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

Timothy J. Davies[1,2], Nicole Stoesser[1,2,3], Anna E Sheppard[1], Manal Abuoun[4], Philip Fowler[1], Jeremy Swann[1], T. Phuong Quan[1,2], David Griffiths[1], Alison Vaughan[1], Marcus Morgan[3], Hang TT Phan[1], Katie J Jeffery[3], Monique Andersson[3], Matt J Ellington[5], Oskar Ekelund[6], Amy J. Mathers[7], Robert A. Bonomo[8], Neil Woodford[2,5], Derrick W. Crook[1,2,9], Tim E.A. Peto[1,2,3], Muna F Anjum[4], A. Sarah Walker[1,2]

1. Nuffield Department of Medicine, Oxford University, Oxford, United Kingdom
2. National Institute for Health Research (NIHR) Health Protection Research Unit on Healthcare Associated Infections and Antimicrobial Resistance at University of Oxford
3. Oxford University Hospitals NHS Foundation Trust, Oxford, UK
4. Bacteriology, Animal and Plant Health Agency, Surrey UK
5. Antimicrobial Resistance and Healthcare Associated Infections (AMRHAI) Reference Unit, National Infection Service, Public Health England, London UK
6. Department of Clinical Microbiology and the EUCAST Development Laboratory, Kronoberg Region, Central Hospital, Växjö, Sweden
7. 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
8. 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
9. National Infection Service, Public Health England, UK

Corresponding Author: Timothy J Davies, Microbiology Level 7, John Radcliffe Hospital, Headley Way, Oxford, OX3 9DU. Email timothy.davies@ndm.ox.ac.uk Tel: +447545593235

ABSTRACT

Background

Resistance to co-amoxiclav in *Escherichia coli* is rising globally, yet susceptibility testing remains challenging as different methods yield different results. Predicting co-amoxiclav susceptibility via whole-genome sequencing (WGS) may be more reliable.

Methods

976 isolates from 968 *E. coli* bloodstream infection cases occurring in Oxfordshire, UK between January 2013-August 2015 were sequenced and phenotyped (BD Phoenix); a random stratified subsample were phenotyped in triplicate by agar dilution using two clavulanate concentration regimes (2:1 amoxicillin:clavulanate ratio; fixed 2mg/L clavulanate concentration). Each phenotype was compared with two WGS-derived resistance predictions, one “basic” excluding, and one “extended” including, features associated with penicillinase hyper-production.

Findings

340(35%) isolates were co-amoxiclav-resistant. Extended genotype features, while crucial for identifying resistance (sensitivity: 82% (277/339) extended vs 23% (78/339) basic; $p<0.0001$), also modestly lowered specificity (93% (591/637) extended vs 100% (634/637) basic). Only 222/261(89%) isolates tested by agar-dilution were concordantly susceptible/non-susceptible between the two clavulanate concentrations. Sensitivity and specificity of WGS to predict resistance depended on phenotyping method (ranging from 85-93% and 47-85% respectively). Several genetic features had inconsistent impacts on phenotype and only caused modest minimum inhibitory concentrations (MICs) changes. However, genotype-model predicted MICs were within one doubling dilution of observed MICs for 681/704 (97%) non-subsample isolates where predictions were possible.

Interpretation

WGS accurately predicts co-amoxiclav MIC in *E. coli*, provided mechanisms causing penicillinase hyper-production are considered. Binary interpretations (susceptible/resistant) for co-amoxiclav are poorly reproducible, regardless of phenotyping methodology, and should be reconsidered. Work is needed to evaluate the impact of genotypes/MIC on clinical outcomes.

Funding

National Institutes for Health Research.

RESEARCH IN CONTEXT

Evidence before the study

We searched PubMed for publications from inception up until January 1, 2019 using the terms "*Escherichia coli*", "co-amoxiclav", "drug resistance" "sequencing", and their synonyms, and also reviewed references of retrieved articles and articles identified as "similar articles" by PubMed. Incidence of co-amoxiclav resistance is rising worldwide. Significant discrepancies in characterising co-amoxiclav resistance by different phenotyping methods are observed, raising doubts about the diagnostic validity of drug susceptibility testing. One potential solution is to instead track the genetic determinants of resistance. Mechanisms of resistance to co-amoxiclav, like all beta-lactam/beta-lactamase inhibitor combinations, are diverse and include mechanisms that regulate expression, so previous studies have used a combination of PCR or microarray-based tests. An alternative would be whole-genome sequencing (WGS): for other organisms this has provided a powerful way of identifying resistance and understanding epidemiological associations. Given the complexity of the relationship between genetic variation and resistance, the evidence to date that WGS can reliably predict co-amoxiclav resistance in *E. coli*, particularly in population representative isolates, is limited.

Added value of this study

Our study investigated whether it was possible to predict co-amoxiclav susceptibility from WGS from unselected *E. coli* bloodstream infections. By examining the WGS and drug susceptibility testing data from many population-representative clinical isolates, we found that:

- Rather than just relying on presence and absence of gene(s) or mutation(s), additional data from WGS such as relative gene copy number and promoter mutations are essential to identify co-amoxiclav resistance accurately.
- As well as disagreeing with each other, neither of two currently widely used reference phenotypic methods aligned perfectly with the genetic features found. Further, several genetic features had inconsistent effects on phenotype, being found in both resistant and susceptible isolates.
- The effects of some resistance features individually on minimum inhibitory concentrations (MICs) were small, variable and potentially additive. As the distribution of co-amoxiclav MICs spans the breakpoint(s) defining resistance and susceptibility, these changes substantially affect reproducibility and interpretation.

Implications of all available evidence

Co-amoxiclav resistance in *E. coli* is better thought of as quantitative, rather than qualitative, and methods that infer resistance from WGS data should be adapted to identify genetic features associated with incremental, but individually modest, changes to MICs. While our results are promising, our efforts to validate our WGS-based approach against "gold-standard" culture-based methods have instead uncovered broader issues of reproducibility. The current paradigm enforces a "resistant" vs "susceptible" dichotomy. In reality, resistance is a continuum built up by many individual features inevitably resulting in poor reproducibility and suboptimal concordance. This was seen in our study, with different culture-based methods frequently failing to agree with one another, and repeated measurements often altering the resistance classification of the sample. Analogously, the binary susceptible/resistant classification derived from our genetics-based predictions failed

to fully agree with either culture-based method. 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.

INTRODUCTION

Rising co-amoxiclav resistance in *E. coli* is perceived as a major healthcare problem, with increasing frequency of resistant bloodstream infections (BSI)[1] threatening co-amoxiclav's utility and status as the most commonly used antibiotic in Europe.[2] Consequently, many hospitals are considering broadening their first-line empiric antibiotics for common infections. However, there is significant uncertainty created by observed differences in the categorization of clinical samples by the two main assays for co-amoxiclav susceptibility.[3] These differences are so large that increasing co-amoxiclav 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 co-amoxiclav resistance,[6] but found no evidence of clonal expansion of any specific co-amoxiclav-resistant strains. However, the genetic epidemiology of co-amoxiclav resistance mechanisms was not investigated.

In addition to its widespread clinical use, co-amoxiclav 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 co-amoxiclav resistance threaten to 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] Rather than resistance being associated with the simple presence/absence of specific genes, previous studies have found much co-amoxiclav resistance is likely attributable to mechanisms which increase the effective concentration of beta-lactamases (e.g. additive effects of multiple beta-lactamases,[12] increasing gene expression[13] or modifying cell permeability[14]). Given this added complexity, studies using WGS to predict phenotypic resistance have either not included co-amoxiclav,[15,16] or have included only small numbers of selected, generally highly resistant isolates.[17] Similar studies investigating other BL/BLIs, such as piperacillin-tazobactam, reported poor accuracy when predicting resistance from genotype.[18]

We therefore investigated concordance between WGS-derived genotypes and co-amoxiclav 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).

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 antimicrobial susceptibility testing (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)[19] 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 -80^{o}C. 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 Table 1) were de-novo assembled using Velvet[20] as previously described.[21] *In silico* Achtman[22] multi-locus sequence types (MLST) types were defined using ARIBA.

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

Using ARIBA[23] (default parameters) and tBLASTn/BLASTn, we compared two genetic resistance prediction algorithms for amoxicillin and co-amoxiclav (Table 1, Supplementary Methods); first a "basic" prediction using only presence/absence of relevant genes in the Resfinder[16] database and second an "extended" prediction, which additionally included bla_{TEM} and *ampC* promoter mutations, estimates of DNA copy number and predicted porin loss-of-function. For bla_{TEM} and *ampC* promoters, sequences identified using ARIBA/BLASTn were searched for variant sites and regions previously associated with significantly increased expression.[24–26] 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 1). 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 using random sampling within strata defined by phenotype-genotype combinations (Supplementary Figures 2, 3) for replicate agar dilution phenotyping with clavulanate concentration and MIC interpretation according to EUCAST (EUCAST-based agar dilution, EAD), and CLSI (CLSI-based agar dilution, CAD) guidelines. 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 (CAD) or a fixed concentration of clavulanate (2 mg/L, EAD), with *E. coli* controls ATCC25922 (wild type) and ATCC35218 (TEM-1 beta-lactamase producer). 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 colonial morphologies were seen. Isolates were included in analyses if two or more MICs were in essential agreement, defining the susceptibility classification for that isolate using the "upper median" MIC (choosing the higher MIC when the median lay between two MIC readings).

Modelling and predicting MICs

Random effects models (Stata 14.2; StataCorp LP, 2015) 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 2). Models included method-specific random effects for each isolate and testing batch, and method-specific (heteroskedastic) errors. All genetic element categories were included *a priori*, but the most predictive effects of each (including presence/absence of genes and/or promoter mutations and/or gene dosage) was selected using the Akaike Information Criterion (AIC) (Supplementary Table 3). Lastly, 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 $p<0{\cdot}05$. Final estimates were then used to predict MICs in all non-subsample isolates which did not contain resistance features not present in the agar

dilution subsample. Predicted MICs were then compared to routine laboratory phenotypes.

Role of the funding source

The study funder had no role in design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

Routine laboratory phenotypes and co-amoxiclav resistance genotypes

From 1039 *E. coli* BSI occurring between January 2013-August 2015, 1008 distinct *E. coli* isolates with complete phenotype data were available from the routine laboratory (representing 1000 [96%] *E. coli* BSI). 976/1008 (97%) sequences passed quality control (representing 968/1039 [93%] *E. coli* BSI) (Supplementary Methods). 340/976 (36%) had co-amoxiclav MIC > 8/2 mg/L by BD Phoenix/EUCAST breakpoints (Supplementary Table 4).

The collection was highly diverse, representing 152 different sequence types (STs). The most common was ST73 (161,17%) (Figure 1), followed by ST131 (124,13%), which had the highest percentage of phenotypically-resistant isolates (N=74,60%) and was the only ST associated with co-amoxiclav resistance (chi-squared p<0·0001 compared with p>0·16 for all other STs).

The most common beta-lactam resistance mechanisms identified were acquired beta-lactamase genes, bla_{TEM} (N=427,44%), $bla_{\text{CTX-M}}$ (N=73,7%), bla_{OXA} (N=62,6%) and bla_{SHV} (N=23,2%) (Figure 1, Supplementary Tables 5-8). For the 594 non-*ampC* beta-lactamases identified, median DNA copy number from mapping coverage was 2·23 (IQR 1·73,3·31), with 237(40%) >2.5 (the threshold derived from ROC analysis). Variant bla_{TEM} and *ampC* promoters considered to be associated with increased expression were identified in 49 (5%) and 20 (2%) isolates respectively (Supplementary Tables 9-10). 30 (3%) of isolates potentially had one non-functional porin, of which 21 also contained a beta-lactamase gene; however, no isolate had "functionally lost" both *ompC* and *ompF*.

WGS-derived resistance prediction compared with routine phenotyping

Including features affecting beta-lactamase expression (i.e. the 'extended' approach – see Methods) had little impact on our ability to identify amoxicillin resistance (sensitivity: 523/531 [98%] extended vs 511/531 [96%] beta-lactamases only [basic]; specificity: 439/445 [99%] extended vs 441/445 [99%] basic). However, including these features proved vital for inferring co-amoxiclav resistance with significantly higher sensitivity with extended 277/339 (82%) versus basic 78/339 (23%) resistance prediction (McNemar's p<0·0001). Increased sensitivity modestly reduced specificity (93% (591/637) extended vs 100% (634/637) basic). Overall categorical agreement of WGS-derived with observed phenotype increased from 712/976 (73%) to 868/976 (89%) when these genetic features were included.

One reason for lower than optimal agreement with the extended algorithm could be poor performance of individual algorithm components. Presence/absence of beta-lactamase (*bla*) genes had high positive predictive value (PPV) (class C/D[27] beta-lactamase: PPV=97%, inhibitor-resistant class A beta-lactamase: PPV=100%), as did promoter mutations and porin loss (Table 1). Even the most common cause of predicted co-amoxiclav resistance, increased beta-lactamase gene DNA copy number (>2·5), was reasonably successful (PPV=0·77), although only isolates with high copy number (>6) were invariably resistant (Figure 2). Of 79 isolates containing non-inhibitor-resistant beta-lactamases where the algorithm incorrectly predicted susceptibility/resistance, 64 (81%) had peri-breakpoint (8/2-16/2mg/L) MICs. Likewise overall, the extended algorithm performed better at predicting susceptibility category in non-peri-breakpoint isolates, correctly identifying 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 (Figure 3). Peri-breakpoint isolates were harder to classify with 82/108 (76%) incorrect WGS resistance predictions occurring in isolates with MICs 8/2-16/2 mg/L.

We therefore investigated two other hypotheses that could explain the low agreement: (*i*) variable accuracy of the different phenotypic methods, and (*ii*) the binary resistant/susceptible classification being too simplistic.

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

261/291 (90%) isolates selected for agar dilution passed quality controls. The stratified random sampling enriching for resistant phenotypes meant that 160/261 (61%) isolates were co-amoxiclav-resistant by routine AST (Supplementary Table 4). 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 4).

As expected, EUCAST agar dilution (EAD, fixed 2mg/L clavulanate concentration) classified more isolates as more resistant than CLSI agar dilution (CAD, ratio 2:1 amoxicillin:clavulanate) ($\log_2$ MIC difference=0·84, p<0.0001). They were also in closer agreement with routine AST (Supplementary Figure 5), as expected, given BD Phoenix is calibrated against EUCAST guidelines. Using upper median MICs, EAD classified 145 (56%) isolates as resistant, whereas CAD classified 27 (10%) as resistant and 79 (30%) as intermediate (total 106 [41%] CLSI non-susceptible). All

CAD resistant/intermediate isolates were EAD-resistant. Combining results, 27 (10%) isolates were agreed resistant (i.e. resistant by both methods), 116 (44%) agreed susceptible (i.e. susceptible by both methods), and 118 'indeterminate' (EAD-resistant and CLSI susceptible (39[15%]) or intermediate (79[30%])). Overall agreement between EAD resistant/susceptible and CAD non-susceptible/susceptible (85%) was thus similar to that between WGS predictions and routine AST (89%).

EAD showed more variability upon retesting than CAD (Supplementary Figure 6), with 158(61%) isolates having at least one doubling dilution between maximum and minimum MICs across EAD repeats versus 73(28%) for CAD. However essential agreement between repeat MICs was high with only one (CAD) and twelve (EAD) isolates having MICs spanning three or more doubling dilutions. Classifications were more variable, with 40(15%) isolates both resistant and susceptible by EAD, and 31(12%) isolates changing susceptible/intermediate/resistant category by CAD.

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

Overall, WGS classified as resistant 23/27 (85%) EAD/CAD agreed resistant isolates, 107/118 (91%) indeterminate isolates (76/79 EAD-resistant/CAD-intermediate, 31/39 EAD-resistant/CAD susceptible) and 17/116 (15%) agreed susceptible isolates (Figure 4). Only 8/27 (30%) of the agreed resistant samples contained inhibitor-resistant beta-lactamases.

Conversely, 24/79 (30%) CAD-intermediate and 10/39 (26%) CAD-susceptible isolates contained $bla_{\text{OXA-1}}$ showing that identification of inhibitor-resistant beta-lactamases was neither necessary nor sufficient to predict CAD resistance. Similarly,

for other genetic features, assessment of their individual contribution to the phenotype was challenging due to co-occurrence of features in the same isolate and (Supplementary Figure 4) and the impact of some features on susceptibility varying both between isolates (Supplementary Figure 7) and within isolate repeats (Supplementary Figure 6). For example, 4/9 isolates with ampC promoter mutations in the agar dilution subset were found both CAD "R" and CAD "I" on repeat testing.

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

WGS predictions were more accurate for non-peri-breakpoint MICs (EAD: (≤4/2 mg/ml, ≥32/2 mg/ml), CAD: (≤4/2 mg/ml, ≥32/16 mg/ml)). For EAD, WGS correctly identified resistance/susceptibility in 169/177 (95%) of isolates with non-peri-breakpoint MICs, versus only 60/84 (71%) with peri-breakpoint MICs. Similarly, for CAD, discounting 79 intermediate isolates, WGS correctly predicted 97/106 (92%) non-breakpoint isolates, but predicted 43/76 (57%) isolates with MIC 8/4 mg/L resistant. Interestingly, however, there were three consistently resistant (EAD MIC ≥ 32/2mg/L, CAD MIC ≥32/16 mg/L) and three consistently susceptible (EAD MIC ≤ 4/2mg/L, CAD MI C ≤4/2 mg/L) discrepants. All three resistant discrepants were explained by complexities inferring phenotype from WGS. One had a novel $bla_{\text{CTX-M}}$ variant (CTX-M-15-like, Ser130Gly mutation). Previous work on mechanisms of beta-lactamase inhibition suggests mutations at Ambler position[28] 130 likely lead to inhibitor resistance[7] and indeed a similar mutation (Ser130Thr CTX-M-190) resulted in sulbactam and tazobactam resistance.[29] The other two 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*, suggesting either assembly issues or insertion of alternative elements upstream of *ampC* drove increased expression. 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.

Impact of individual resistance features on a continuous measure of susceptibility

Independently, all beta-lactamases were associated with increased MICs in multivariable modelling (Supplementary Table 11, Figure 5). The largest effects of beta-lactamase presence/absence were for $bla_{\text{OXA-1}}$ (i.e. blaOXA:2d model term, Supplementary Methods) and members of the 'other' group of beta-lactamases, comprising either inhibitor resistant beta-lactamases (N=15), or those with unknown impact on beta-lactam susceptibility (N=4) (Supplementary Table 2). These caused two-three fold and four fold doubling dilution increases in EAD MIC respectively. While there was at most weak evidence of a presence/absence effect of non-inhibitor resistant bla_{TEM} (*bla*TEM:2b model term) (CAD p=0·01, EAD p=0·52) or bla_{SHV} (blaSHV:2b model term) (CAD p=0·92, EAD p=0·26), there was stronger evidence of associations with increased copy number for both (i.e. a gene dosage effect). Consistent with unadjusted analyses (Supplementary Table 12), promoter mutations were associated with increased MICs (p<0·0001). However, there was no clear increase in MIC independently associated with suspected porin loss (p>0·10).

Of note, when increased copy number effects were included, EAD (EUCAST) testing methodology accentuated increases in MIC caused by genetic resistance features other than for suspected porin loss and blaCTX-M:2be ($p_{heterogeneity} \leq 0{\cdot}05$). EAD methodology however was also associated with increased between and within sample standard deviation (Supplementary Table 13).

Predictions of MIC in an independent validation set

Overall, the genotype-based random effects model accurately predicted AST MIC for 704/715 non-subsample isolates where predictions were possible (Supplementary Methods). Predicted MIC agreed with observed AST MIC for 548/704 (78%) isolates (Figure 6) and was within one doubling dilution for 681/704 (97%) isolates. As for phenotypic method comparisons, discordance between predicted and observed resistant/susceptible classifications was high despite having high essential agreement of MICs. Three isolates had predicted MICs three doubling dilutions lower than observed. One had an inconsistent phenotype [ampicillin susceptible, co-amoxiclav resistant], suggesting a phenotyping problem. The other two both had observed MIC, ≥32/2mg/L but only contained a low copy number bla_{TEM-1} and had predicted MIC 8/2mg/L.

DISCUSSION

Decisions about broadening recommended empiric antimicrobial regimens from co-amoxiclav are currently being made based on potentially unreliable AST data and an incomplete understanding of the genetic causes of co-amoxiclav resistance. Here, we have confirmed that the mechanisms of resistance are multifactorial, resulting from combinations of multi-copy beta-lactamase genes, mutations in resistance gene-associated promoters, and inhibitor resistance (resistance to clavulanic acid inactivation). Further, the individual effects of some of these features on MIC were small, variable and potentially additive, resulting in only minor shifts around clinical breakpoints, potentially explaining inconsistencies on repeated phenotyping for the same isolate and discrepancies between genotypic predictions and phenotypic susceptible/resistant classifications. Finally, the phenotypic testing methodology significantly affected the magnitude of the effect of these resistance features on the MIC. Thus, despite only moderate success in predicting co-amoxiclav resistance category due to phenotypic inconsistencies, a WGS-based approach could predict the MIC to within one doubling dilution (essential agreement) of the observed MIC for 97% of isolates from a population-representative set of *E. coli* BSI

Our study highlights the importance of isolate sampling frame, phenotyping method and breakpoint selection. A previous study of 76 *E. coli* isolated from cattle[17] which reported high sensitivity and specificity of WGS to predict co-amoxiclav resistance, contained highly-resistant isolates (30% containing bla_{CMY-2}), and only attempted to predict CLSI-defined resistance (>32/16 mg/L). In contrast, in our study, similar to other population representative studies of human isolates,[6,13] only a small proportion

of co-amoxiclav resistance was due to inhibitor-resistant beta-lactamases, with most of the resistance being due to hyper-production of beta-lactamases. Further, given there is lack of consensus as to which breakpoint and clavulanate concentration should be used to compare genotype with phenotype in co-amoxiclav, we assessed against both commonly used methods (EUCAST, CLSI).

Compared with other studies of BL/BLIs and *E. coli* causing human infections, less BL/BLI resistance was accounted for by inhibitor resistant beta-lactamases.[18] To identify resistance in our population-representative set of isolates, we found it critical to consider genetic features that alter expression of beta-lactamases. Although the individual effects of some of these features on MICs were small, they were important, because MICs for many isolates were around 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. Finally, 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 main study strengths are the unbiased, large, population-representative sampling frame; detailed, replicated, reference-grade phenotyping for a substantial subset of isolates; detailed and complete genotyping; and the statistical modelling. 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.

One limitation was that agar dilution phenotypes were determined for only 261 isolates, resulting in limited representation of some, rarer resistance elements. This meant firstly, that some infrequent features had to be categorized together for modelling, and secondly, we were unable to assess definitively interactions between all features (e.g. combinations of beta-lactamases). Reassuringly, however, the features causing the greatest MIC increases were those traditionally associated with co-amoxiclav resistance[7], their specific impact being modelled here for the first time. Another limitation is that many genetic features, e.g. DNA copy number, are proxies for increased expression. While we assessed the independent effects of these proxies on MIC, WGS is unable to directly quantify expression, adding uncertainty in interpretation. Likewise, feature absence/disruption in WGS data is a proxy of loss of function. Such predictions need to be interpreted with caution[30] as several genotype-phenotype discrepancies occurred in isolates with no identifiable *ampC* promoter but showing phenotypes consistent with increased *ampC* expression. Finally, we did not attempt to predict MICs in 11 isolates with resistance features not seen in our agar dilution subset as we had no reliable measure of their impact on co-amoxiclav MIC. This issue is similar in nature to the problem of predicting resistance caused by the novel $bla_{\text{CTX-M}}$ we encountered. These issues highlight the need to utilize knowledge about mechanisms of inhibitor resistance as opposed to relying solely on observed data when predicting resistance in isolates with previously unobserved resistance features.

In summary, WGS can identify the causes of co-amoxiclav resistance in *E. coli* provided the approach is extended to consider the complicated, polygenic, and

expression-related nature of this resistance. Further, our method enabled assay-specific MIC predictions from genetic data. 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. Further, 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. A genetic approach potentially offers a more reliable method to identify and monitor resistance to co-amoxiclav, as well as resistance to other BL/BLI combinations. Ultimately, given inherent uncertainties in phenotyping, future approaches need to incorporate patient outcomes.

FUNDING

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

DECLARATION OF INTERESTS

NW and MJE: none to declare. However, 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.

ACKNOWLEDGEMENTS

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

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 of genetic results and established rules regarding the relationship with phenotype. TJD and ASW fitted mixed level 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.

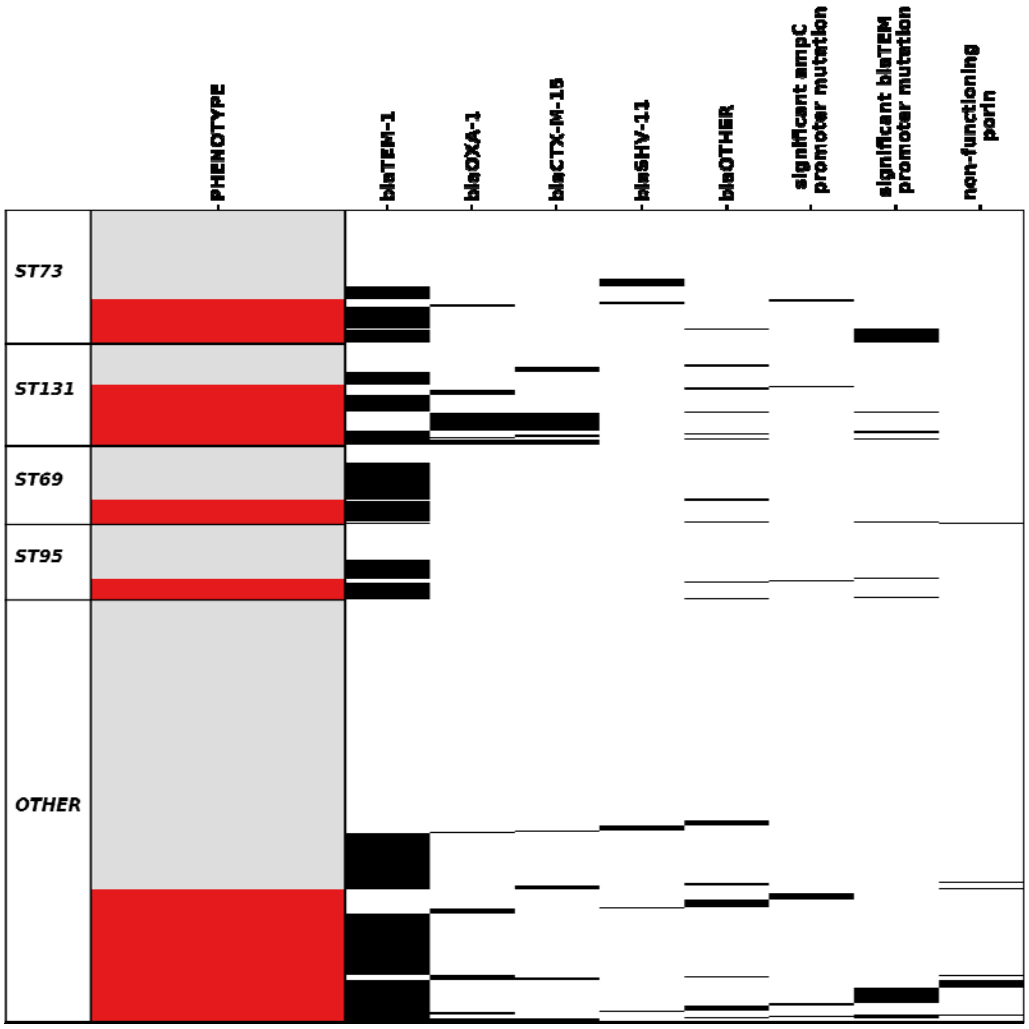
TABLES AND FIGURES

Table 1: Resistance prediction feature performance

Feature		Prediction	Number of isolates resistant by routine AST/total isolates with this feature (%)	
			PPV	**PPV restricted to isolates not predicted as co-amoxiclav resistant by any other feature**
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	21/21 (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				
6	Features suggesting disruption of either *ompC* or *ompF* in an isolate containing an additional beta-lactamase (see supplementary methods)	Extended	18/21 (86%)	2/2 (100%)

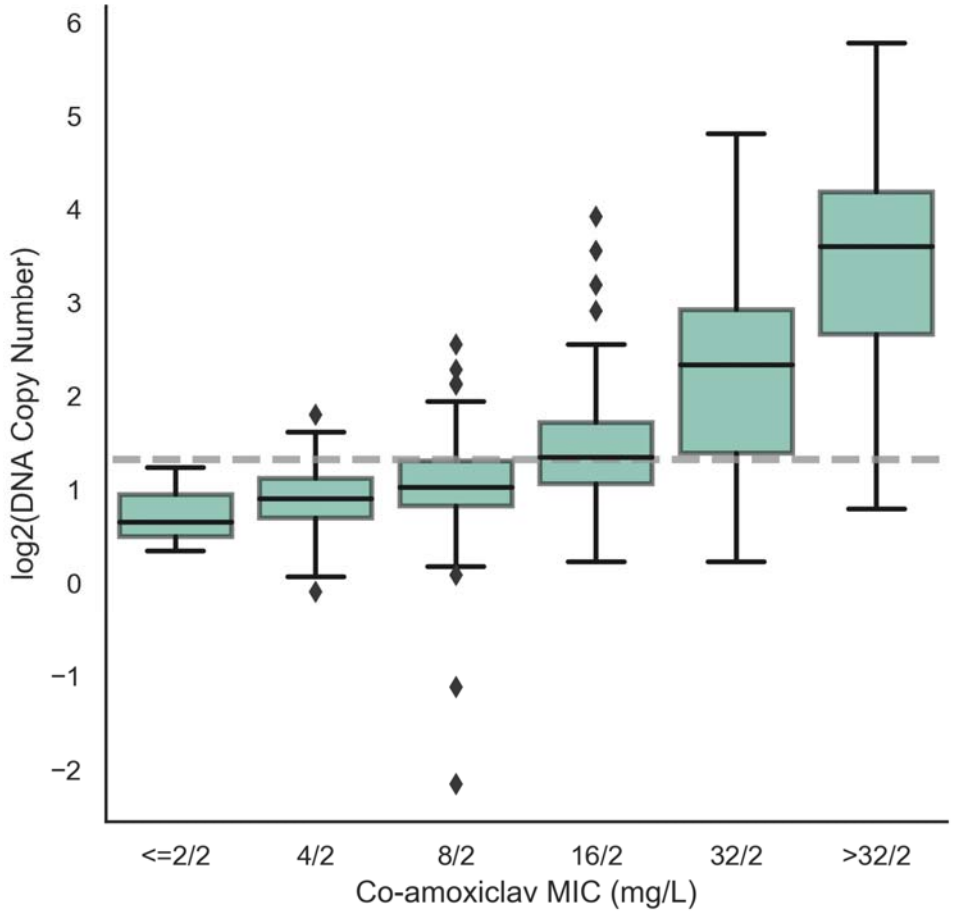
***: Cut-off chosen following a ROC analysis (Supplementary Methods)**

Figure 1: Phenotypes and genetic features found by ST (n=976)



Note: single horizontal lines represent each isolate. Red indicates resistant by BD Phoenix/EUCAST breakpoints (>8/2 mg/L) and grey susceptible. Black lines indicate the presence of each genetic feature with blaOTHER being any non $bla_{\text{TEM-1}}$, $bla_{\text{OXA-1}}$, $bla_{\text{SHV-11}}$ or $bla_{\text{CTX-M-15}}$ beta-lactamase (see Supplementary Tables).For promoter mutations/non-functional porin definitions, see Supplementary Methods.

Figure 2: Association between DNA copy number and co-amoxiclav MIC in isolates with no alternate resistance features.



Note: grey line indicates 2.5 threshold used to define resistance in the extended algorithm based on ROC analysis.

Figure 3: Proportion WGS predicted resistant (extended algorithm) by routine laboratory MIC.

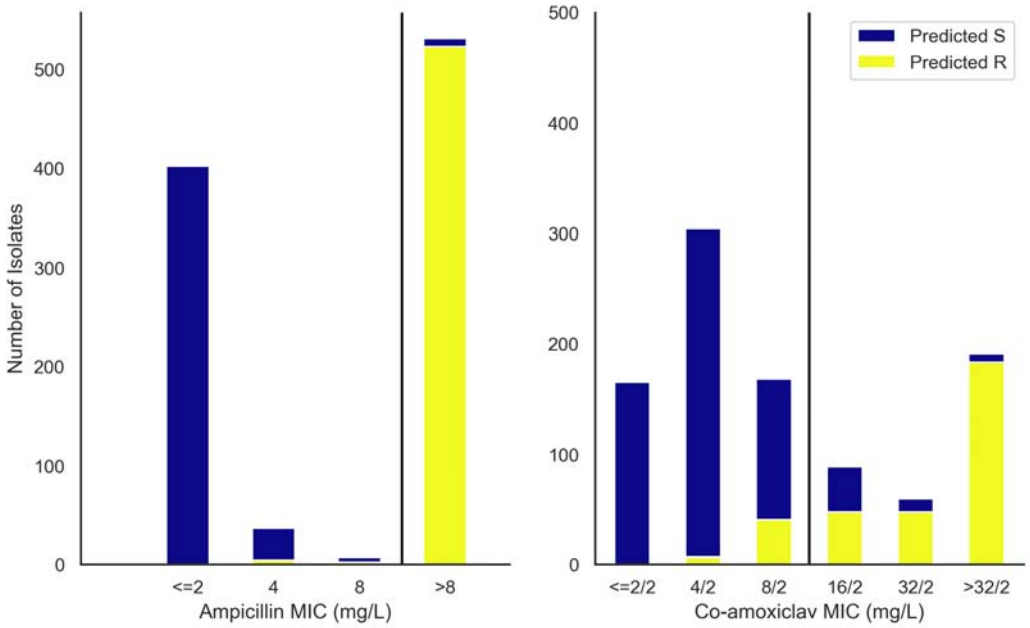
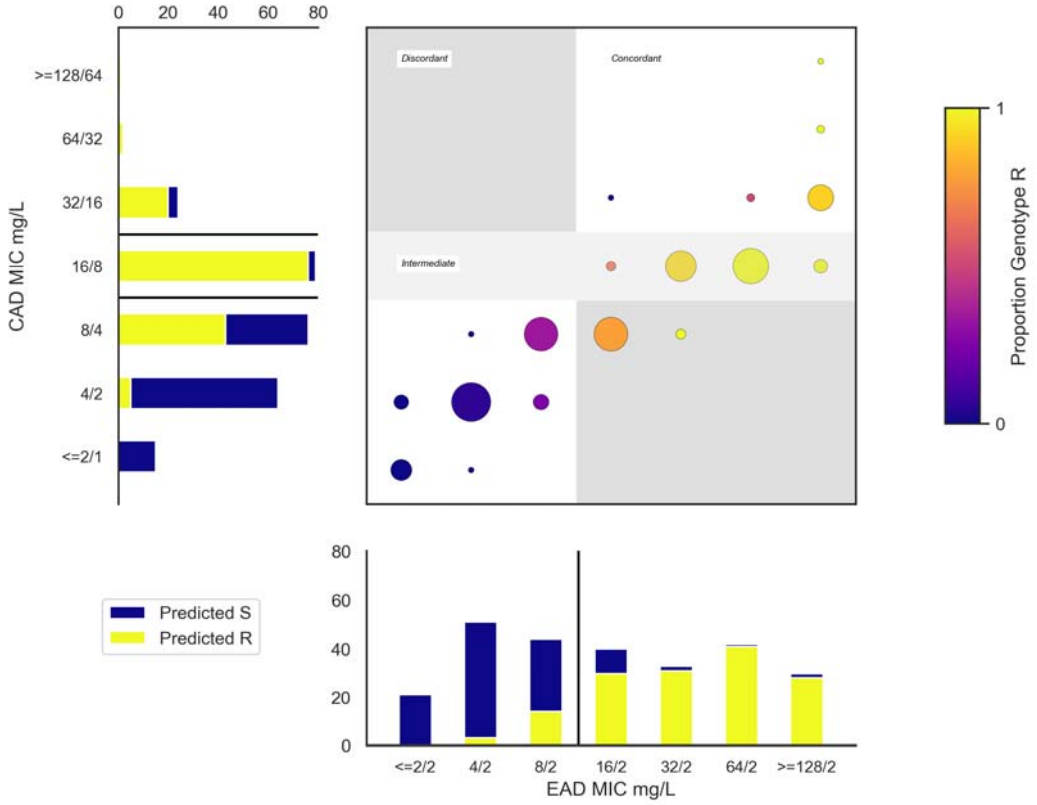
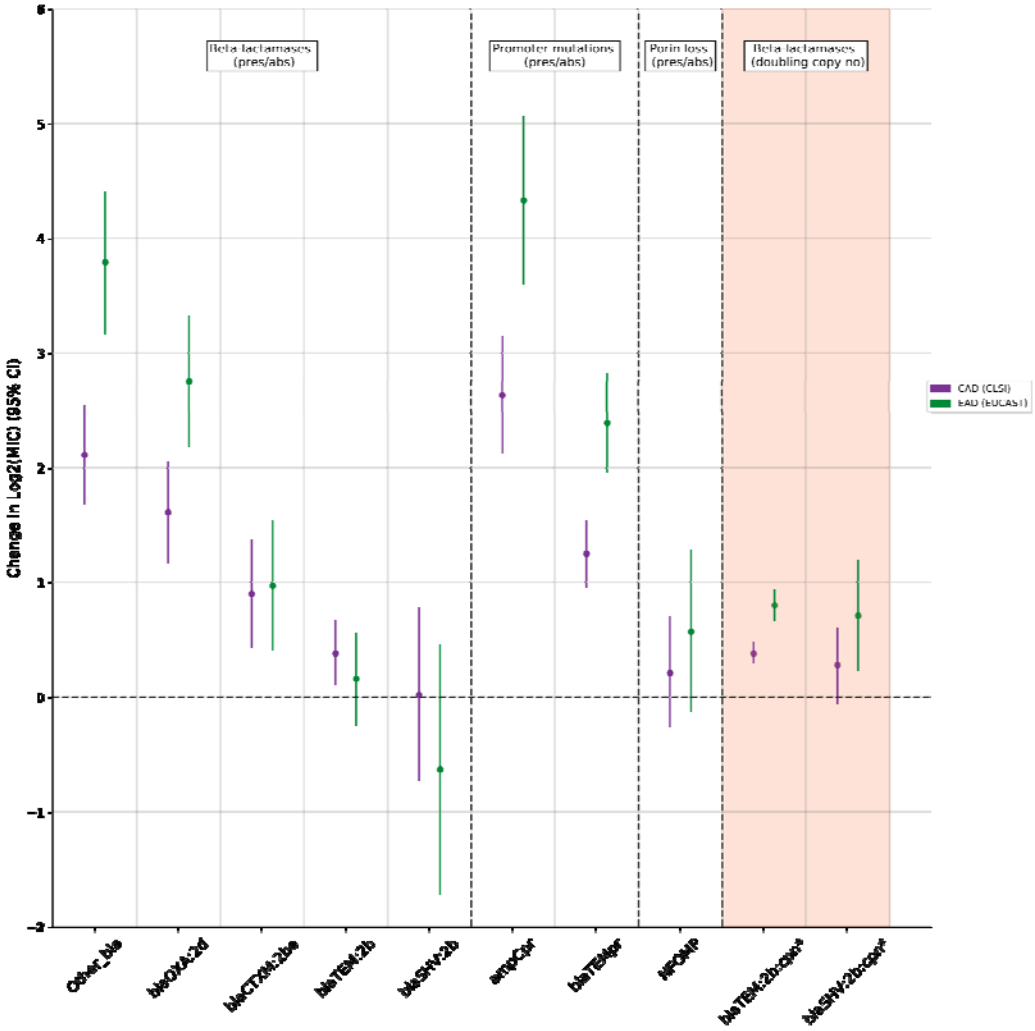


Figure 4 Proportion WGS predicted resistant (extended algorithm) by EAD and CAD MIC.



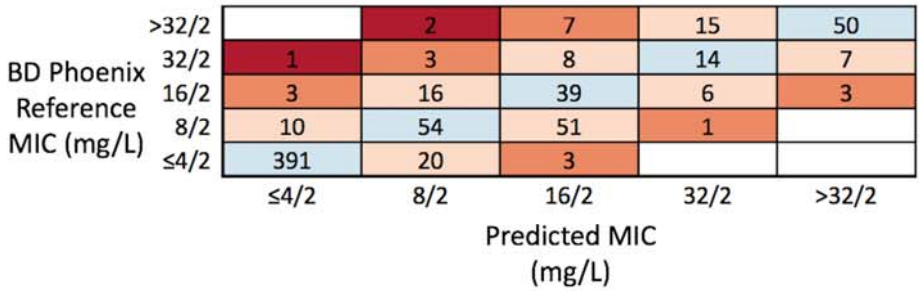
Note: Sub axes show proportion of samples predicted R by MIC for each of EAD and CAD. Main axis (x,y) coordinate represents (EAD MIC, CAD MIC), circle size represents the number of isolates with this combination of EAD and CAD MICs, and colour denotes proportion identified as resistant by WGS.

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



*Note: Purple represents testing using CAD (CLSI, amoxicillin/clavulanate 2:1 ratio)), and green EAD (EUCAST, fixed 2mg/L clavulanate). All elements except those denoted by * and shaded in orange are modelled as binary presence vs absence effects (see Supplementary Table 11).), 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=701)



Essential agreement	681/704 (97%)
Major Error	55/704 (8%)
Very Major Error	25/704 (4%)

Note: blue shading indicates correct correctly predicted observed AST MIC (548/704 (78%) isolates), light pink predicted within one doubling dilution (total 681/704 (97%) isolates, essential agreement), orange within two doubling dilutions (total 701/704 (100%)) and red greater than 2

REFERENCES

1. Public Health England. English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR) [Internet]. 2018 [cited 2018 Dec 4]. Available from: www.facebook.com/PublicHealthEngland
2. Plachouras D, Kärki T, Hansen S, et al. Antimicrobial use in European acute care hospitals: results from the second point prevalence survey (PPS) of healthcare-associated infections and antimicrobial use, 2016 to 2017. Eurosurveillance [Internet]. 2018 Nov 15 [cited 2019 Jan 2];23(46):1800393. Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.23.46.1800393
3. Diez-Aguilar M, Morosini M-I, Lopez-Cerero L, et al. Performance of EUCAST and 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. 2015 Aug;70(8):2306–10.
4. Mouton J. Antimicrobial susceptibility testing with EUCAST breakpoints and methods. In: Mouton J, editor. ECCMID. Vienna, Austria; 2017.
5. Vihta K-D, Stoesser N, Llewelyn MJ, et al. Trends over time in Escherichia coli bloodstream infections, urinary tract infections, and antibiotic susceptibilities in Oxfordshire, UK, 1998-2016: a study of electronic health records. Lancet Infect Dis. 2018 Oct;18(10):1138–49.
6. Kallonen T, Brodrick HJ, Harris SR, et al. Systematic longitudinal survey of invasive Escherichia coli in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. Genome Res. 2017 Jul;

7. Drawz SM, Bonomo RA. Three decades of beta-lactamase inhibitors. Clin Microbiol Rev. 2010 Jan;23(1):160–201.

8. Bush K. A resurgence of beta-lactamase inhibitor combinations effective against multidrug-resistant Gram-negative pathogens. Int J Antimicrob Agents. 2015 Nov;46(5):483–93.

9. EUCAST. Guidance document on beta-lactamase inhibitor combinations [Internet]. 2017 [cited 2019 Jan 2]. Available from: http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/General_documents/Inhibitor_combinations_-_Guidance_for_drug_developers_20171001.pdf

10. Humphries R. Mechanisms of resistance to ceftazidime- avibactam. In: ECCMID. Madrid, Spain; 2018.

11. Ellington MJ, Ekelund O, Aarestrup FM, et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. Clin Microbiol Infect. 2017 Jan;23(1):2–22.

12. Livermore DM, Day M, Cleary P, et al. OXA-1 β-lactamase and non-susceptibility to penicillin/β-lactamase inhibitor combinations among ESBL-producing Escherichia coli. J Antimicrob Chemother [Internet]. 2018 Nov 2;dky453-dky453. Available from: http://dx.doi.org/10.1093/jac/dky453

13. Ortega A, Oteo J, Aranzamendi-Zaldumbide M, et al. Spanish multicenter study of the epidemiology and mechanisms of amoxicillin-clavulanate resistance in Escherichia coli. Antimicrob Agents Chemother. 2012 Jul;56(7):3576–81.

14. Livermore DM. beta-Lactamases: quantity and resistance. Clin Microbiol Infect. 1997 Feb;3 Suppl 4:S10–9.

15. Stoesser N, Batty EM, Eyre DW, et al. Predicting antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. J Antimicrob Chemother. 2013 Oct;68(10):2234–44.

16. Zankari E, Hasman H, Kaas RS, et al. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J Antimicrob Chemother. 2013 Apr;68(4):771–7.

17. Tyson GH, McDermott PF, Li C, et al. WGS accurately predicts antimicrobial resistance in Escherichia coli. J Antimicrob Chemother. 2015 Oct;70(10):2763–9.

18. Shelburne SA, Kim J, Munita JM, et al. Whole-Genome Sequencing Accurately Identifies Resistance to Extended-Spectrum beta-Lactams for Major Gram-Negative Bacterial Pathogens. Clin Infect Dis. 2017 Sep;65(5):738–45.

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

20. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008 May;18(5):821–9.

21. Stoesser N, Sheppard AE, Peirano G, et al. Genomic epidemiology of global Klebsiella pneumoniae carbapenemase (KPC)-producing Escherichia coli. Sci Rep. 2017 Jul;7(1):5917.

22. Wirth T, Falush D, Lan R, et al. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol Microbiol. 2006 Jun;60(5):1136–51.

23. Hunt M, Mather AE, Sánchez-Busó L, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microb genomics. 2017

Oct;3(10):e000131.

24. Lartigue MF, Leflon-Guibout V, Poirel L, Nordmann P, Nicolas-Chanoine M-H. Promoters P3, Pa/Pb, P4, and P5 upstream from bla(TEM) genes and their relationship to beta-lactam resistance. Antimicrob Agents Chemother. 2002 Dec;46(12):4035–7.

25. Peter-Getzlaff S, Polsfuss S, Poledica M, et al. Detection of AmpC beta-lactamase in Escherichia coli: comparison of three phenotypic confirmation assays and genetic analysis. J Clin Microbiol. 2011 Aug;49(8):2924–32.

26. Caroff N, Espaze E, Bérard I, Richet H, Reynaud A. Mutations in the ampC promoter of Escherichia coli isolates resistant to oxyiminocephalosporins without extended spectrum β-lactamase production. FEMS Microbiol Lett [Internet]. 1999 Apr 1 [cited 2019 Jan 2];173(2):459–65. Available from: https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.1999.tb13539.x

27. Ambler RP. The structure of beta-lactamases. Philos Trans R Soc Lond B Biol Sci. 1980 May;289(1036):321–31.

28. Ambler RP, Coulson AF, Frère JM, et al. A standard numbering scheme for the class A beta-lactamases. Biochem J [Internet]. 1991 May 15 [cited 2019 Jan 6];276 (Pt 1)(Pt 1):269–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2039479

29. Shen Z, Ding B, Bi Y, et al. CTX-M-190, a Novel beta-Lactamase Resistant to Tazobactam and Sulbactam, Identified in an Escherichia coli Clinical Isolate. Antimicrob Agents Chemother. 2017 Jan;61(1).

30. Walker TM, Kohl TA, Omar S V, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a

retrospective cohort study. Lancet Infect Dis. 2015 Oct;15(10):1193–202.