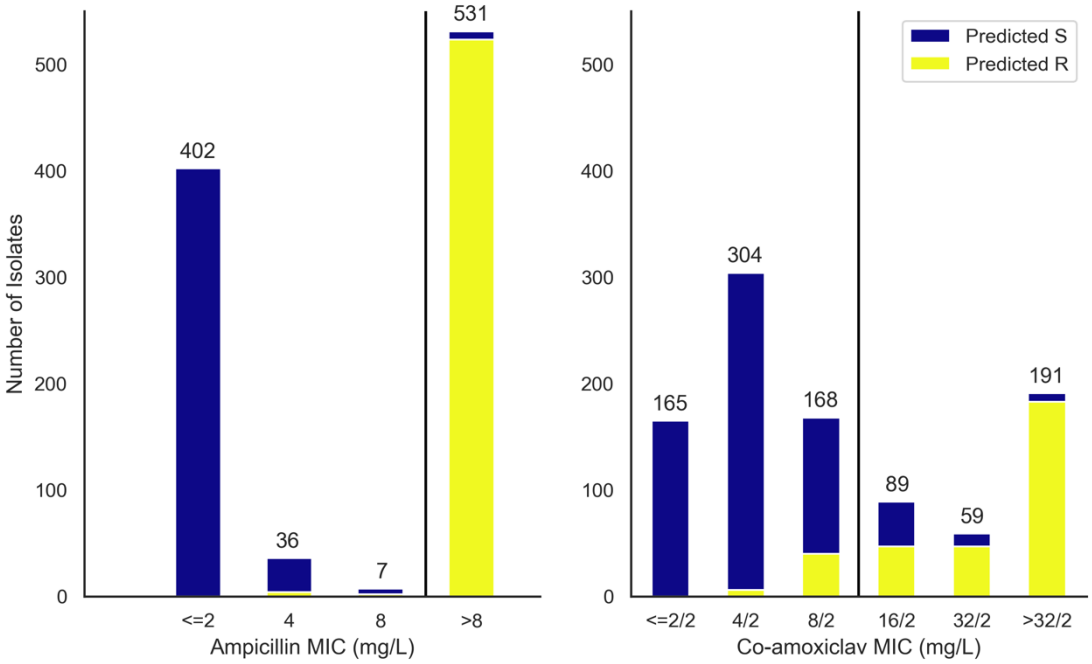
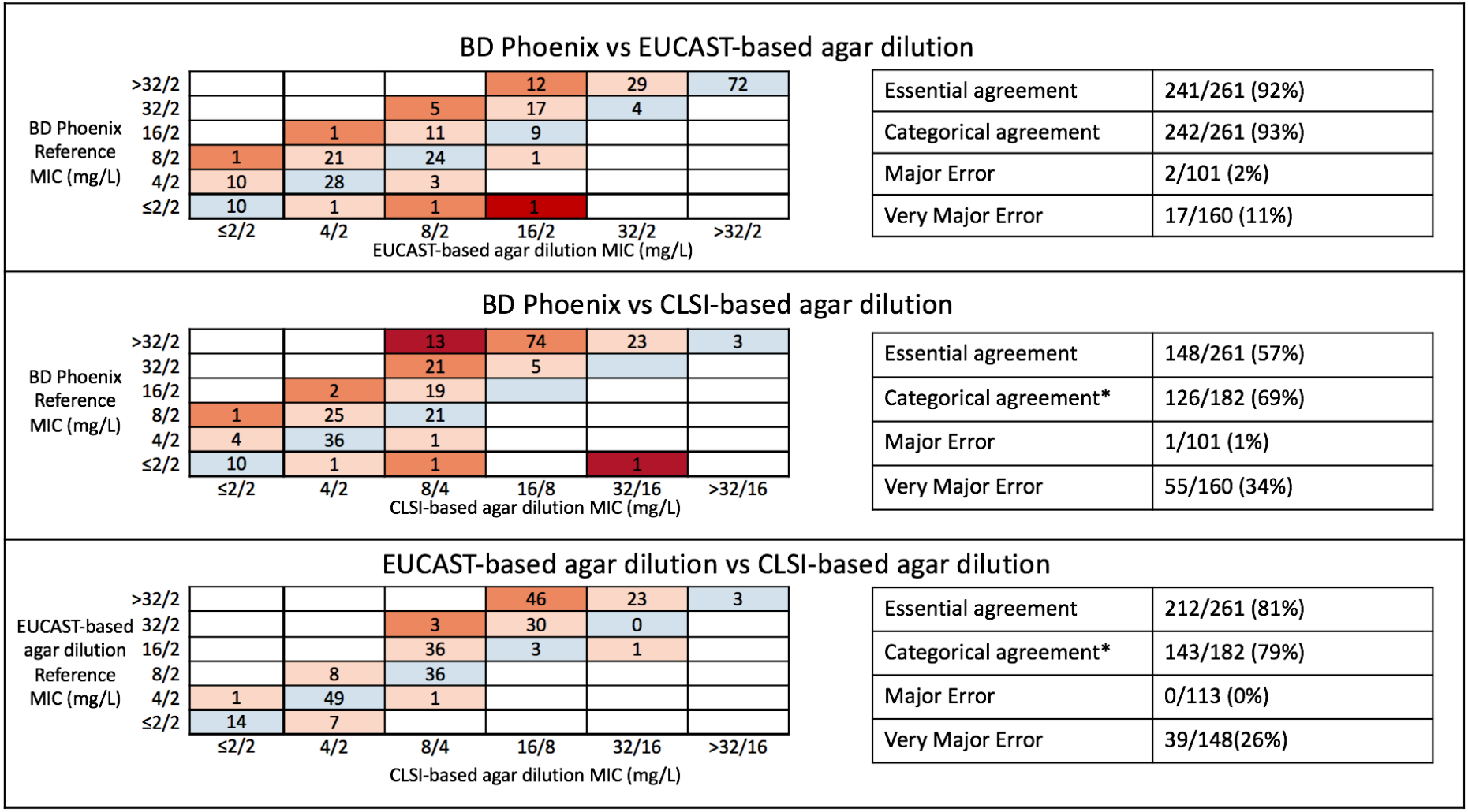
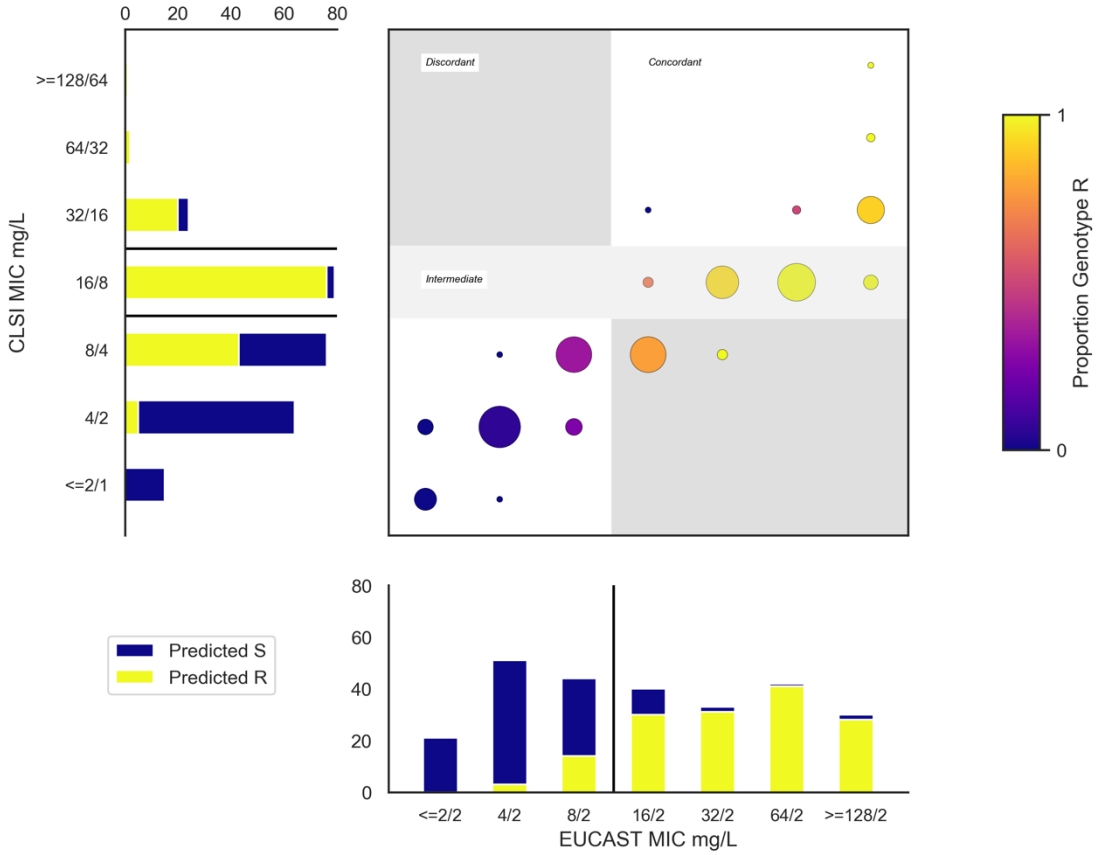


Figure 3: Comparison of the three different phenotyping methods on the agar-dilution subsample isolates (N=261)



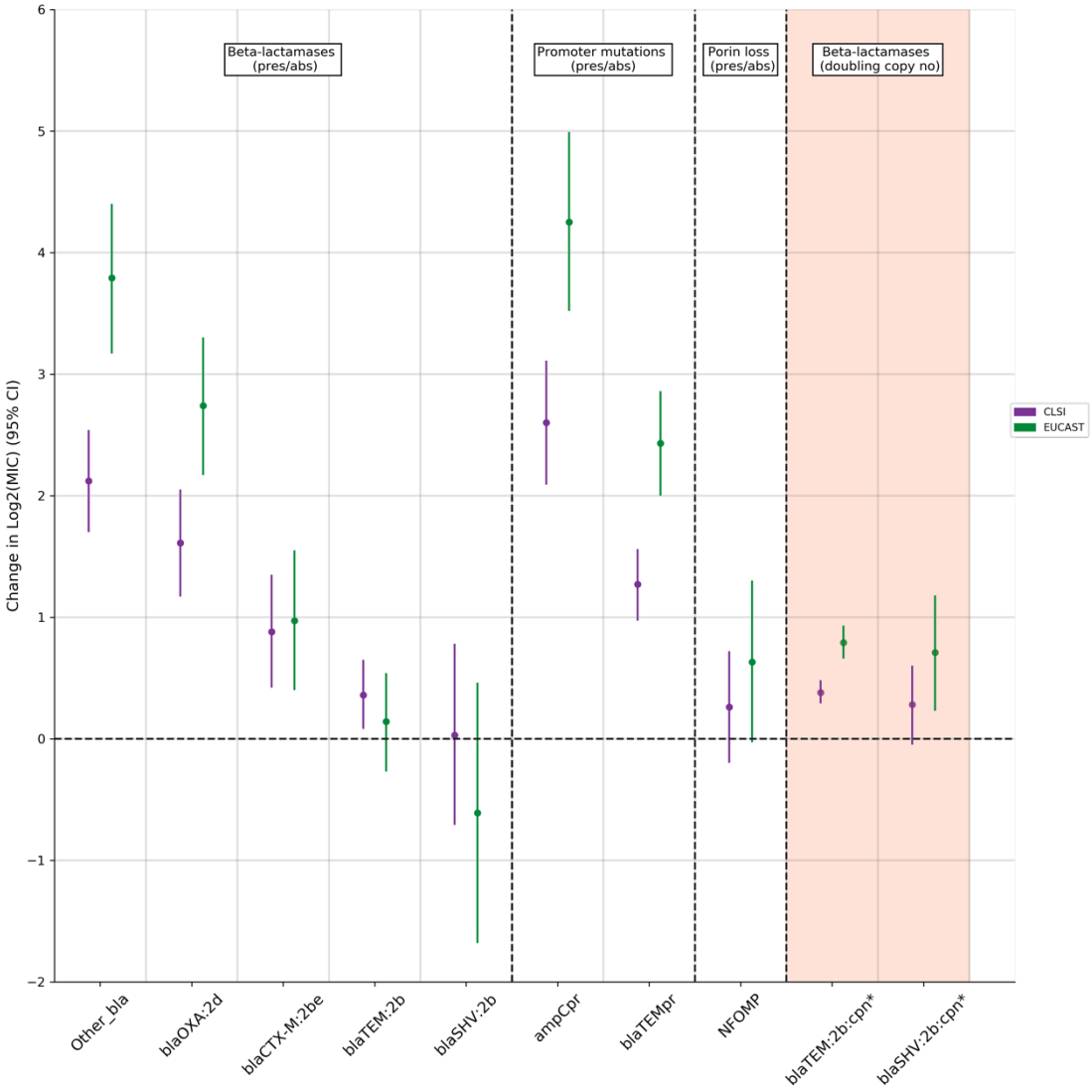
Note: Comparison of MICs obtained using the three different phenotyping methods; EUCAST-based agar-dilution, CLSI-based agar dilution and BD Phoenix (performed in the OUH microbiology laboratory and using panels calibrated against EUCAST guidelines). *: Isolates are in categorical agreement if they are reported as either resistant by both methods (i.e. BD Phoenix/EUCAST-based agar dilution MIC >8/2 mg/L and CLSI-based agar dilution > 16/8 mg/L) or susceptible by both methods (i.e. BD Phoenix/EUCAST-based agar dilution MIC <=8/2 mg/L and CLSI-based agar dilution MIC <= 8/4 mg/L). Intermediate isolates were excluded from these comparisons (but are shown above) as BD phoenix/EUCAST-based agar dilution have no intermediate category. Blue: full agreement of MICs, light orange: essential agreement, dark orange: within two doubling dilutions (theoretically feasible believing both tests having an error of +/- 1 dilution) and red: disagreement.

Figure 4 Proportion WGS predicted resistant (extended algorithm) by MICs from EUCAST and CLSI-based methods



Note: Main panel, each (x,y) coordinate represents (EUCAST-based MIC,CLSI-based MIC) combination. At each coordinate, circle size represents the number of isolates with this combination of fixed and ratio MICs, and color denotes proportion identified as resistant by WGS as indicated by the color bar to the right of the figure. The two sub-panels (bar charts to the left and bottom of the main panel) show the number of isolates with each MIC (in line with the main panel). Yellow/blue coloring indicate which of these were predicted resistant/susceptible respectively, and black lines indicate cut-offs used to determine resistance classification (susceptible/resistant for EUCAST-based agar dilution, susceptible/intermediate/resistant for CLSI-based agar dilution).

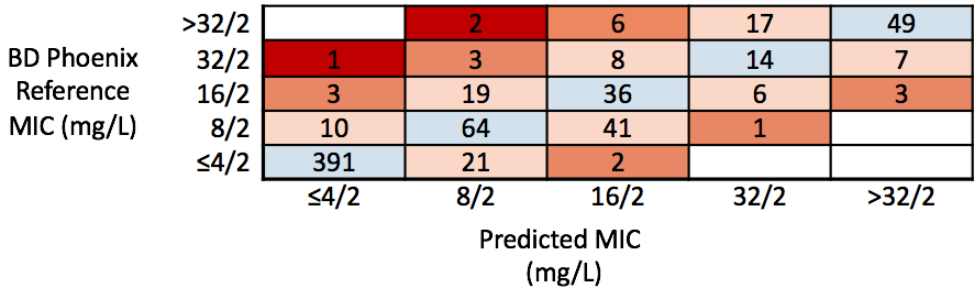
Figure 5: Changes in doubling dilution MIC independently associated with each feature/testing method (multivariable random-effects model).



Note: Purple represents testing using 2:1 CLSI-based agar dilution (CLSI), and green using EUCAST-based agar diltuion. All elements except those denoted by * and shaded in orange are modelled as binary presence vs absence effects (see Supplementary Methods).), Other_bla (grouped other bla genes, includes bla_{TEM-40} (N=2), bla_{TEM-30}(N=3), bla_{CMY-2} (N=3), bla_{OXA-48} (N=1), $bla_{TEM-190}$ (N=1), bla_{TEM-33} (N=1), Supplementary Table 2) blaOXA:2d (Bush-Jacoby 2d, bla_{OXA}), blaCTXM:2be (Bush-Jacoby 2be, CTXM), blaTEM:2b (Bush-Jacoby 2b, bla_{TEM}), blaSHV:2b, (Bush-Jacoby 2b, SHV), ampCpr (ampC promoter mutation suggesting increased

expression), blaTEMpr(bla_{TEM} hyper-producing promoter), NFOMP(non-functional ompF/ompC), blaTEM:2b:cpn (copy number) effect modelled as effect of doubling copy number, blaSHV:2b:cpn (copy number) effect modelled as effect of doubling copy number

Figure 6 Model based MIC prediction for non-subsample isolates (N=704)



Essential agreement	683/704 (97%)
Categorical agreement	632/704 (90%)
Major Error	44/530 (8%)
Very Major Error	28/174(16%)

Note: blue shading indicates correct correctly predicted observed AST MIC (554/704 (79%) isolates), light pink predicted within one doubling dilution (total 683/704 (97%) isolates, essential agreement), orange within two doubling dilutions (total 701/704 (100%)) and red greater than 2 doubling dilutions. Excluding 11 isolates with resistance mechanisms not included in the agar-dilution subsample on which the prediction model was derived (similar overall performance including these).

Table 1: Performance of WGS-based prediction using both basic and extended algorithms

WGS-predicted phenotype (No. isolates)	**AST Phenotype (No. isolates)**		**Sensitivity (%)**	**Specificity (%)**
Ampicillin: basic	S	R	96	99
S	441	20		
R	4	511		
Ampicillin: extended prediction	S	R	98	99
S	439	8		
R	6	523		
Amoxicillin-clavulanate: basic	S	R	23	100
S	634	261		
R	3	78		
Amoxicillin-clavulanate: extended	S	R	82	93
S	591	62		
R	46	277		

Table 2: Resistance prediction feature performance

Feature	Prediction	Number of isolates resistant by routine AST/total isolates with this feature (%)	
		PPV	PPV when in isolation[+]
Beta-lactamases			
1 Any class C or D serine beta-lactamase	Basic	66/69 (96%)	32/34 (94%)
2 Any inhibitor resistant class A beta-lactamase	Basic	12/12 (100%)	2/2 (100%)
Promoter mutations			
3 Non-P3 *bla*$_{TEM}$ promoter associated with *bla*$_{TEM}$ hyper-production	Extended	48/49 (98%)	29/30 (97%)
4 - *ampC* promoter mutation associated with a*mpC* hyper-production	Extended	20/20 (100%)	13/13 (100%)
Increased DNA copy number			
5 Relative coverage of any transmissible beta-lactamase > 2.5*	Extended	184/227 (81%)	128/167 (77%)
Decreased permeability[o]			
6 Features suggesting disruption of either *ompC* or *ompF* in an isolate containing an additional beta-lactamase (see supplementary methods)	Extended	19/22 (86%)	2/2 (100%)

+: PPV restricted to isolates not predicted as amoxicillin-clavulanate resistant by any other feature

*: Cut-off chosen following a receiver operating curve (ROC) analysis (Supplementary Methods, Supplementary Figure 3A)

[o]: Beta-lactam resistance features in isolates shown in Supplementary Tables S2, S3, S4

Note: PPV=positive predictive value

Supplementary material for: Reconciling the potentially irreconcilable? Genotypic and phenotypic amoxicillin-clavulanate resistance in *Escherichia coli*

List of contents