Heyckendorf, J; Andres, S; Kser, CU; Olaru, ID; Schn, T; Sturegrd, E; Beckert, P; Schleusener, V; Kohl, TA; Hillemann, D; Moradigavand, D; Parkhill, J; Peacock, SJ; Niemann, S; Lange, C; Merker, M (2017) What is resistance? Impact of phenotypic versus molecular drug resistance testing on multi- and extensively drug-resistant tuberculosis therapy. Antimicrobial agents and chemotherapy. ISSN 0066-4804 DOI: https://doi.org/10.1128/AAC.01550-17

Downloaded from: http://researchonline.lshtm.ac.uk/4645771/

DOI: 10.1128/AAC.01550-17

Usage Guidelines

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by/2.5/
What is resistance? Impact of phenotypic versus molecular drug resistance testing on multi- and extensively drug-resistant tuberculosis therapy

Authors
Jan Heyckendorf¹⁻³*, Sönke Andres⁴*, Claudio U. Köser⁵*, Ioana D. Olaru¹⁻², Thomas Schön⁶⁻⁷, Erik Sturegård⁸, Patrick Beckert²⁻⁹, Viola Schleusener⁰, Thomas A. Kohl²⁻⁹, Doris Hillemann⁴, Danesh Moradigaravand¹⁰, Julian Parkhill¹⁰, Sharon J. Peacock⁵,¹⁰,¹¹, Stefan Niemann²⁻⁹, Christoph Lange¹⁻³,¹²,¹³, Matthias Merker²⁻⁹

Author affiliations
¹ Division of Clinical Infectious Diseases, Research Center Borstel, Borstel, Germany; ² German Center for Infection Research (DZIF), Partner site Hamburg-Lübeck-Borstel, Germany; ³ International Health/Infectious Diseases, University of Lübeck, Lübeck, Germany; ⁴ Division of Mycobacteriology (National Tuberculosis Reference Laboratory), Research Center Borstel, Borstel, Germany; ⁵ Department of Genetics, University of Cambridge, UK; ⁶ Department of Infectious Diseases and Clinical Microbiology, Kalmar County Hospital, Kalmar, Sweden; ⁷ Department of Clinical and Experimental Medicine, Division of Medical Microbiology, Linköping University, Linköping, Sweden; ⁸ Clinical Microbiology, Department of Translational Medicine, Lund University, Malmö, Sweden; ⁹ Division of Molecular and Experimental Mycobacteriology, Research Center Borstel, Borstel, Germany; ¹⁰ Wellcome Trust Sanger Institute, Hinxton, UK; ¹¹ London School of Hygiene & Tropical Medicine, London, UK; ¹² Department of Medicine, Karolinska Institute, Stockholm, Sweden; ¹³ Department of Medicine, University of Namibia School of Medicine, Windhoek, Namibia; *Contributed equally
Corresponding authors:

Prof. Dr. med. Dr. h.c. Christoph Lange, Clinical Infectious Diseases, Research Center Borstel, Parkallee 35, D-23845 Borstel, Germany. Phone +49-45371883320, email: clange@fz-borstel.de

Prof. Dr. rer nat. Stefan Niemann, Molecular and Experimental Bacteriology, Research Center Borstel, Parkallee 1, D-23845 Borstel, Germany. Phone: +49-4537188-2750, email: sniemann@fz-borstel.de

Word count: 3692

Running title: What is resistance?
Rapid and accurate drug-susceptibility testing (DST) is essential for the treatment of multi- and extensively drug-resistant tuberculosis (M/XDR-TB). We compared the utility of genotypic DST assays with phenotypic DST (pDST) using BACTEC 960 MGIT or Löwenstein-Jensen to construct M/XDR-TB treatment regimens for a cohort of 25 consecutive M/XDR-TB patients and 15 possible anti-TB drugs. Genotypic DST results from Cepheid GeneXpert MTB/RIF (Xpert) and line probe assays (LPAs: Hain GenoType MTBDRplus 2.0 and MTBDRsl 2.0) and whole genome sequencing (WGS) were translated into individual algorithm-derived treatment regimens for each patient. We further analysed if discrepancies between the various methods were due to flaws in the genotypic or phenotypic test using MIC results. Compared with pDST, the average agreement in the number of drugs prescribed in ‘genotypic’ regimens ranged from just 49% (95% CI 39-59%) for Xpert and 63% (95% CI 56-70%) for LPAs to 93% (95% CI 88-98%) for WGS. Only the WGS regimens did not comprise any drugs to which pDST showed resistance. Importantly, MIC testing revealed that pDST likely underestimated the true rate of resistance for key drugs (rifampicin, levofloxacin, moxifloxacin, and kanamycin) because critical concentrations (CCs) were too high. WGS can be used to rule-in resistance even in M/XDR strains with complex resistance patterns, but pDST for some drugs is still needed to confirm susceptibility and construct the final regimens. Some CCs for pDST need to be re-examined to avoid systematic false-susceptible results in low-level resistant isolates.
INTRODUCTION

Tuberculosis (TB) is a leading cause of morbidity and mortality worldwide (1). Although the global incidence of TB has been slowly declining, the emergence of multidrug-resistant (MDR)-TB, defined as resistance to rifampicin and isoniazid, challenges TB-control (1). Extensively drug-resistant (XDR)-TB, defined as MDR-TB and resistance to at least one fluoroquinolone [e.g. ofloxacin, levofloxacin, or moxifloxacin; World Health Organization (WHO) group A] and any second-line injectable drug (SLID, amikacin, kanamycin, or capreomycin; WHO group B) has been reported in 117 countries (1).

Therapy of M/XDR-TB is complex and requires a long duration of treatment with a combination of at least four drugs often leading to adverse-events and poor treatment outcomes (2, 3). Moreover, the initiation of appropriate therapy is often delayed due to the slow growth rate of Mycobacterium tuberculosis complex isolates, which means that phenotypic drug-susceptibility testing (pDST) can take weeks to months (4, 5). To accelerate this rate-limiting step, a number of genotypic DST assays that detect resistance mutations have been endorsed by the WHO (6). The Cepheid GeneXpert (Xpert) is an automated point-of-care assay with a high diagnostic accuracy for rifampicin-resistance detection, providing results within 1.5 hours (7). Line probe assays (LPAs, e.g. Hain GenoType MTBDRplus 2.0 and MTBDRsl 2.0) can also be performed directly from sputum to provide results within 1-2 days with a high diagnostic accuracy for resistance to isoniazid, rifampicin, fluoroquinolones, and SLIDs (6). Because these assays only target a limited number of resistance variants, their sensitivity compared with pDST is limited. Whole genome sequencing (WGS) can theoretically overcome this shortcoming by interrogating the entire genetic repertoire (4, 5, 8). Nevertheless, the utility of WGS is currently limited by the need for expensive equipment, highly trained personnel, and complex bioinformatic procedures. Moreover, WGS requires an initial culture, which introduces a delay compared with the aforementioned targeted assays (6,
9). More fundamentally, there is a lack of understanding of the genetic basis of antibiotic resistance, which complicates the interpretation of WGS data (10).

However, it is important to appreciate that discrepancies observed between pDST and genotypic methods are not exclusively due to problems related to the interpretation of the genotype (6). Instead, the evidence is mounting that some critical concentrations (CCs), which are set by the Clinical and Laboratory Standards Institute (CLSI) and/or WHO and define resistance on a phenotypic level, are higher than the epidemiological cut-off values (ECOFFs), which represent the highest concentration of the wild-type MIC distribution (6, 11-15). As a result, some isolates with elevated MICs compared to the ECOFF due to known mutations are classified as susceptible even though limited pharmacokinetic/pharmacodynamics or clinical outcome data evidence exists that these isolates are still treatable (6, 12, 13, 16).

Therefore, this study had two main goals. First, we compared the utility of genotypic methods (Xpert, LPAs, and WGS) with pDST to design M/XDR regimens using standardised algorithms. Second, we analysed whether discrepancies between the various methods were due to flaws in pDST or the genotype.
RESULTS

Patient cohort

20 patients with MDR-TB and 5 with XDR-TB admitted to the Medical Clinic of the Research Center Borstel (Germany) were enrolled (Table S1).

Comparison of M/XDR TB regimens based on pDST with molecular methods

367 pDST results for a total of 15 drugs served as the reference standard (Figure 1). Xpert classified all 25 patients as having rifampicin resistance, yet one isolate was phenotypically susceptible, resulting in an agreement of 96% (95% CI 80-100%). LPA and pDST results agreed in 228 of 243 cases [94% (95% CI 90-97%)]. 340 of the 367 WGS-based drug resistance predictions [93% (95% CI 89-95%)] were concordant with pDST (Figure 1A, Table S2).

There was a 49% (95% CI 39-59%) average agreement in number of antibiotics prescribed between the regimens based on Xpert results alone and those based on pDST (Figure 2 and Table S3) (3). This increased to 68% (95% CI 56-80%), if resistance to both ethambutol and pyrazinamide was also assumed based on the discovery of rifampicin resistance. Making the equivalent assumption for LPAs increased the agreement from 63% (95% CI 56-70%) to 87% (95% CI 80-94%). The best agreement with pDST regimens was achieved with WGS [93% (95% CI 88-98%)] (Figure 2 and Table S3). Importantly, the WGS regimens did not feature any drugs to which resistance was found using pDST. In contrast, the 25 regimens that were designed using LPAs or the Xpert contained 56/152 [37% (95% CI 29-56)] and 77/150 [51% (95% CI 43-60%)] drugs respectively, for which pDST showed resistance (Table S4).

A more detailed analysis of drug categories revealed that the Xpert regimens involved an increased administration of group A, B, and D1 drugs compared with pDST (P<0.001) (Table
Moreover, no D2 and D3 drugs were part of these regimens ($P<0.001$). For the LPA regimens, only the increase in the number of D1 drugs was statistically significant. By contrast, the use of WGS resulted in a significant decrease in the use of D1 drugs because more ethambutol resistance was predicted (Table S5).

**Analysis of the discrepancies between different DST methods**

We determined the MICs for selected isolates and antibiotics to investigate the potential causes of the discrepancies observed with the different DST methods (Table S2).

**Rifampicin and rifabutin**

One isolate (11102-14) with an $rpoB$ D435Y mutation had an MIC for rifampicin that was below the CC, but above the tentative ECOFF defined in this study (tentative ECOFF=0.25 µg/ml < $rpoB$ mutant=0.5 µg/ml < CC=1 µg/ml), which suggested that the susceptible pDST result likely represented a breakpoint artefact (Figure 3A). This isolate also tested susceptible to rifabutin at the CC of 0.5 µg/ml (Figure 3B). In this case, however, the result was likely valid as its MIC (0.06 µg/ml) was even lower than the tentative ECOFF (0.12 µg/ml). By contrast, the susceptible pDST results to rifabutin for the D435Y and L452P/E481A isolates (12041-13 and 999-13) were again likely the result of a breakpoint artefacts (17).

**Isoniazid and prothionamide**

All gWT isolates tested susceptible at the CLSI and WHO CC of 0.1 µg/ml. Conversely, all isolates with elevated MICs had known resistance mutations. Although not endorsed by WHO and not considered for our hypothetical regimens, CLSI has set 0.4 µg/ml as an additional breakpoint to define low-level resistance that can be treated with a high dose of isoniazid according to some recommendations (Figure 3C) (18). Based on our WGS results, we were able to predict that all gNWT isolates were resistant even at this higher concentration [either
because of the katG S315T mutation, which is known to confer predominantly high-level resistance, or because the isolates harboured both the inhA -15c/t promoter mutation and inhA coding changes (S94A or I194T) (18, 19). It was not possible to predict the correct level of resistance for the inhA double mutants using the MTBDRplus given that this assay only interrogates promoter mutations (20).

For prothionamide, we only observed a single disagreement between our WGS predictions and pDST (21). Isolate 3758-14 originally tested susceptible despite a frameshift mutation in ethA (22). However, this discrepancy was likely a random error since the isolate was found to have an elevated MIC compared with the CC (>25 μg/ml vs. 2.5 μg/ml, respectively).

Levofloxacin and moxifloxacin

All seven isolates with known gyrA resistance mutations were resistant to levofloxacin at the CC of 1.5 μg/ml (23). However, a review of MIC data from the literature revealed a tentative ECOFF of 0.75 μg/ml, which resulted in the misclassification of 9 gyrA isolates from the literature (Figure 4A).

WHO has set two CCs for moxifloxacin. The lower CC at 0.5 μg/ml is supposed to correspond to the ECOFF and is intended as a surrogate for ofloxacin and levofloxacin resistance (14, 24). However, our pooled MIC data suggested that the tentative ECOFF was actually 0.25 μg/ml, which was in agreement with the current CLSI guidelines (Figure 4B) (11). All of our gyrA mutants were resistant at 2 μg/ml, the second WHO CC, which should define resistance to moxifloxacin itself (i.e. isolates with only slightly elevated MICs of 1 and 2 μg/ml are deemed to still be treatable with moxifloxacin). However, in light of the fact that WHO has already acknowledged that this CC may be too high and given that predicting the
precise MIC based on genotypic data alone is challenging, we simply classified our isolates as
gNWT (24).
SLIDs
The MIC distribution for isolates with known mutations in the resistance genes eis and whiB7
ranged from 2.5 to 10-12.5 μg/ml and was truncated by the current CC of 2.5 μg/ml, whereas
all gWT isolates had MICs ≤0.125 μg/ml (25-27). Therefore, the two isolates with an MIC of
2.5 μg/ml (12471-13 and 11411-14) would have tested resistant if the CC was lowered to the
tentative ECOFF of 1.25 μg/ml (Figure 5A and Table S2). Moreover, we would predict isolate
811-15, which had a known whiB7 resistance mutation (-56 g/a), to retest resistant at 1.25
μg/ml (it tested susceptible at 2.5 μg/ml and no MIC data were available for this isolate) (26).
Two isolates had a previously unknown deletion of the upstream and coding region of eis,
which resulted in an invalid result with the MTBDRsl assay. The effect of this change on
kanamycin resistance remains to be determined.
No discrepancies were observed for amikacin and capreomycin (28).
Other antibiotics
No discrepancies were found for streptomycin and pyrazinamide (29-33). For linezolid,
isolate 9685-14 had a novel 23S mutation (rrl 906 g/a) that was observed in a susceptible
isolate.
For the remaining antibiotics, we found evidence of false-susceptible pDST results. In the
case of ethambutol, all 25 isolates were classified as gNWT but four tested susceptible (34-
36). Up to five isolates, as opposed to two just phenotypically confirmed isolates, might have
been cycloserine resistant given that the recently proposed tentative ECOFF of 20 μg/ml is
below the CC of 30 µg/ml (37). Finally, up to six additional isolates could have been resistant to para-aminosalicylic acid based on the WGS data (see supplementary results).
DISCUSSION

We investigated how different genotypic DST assays influence the design of standardised algorithm-derived M/XDR-TB regimens. As expected, the accuracy of predicting resistance and, consequently, the ability to design appropriate treatment regimen correlated with the proportion of the genome analysed. Moreover, we demonstrated that the pDST results were flawed in some cases.

Although LPAs have been endorsed by the WHO for the rapid molecular prediction of drug-resistance of rifampicin, isoniazid, fluoroquinolones, and SLIDs, the Xpert is the most frequently used assay for initial routine molecular DST in many high-burden countries (6). Based on our results, it is a good test to rule-in rifampicin resistant TB that can be used as surrogate marker for M/XDR-TB depending on the geographical region. However, it is paramount that these results are complemented with additional DST since a treatment regimens based only on an Xpert result would have led to the ineffective administration of approximately half of the drugs in this cohort of patients who were predominantly from Eastern Europe. This will be different in other geographic settings, where the extent of drug resistance beyond rifampicin and isoniazid is lower (38, 39).

The prediction of resistance to fluoroquinolones and SLIDs by LPAs was generally accurate for patients in this cohort. However, this test was also insufficient to construct appropriate M/XDR-TB regimens compared with pDST, especially in patients with XDR-TB. For example, almost all of the patients with M/XDR-TB from this cohort had strains that were resistant to ethambutol and pyrazinamide, which are not covered by the MTBDRs/ 2.0. This was in line with results from a European study at 26 different centres in high-intermediate- and low-burden countries of TB that reported resistance to pyrazinamide and ethambutol in...
59.7% and 59.3% of all patients with MDR-TB (94.4% and 81.8% of patients with XDR-TB), respectively (38, 39).

The M/XDR-TB treatment regimens based on WGS showed the highest agreement [93% (95% CI 88-98%)] with those based on pDST. Unlike the other genotypic assays, WGS did not miss any phenotypically confirmed resistances, but did predict resistance in some phenotypically susceptible isolates. This was partly due to the fact that we identified novel or poorly defined mutations that we could not interpret with regard to their impact on resistance development (e.g. mutations in \textit{rrl} or \textit{gyrB}; Table S2). Here, we adopted a conservative approach and assumed that these mutations conferred resistance, until disproved by another method, e.g. MIC determination of mutants derived from allelic exchange experiments and sequential patient derived isolates that allow the interpretation of individual mutations and their effect on the drug resistance level in a particular phylogenetic strain background.

In other cases, problems with pDST played a role. The false-susceptible pDST results for ethambutol were likely due to the fact that some resistance mutations only result in slight MIC increases, which means that it can be difficult to distinguish the gWT strains from gNWT strains using pDST, unless secondary mutations increase the MICs even further (14, 40-42). The lack of reproducibility of pDST was also apparent for isolate 3758-14, which initially tested susceptible to prothionamide but became resistant upon retesting (Table S2).

Our results highlighted breakpoint artefacts (i.e. cases in which the current CCs were likely set above the tentative ECOFFs) as a major cause for systematic errors. In the absence of well-documented, high-quality evidence that isolates with elevated MICs can be treated with the standard or an elevated dose, the CCs for these drugs should be lowered to the tentative ECOFFs to avoid misdiagnosing isolates with elevated MICs as susceptible (12, 13). One
possibility to gather such evidence would be to conduct a placebo-controlled study in which
high-dose rifampicin or rifabutin is used to treat low-level rpoB resistance mutations as part
of a backbone M/XDR-TB regimen (43).

Importantly, we raised the possibility that breakpoint artefacts may exist for six drugs that
constitute the backbone of the treatment of drug-susceptible TB or MDR TB (i.e. rifampicin,
levofloxacin, moxifloxacin, and kanamycin) in addition to less widely used drugs (i.e.
rifabutin and cycloserine). The impact of this phenomenon depends on the geographic setting.

For example, low-level resistance mutations in rpoB account for more than 10% of rifampicin
resistance in Bangladesh, but are less frequent in other countries (44, 45). Problems related to
kanamycin DST are likely to be important in Eastern Europe where eis mutations are
widespread amongst the dominant MDR TB clones (46, 47).

This study was limited given that it was retrospective and only featured a small number of
MDR and XDR patients from a single centre although the comparison between genotypic
DST and pDST was strengthened by inclusion of MIC determinations of fully susceptible
isolates from Sweden (n=15). Our results did not provide direct evidence that treatment
regimen based on different genotypic DST methods have an impact on clinical outcomes.

Moreover, data from more laboratories including both drug resistant and drug susceptible
isolates are required to set ECOFFs with confidence (16, 48). Nevertheless, the fact that
potential breakpoint artefacts were found for so many key drugs underlines the urgent need
for both CLSI and WHO to re-examine their CCs, which were largely set based on expert
opinion using evidence that was not or insufficiently documented, as opposed to modern and
transparent principles pioneered by the European Committee on Antimicrobial Susceptibility
Testing (EUCAST) (6, 12, 16). Importantly, this should include clear recommendations about
how to proceed when discrepant results between genotypic assays and pDST are found (49).

Ideally, these recommendations should consider MICs as well as clinical outcome data.

In conclusion, the strength of this study was that instead of merely calculating the concordance of genotypic DST results compared with pDST, as is customary for these assessments, we also compared the resulting regimens. In our view, this is more clinically meaningful as TB is never treated with a single drug (in effect, we assessed the situation in settings that lack the laboratory infrastructure for pDST or, alternatively, the period whilst pDST is being carried out but its results are not yet available). This is an important distinction since the concordance of a genotypic DST assay with pDST can be deceptively high [96% (95% CI 80-100%) for Xpert in our case], yet more than half of the drugs in the resulting regimens would still be prescribed inappropriately. Xpert and LPA results should therefore only be used to rule-in resistance to WHO group A/B drugs and need to be complemented with further testing. WGS can provide important additional information on resistance to WHO group C/D drugs but cannot replace pDST completely either (e.g. pDST is still needed for novel mutations and to detect resistance caused by known resistance mutations that occur at frequencies below the detection limit of WGS (6)). Finally, the CCs need to be re-evaluated to avoid systematic false susceptible pDST results for a variety of first and second line drugs.
MATERIALS AND METHODS

Study population

All patients (n=25) with a diagnosis of M/XDR-TB admitted to the Medical Clinic of the Research Center Borstel (Germany) between March 2013 and March 2015 were included consecutively in the study.

Microbiology, pDST and MIC testing

The primary detection, enrichment, DST, and MIC testing for the Germany isolates were done under routine conditions at the German National Reference Laboratory for Mycobacteria, Borstel. The following CCs in μg/ml were used for pDST with the BACTEC 960 MGIT system using a critical proportion of 1% for all drugs, with the exception of pyrazinamide, for which 10% was employed: rifampicin (1.0), rifabutin (0.5), isoniazid (0.1), prothionamide (2.5), ofloxacin (2.0), levofloxacin (1.5), moxifloxacin (0.5 & 2.0), kanamycin (2.5), amikacin (1.0), capreomycin (2.5), para-aminosalicylic acid (4.0), streptomycin (1.0), ethambutol (5.0), pyrazinamide (100.0), and linezolid (1.0) (11, 14). Cycloserine was tested using the proportion method on Löwenstein-Jensen medium using a CC of 30 μg/ml and a critical proportion of 1% (14).

The following concentrations in μg/ml were included for MGIT MIC testing for clinical isolates: rifampicin (0.12, 0.25, 0.5, 1.0, 4.0, 20.0), rifabutin (0.06, 0.12, 0.25, 0.5, 2.0, 10.0), isoniazid (0.1, 0.4, 1.0, 3.0, 10.0), prothionamide (0.62, 1.25, 2.5, 5.0, 10.0, 25.0), levofloxacin (0.18, 0.37, 0.75, 1.5), moxifloxacin (0.06, 0.12, 0.25, 0.5), kanamycin (0.31, 0.62, 1.25, 2.5, 5.0, 12.5, 25.0), amikacin (0.12, 0.25, 0.5, 1.0, 4.0, 20.0, 40.0), capreomycin (0.31, 0.62, 1.25, 2.5, 5.0, 12.5, 25.0), and para-aminosalicylic acid (0.5, 1.0, 2.0, 4.0). The
following concentrations ranges in μg/ml were tested in two-fold dilutions for the *M. tuberculosis* H37Rv ATCC 27294 reference strain: rifampicin (0.06-0.5), rifabutin (0.06-0.5), isoniazid (0.006-0.05), prothionamide (0.31-2.5), levofloxacin (0.09-1.5), moxifloxacin (0.06-0.5), kanamycin (0.31-2.5), amikacin (0.12-1), capreomycin (0.31-2.5), *para*-aminosalicylic acid (0.5-4), and linezolid (0.12-1).

**Molecular DSTs**

All baseline sputum specimens were analysed with the Xpert assay according to the recommendation of the manufacturer. Genomic DNA extracted with cetyltrimethylammonium bromide from Löwenstein-Jensen cultures was used for the MTBDRplus 2.0 and MTBDRsl 2.0 LPAs as well as for WGS using a modified Illumina NexteraXT protocol and the MiSeq or NextSeq sequencers (20, 50-52). The detection of a *inhA* promotor variant with the MTBDRplus was used to infer prothionamide resistance (18). The raw data (fastq files) was submitted to the European Nucleotide Archive (Table S2). Resulting reads were aligned to the *M. tuberculosis* H37Rv genome (GenBank ID: NC_000962.3) using BWA-MEM (53). The GATK software package was utilized for base quality re-calibration and alignment correction for possible PCR or insertion/deletion artefacts (54). Polymorphisms with a minimum of 10x coverage and 75% variant frequency were extracted and combined for all isolates using customized perl scripts. We focused our analysis on 33 resistance genes (Table S6), for which known polymorphisms that do not correlate with resistance (i.e. phylogenetic variants) were excluded (Table S7) (5, 55, 56).

WGS data were analysed as follows (15). Isolates that did not have any mutations or only harboured neutral polymorphisms in drug-resistance genes (Table S7) were classified as genotypically wild-type and were assumed to be susceptible (gWT-S). Isolates with mutations...
known to result in MICs above the current CC that defines resistance [i.e. MICs > CC(R)]
were classified as genotypically non-wild-type and resistant (gNWT-R). Where two CCs have
been set to define intermediate resistance (i.e. isolates that are treatable with an elevated dose
of the drug), isolates with mutations that result in MICs within this range [i.e. CC(S) < MIC ≤
CC(R)] were gNWT intermediate (gNWT-I). gNWT susceptible (gNWT-S) was used to refer
to isolates with mutations that confer elevated MICs below the lowest CC [i.e. ECOFF < MIC ≤
CC(S)]. Isolates with likely or known resistance mutations that do not necessarily result in
MICs above the CC(S/R) (i.e. in the case of ethambutol and kanamycin) or that confer MIC
increases above the CC(S) but not necessarily above the CC(R) were classified as simply
gNWT. Mutations with no or insufficient evidence with regards to their effect on MICs were
classified as ‘unclear’.

Algorithm-derived treatment regimens

We retrospectively designed treatment regimens based on the results obtained from each DST
method (pDST, Xpert, LPAs, and WGS) using current MDR-TB treatment recommendations,
as outlined in the supplementary methods (3). To err on the side of caution, unclear and
gNWT mutations from WGS were considered to be resistant. The 367 initial pDST results
served as reference standard for all comparisons (15 drugs for 25 patients with eight missing
results, which could not be conducted because of biosafety concerns).

Statistics

Concordance between each diagnostic test result with phenotypic DST was scored for every
individual on a scale from 0 to 1 with 0 representing no concordance and 1 perfect
concordance for each individual test result. The same approach was used to assess the overlap
between the different treatment regimens for each individual regimen. Differences in scores
were evaluated using the Mann Whitney U test. The overlap between different diagnostic
methods and the agreement between the different treatment regimens were evaluated using
the differences in proportions where each drug from a given group was considered
independently. Graphs were created and statistics calculated using STATA version 14
(STATA Corp., Texas, USA) and Prism Version 5 (Pad Software Inc., La Jolla, California,
USA). P-values below 0.05 were considered as significant.
Determining tentative ECOFFs

We set tentative ECOFFs by visual inspection for a variety of antibiotics (statistical methods could not be used given the MIC data did not meet the minimum requirements specified by EUCAST to set ECOFFs (48)). For this purpose, we pooled the MICs from the German patient cohort with MICs from a Swedish collection (see supplementary methods) and the literature, wherever the individual concentrations and concentration ranges were sufficiently similar (17, 19, 27, 57, 58). As shown in Table S8, we had to truncate some of the distributions for this purpose. For Kambli et al. we excluded one isolate, for which the genetic basis of the elevated MICs was not clear (27). We did not display the MICs for gyrB mutations from Nosova et al. given the mutations differed from the gyrB A504V mutation observed in our study (57). We only included MIC data for rpoB mutations from Berrada et al. that also occurred in the German isolates (17).

Ethics

The ethics committee of the University of Lübeck, Germany approved the study (#15-195A). Approval for whole genome sequencing and analysis of the isolates from Sweden was granted by the UK National Research Ethics Service (12/EE/0439) and the Cambridge University Hospitals NHS Foundation Trust R&D Department (A092685).
Funding

German Ministry of Education and Research (BMBF) for the German Center of Infection Research (DZIF) and Health Innovation Challenge Fund (HICF-T5-342 and WT098600), a parallel funding partnership between the UK Department of Health and Wellcome Trust. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health, Public Health England, or the Wellcome Trust.

Transparency declarations

JP, SJP, and CUK have collaborated with Illumina Inc. on a number of scientific projects. JP has received funding for travel and accommodation from Pacific Biosciences Inc. and Illumina Inc. SJP has received funding for travel and accommodation from Illumina Inc. CUK, SN and CL are consultants for the Foundation for Innovative New Diagnostics. The Bill & Melinda Gates Foundation, PerkinElmer, and Janssen Pharmaceutica covered CUK’s travel and accommodation to present at meetings. The European Society of Mycobacteriology awarded CUK and MM the Gertrud Meissner Award, which is sponsored by Hain Lifescience.
References


Figure 1: Comparison of pDST, Xpert, LPA, and WGS results and corresponding regimens

Upper panels: Results for pDST and molecular methods (Xpert, LPAs, and WGS) for 25 *M. tuberculosis* isolates from patients with M/XDR-TB. Test results denoting either confirmed phenotypic susceptibility or assumed susceptibility based on genotypic methods are shown in green, those denoting resistance are in red, gNWT variants with elevated MICs are in orange, whereas mutations with unclear effects are in grey. Differences between Xpert, LPA, or WGS results compared to the pDST are outlined by black margins (both gNWT and unclear variants were assumed to be resistant for the purposes of designing the regimens and results between DST methods).

Lower Panels: Standard algorithm-derived treatment regimens based on respective results of pDST, LPAs, WGS, and Xpert. Differences of resulting therapy regimens in comparison to the pDST-derived treatments are highlighted by black boxes. Vertical bars indicate data for 15 drugs for each patient, i.e. from left to right isoniazid (H), rifampicin (R), rifabutin (Rb), ethambutol (E), pyrazinamide (Z), kanamycin (Km), amikacin (Am), capreomycin (Cm), ofloxacin (Ox), moxifloxacin (Mx), levofloxacin (Lx), prothionamide (Pt), para-aminosalicylic acid (Pa), cycloserine (Cs), terizidone (Tz), amoxicillin/clavulanic acid (Ac), Meprenem (Me), clofazimine (Cf), delamanid (De), bedaquiline (Bq)

Figure 2: Average overlap of different regimens based on molecular DST assays compared with pDST results.

Standard algorithm-derived treatment regimens based on results of Xpert, LPAs, and WGS (X-axis) with their mean overlap to standard algorithm-derived treatment regimens based on pDST results (Y-axis). Mean overlaps (dots) are expressed with 95% confidence intervals
P values assessing the differences between the mean overlaps between the treatment regimens are shown above.
A+B) The CCs for rifampicin and rifabutin were two dilutions higher than the tentative ECOFFs defined based on the pooled MIC data from this study and the literature (i.e. 1 vs. 0.25 μg/ml for rifampicin and 0.5 vs. 0.12 μg/ml for rifabutin) (17). These distinctions did not make a difference for isolates with rpoB S450F or S450L mutations, which resulted in large MIC increases for both drugs. By contrast, the susceptible resistance result to rifampicin by pDST for the rpoB D435Y isolate (11102-14), as well as the rifabutin results for the rpoB D435V and L452P/E481A isolates (12041-13 and 999-13) likely were breakpoints artefacts, as the isolates had elevated MIC levels compared with gWT isolates and the H37Rv laboratory strain. By contrast, the rpoB D435Y isolate appeared to be genuinely susceptible to rifabutin. However, lowering the CCs for both drugs to the ECOFFs would not necessarily ensure that isolates with elevated MICs always test resistant phenotypically. For example, because the MIC distribution of rpoB D435V (0.12-0.5 μg/ml) overlapped with the gWT distribution of rifabutin, the normal variation in MIC testing would result in a poor reproducibility of pDST for this mutation.

C) WHO has only endorsed a single critical concentration for isoniazid, whereas CLSI has set an additional breakpoint that defines high-level resistance. Some treatment guidelines recommend the treatment of low-level resistant strains with a high dose of isoniazid (18). All mutant isolates were found to be resistant even at the second CLSI breakpoint, which was in accordance with our prediction based on WGS data (18). This would not have been apparent using the GenoType MTBDRplus assay given that it only interrogates inhA promoter mutations, which typically result in low MICs, although this did not affect our interpretation of the assay since we only relied on the WHO CC (18).
Figure 4: MIC distributions for levofloxacin and moxifloxacin

The pooled MIC data identified potential breakpoint artefacts for both agents. First, the CLSI and WHO critical concentrations for levofloxacin were one dilution higher than the tentative ECOFF defined in this study (1.5 vs 0.75 μg/ml) (11, 14). Second, the pooled data supported the current CLSI critical concentration (0.25 μg/ml) as the tentative ECOFF for moxifloxacin rather than the value set by WHO (0.5 μg/ml), which is designed as a surrogate for testing resistance to ofloxacin and levofloxacin (24). Moreover, WHO has acknowledged that the critical concentration at 2 μg/ml that defines resistance to moxifloxacin may be too high (24).

Because two isolates with different genetic backgrounds shared the same gyrB A504V mutations, which is typically a signal of positive selection, these isolates were categorized as unclear. However, MIC testing revealed MICs that were equal or below even the tentative ECOFFs for both fluoroquinolones, which was in line with allelic exchange experiments (59).

Figure 5: MIC distributions for kanamycin, amikacin and capreomycin

The direct alteration of rrs, the shared target of kanamycin, amikacin, and capreomycin, via the A1401G mutation is known to confer unequivocal cross-resistance to all three drugs, which was in agreement with the pooled MIC data (60). By contrast, the current CCs for kanamycin was found to truncate the MIC distribution for isolates with eis and whiB7 mutations (27). This meant that isolates with an MIC of 2.5 μg/ml were misclassified as susceptible despite the fact these included mutations that had been shown to result in elevated MICs using allelic exchange experiments (i.e. eis -37 g/t, eis -10 g/a and whiB7 -116 a/g) (25, 26). By contrast, neither eis nor whiB7 mutations had a significant impact on the MICs of amikacin or capreomycin (based on previous data, the fact that the tentative ECOFF for capreomycin for our study was below the critical concentration was likely an artefact due to the small number of gWT isolates included in this study) (61).
<table>
<thead>
<tr>
<th>ID</th>
<th>Xpert</th>
<th>LPA</th>
<th>WGS</th>
<th>pDOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Legend:
- **Green**: Susceptible test result (either confirmed by pDOT or assumed by molecular method)
- **Red**: Resistant test result
- **Yellow**: Confirmed resistant based on WGS, i.e., isolate with mutation that results in elevated MIC, but for which the result could not be predicted
- **Pink**: Susceptible or borderline test result confirmed by WGS, i.e., isolate with mutation that results in elevated MIC, but for which the result could not be predicted
- **Black**: Differences between pDOT and molecular assay
- **Blue**: Drug administered
- **Orange**: Drug not administered

### Notes:
- The data presented in the table and diagram indicates the results of various tests (Xpert, LPA, WGS, pDOT) for different drug samples (ID 1-20). The color coding helps in identifying the susceptibility status of the isolates.
- The diagram illustrates the distribution of results across different drug categories and the correspondence with molecular methods (WGS).
- The legend provides a key for interpreting the colors used in the table and diagram.