




# Experiences from 4 Years of Organization of an External Quality Assessment for *Mycobacterium tuberculosis* Whole-Genome Sequencing in the European Union/European Economic Area

R. M. Anthony,<sup>a</sup> E. Tagliani,<sup>b</sup> V. Nikolayevskyy,<sup>c</sup> R. de Zwaan,<sup>a</sup> A. Mulder,<sup>a</sup> M. Kamst,<sup>a</sup> C. Ködmön,<sup>d</sup> M. J. van der Werf,<sup>d</sup>  D. Cirillo,<sup>b</sup> D. van Soolingen,<sup>a</sup> ERLTB-Net members

<sup>a</sup>National Tuberculosis Reference Laboratory, Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, The Netherlands

<sup>b</sup>Emerging Bacterial Pathogens Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Milan, Italy

<sup>c</sup>Department of Infectious Diseases, Imperial College London, London, United Kingdom

<sup>d</sup>European Centre for Disease Prevention and Control, Stockholm, Sweden

R. M. Anthony and E. Tagliani contributed equally to this paper. Author order is alphabetical.

**ABSTRACT** Here, we report the development and key features of the first external quality assessment (EQA) scheme for *Mycobacterium tuberculosis* whole-genome sequencing (WGS). The results of four rounds (2017 to 2020) of implementation within the European tuberculosis reference laboratories network (ERLTB-Net-2) are presented and discussed. EQA panels comprising 10 genomic DNAs were distributed to ERLTB-Net 2 laboratories volunteering to participate in this exercise. Since 2018, five FASTQ files were added to better assess the dry WGS processes, and in 2020, three of the five files were replaced by synthetic files (providing additional flexibility for the mutations included in the panels). Ten National tuberculosis reference laboratories participated in all four EQA rounds, and seven participated in at least one. High-confidence resistance mutations were correctly identified by all laboratories, but challenges remained with respect to the identification of mixed loci and interpretation of rare mutations. *M. tuberculosis* genotyping and clustering analysis was >90% accurate for pure samples with the main challenges being related to the analysis of mixed genotypes and DNA FASTQ files. The development and implementation of this WGS EQA scheme has contributed to the continuous improvement in performance of participating laboratories in *M. tuberculosis* WGS and data analysis. This scheme can serve as a model of comprehensive quality assessment for *M. tuberculosis* WGS that can be replicated in different settings worldwide.

**IMPORTANCE** The wider availability of whole-genome sequencing (WGS) coupled to new developments in bioinformatic tools and databases to interpret *Mycobacterium tuberculosis* complex WGS data has accelerated the adoption of this method for the routine prediction of antimycobacterial drug resistance and genotyping, thus necessitating the establishment of a comprehensive external quality control system. Here, we report 4 years of development and results from such a panel.

**KEYWORDS** *Mycobacterium tuberculosis*, *Mycobacterium*, genome analysis, quality assurance

A dramatic decline in the cost of whole-genome sequencing (WGS) and availability of WGS infrastructure, along with analysis methods for *Mycobacterium tuberculosis* sequence data (1–6), has resulted in WGS being implemented by many tuberculosis (TB) national reference laboratories (NRLs) in the European Union/European Economic

**Editor** Po-Yu Liu, Taichung Veterans General Hospital

**Copyright** © 2022 Anthony et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to R. M. Anthony, r.anthony@planet.nl.

ERLTB-Net members participating in the EQA are listed in the Acknowledgments.

The authors declare a conflict of interest. M. J. van der Werf and C. Ködmön are employed by the European Centre for Disease Prevention and Control the funding body of the ECDC/GRANT/2018/001.

**Received** 20 June 2022

**Accepted** 16 November 2022

**Published** 8 December 2022

Area (EU/EEA). WGS can be used for both resistance detection and epidemiological typing/(sub)species identification, meaning that WGS, in addition to being more accurate, is also potentially more cost efficient than previous approaches (7). WGS results generally become available before (full) culture-based susceptibility testing results, and the increased typing resolution allows more accurately targeted epidemiological investigations (8, 9). For these reasons, WGS results are increasingly used to support patient treatment and for public health interventions. *M. tuberculosis* WGS data generation and analysis involves many (nonstandardized) steps and complex interpretation (2, 10); thus, quality control is particularly important. Development of international EQA schemes would also support *M. tuberculosis* WGS standardization and interpretation of the results in the international context, thus contributing to the harmonization of drug resistance prediction (11) and the quality of epidemiological investigations and supporting the potential of WGS to detect uncommon forms of resistance, which are only rarely routinely encountered by any individual laboratory. Furthermore, participation in an EQA scheme is a requirement for ISO 15189 accreditation and contributes to maintaining and improving laboratory proficiency.

The European Reference Laboratory Network for TB (ERLTB-Net-2), comprising 31 TB NRLs in the European Union/European Economic Area member states and the United Kingdom, was established with the aim of consolidating and strengthening TB laboratory capacity, improving quality, and achieving sustainability in TB laboratory diagnosis through provision of training, harmonization of laboratory methods, and development and implementation of reliable EQA systems and standards in TB laboratory diagnosis within the EU/EEA and beyond, as well as supporting the functionality of national TB laboratory networks.

One of the major network activities is related to the development of a sustainable external quality assurance (EQA) system. Since 2010, the ERLTB-Net has developed and implemented several EQA schemes covering all TB laboratory diagnostic methods, from microscopy to *M. tuberculosis* identification using rapid molecular diagnostics, culture, genotypic- and phenotypic-based antimicrobial susceptibility testing, and molecular typing (i.e., mycobacterial interspersed repetitive units variable number of tandem repeats [MIRU-VNTR]) (8, 12–14).

With the gradual shifting from MIRU-VNTR to WGS for *M. tuberculosis* genotyping in several EU/EEA member states and with its increased use for *M. tuberculosis* drug resistance prediction, the ERLTB-Net-2 coordination team decided in 2016 to develop and implement a specific EQA scheme to assess the EU laboratory proficiency for detecting drug resistance and perform *M. tuberculosis* relatedness analysis.

To our knowledge, this is the first EQA scheme for *M. tuberculosis* WGS being developed and implemented on a large scale. Here, we describe its development and report the results of the 4-year (2017 to 2020) implementation within ERLTB-Net-2.

## RESULTS

In 2017, 2018, 2019, and 2020, a total of 13, 12, 14, and 15 EU/EEA TB NRLs of 32, participated in the EQA exercise, respectively.

**Sequencing technology used.** All laboratories utilized Illumina (San Diego, CA, USA) technology to generate their sequence data, apart from one laboratory that used IonTorrent (Thermo Fisher Scientific, USA) for all four EQA rounds.

**Time line for result submission.** After the samples were shipped to participating laboratories, the results were expected within 6 weeks. In 2017, only 5 (38.5%) of the 13 laboratories that joined the EQA activity returned the results within the deadline; the remaining labs requested extensions of up to 21 days. In 2018, only three (25%) laboratories requested an extension. In 2019, two (14.3%) laboratories requested a short extension and returned the results within 8 days of the deadline. In 2020, the great majority, 14 of 15 laboratories, met the deadline, while one laboratory requested an extension due to a COVID-related delay in their sequencing provider.

**Scores.** In 2017, 12 (92.3%) of 13 laboratories scored >75 points and received the EQA proficiency certificate. The remaining laboratory missed the identification of the

**TABLE 1** Summary of the types of determinations scored over the four external quality assessment rounds and the number of errors<sup>a</sup>

Determination	Expected no. of calls	Correctly reported	%
No SNPs associated with resistance	428	428	100
Resistance-associated SNPs	166	151	90
SNPs with moderate or weak association with resistance	46	29	63
Deletion/insertion associated with resistance	26	17	65
Resistance-associated SNPs at $\geq 50\%$ and $< 90\%$ frequency	34	17	50
Resistance-associated SNPs at $\sim 25\%$ frequency	14	7 <sup>b</sup>	50
Resistance-associated SNPs at $< 20\%$ frequency	0	2 <sup>c</sup>	NA

<sup>a</sup>NA, not applicable; SNP, single nucleotide polymorphism.

<sup>b</sup>Twelve laboratories identified the mixed genotype but did not detect the resistant loci.

<sup>c</sup>Two different laboratories detected 5% L449Q (ctg/cAg) in *rpoB* in two different FASTQ files. These loci were flagged as suspicious and potentially associated with (developing) resistance. These positions were checked by the coordinating laboratory and were found to indeed contain a low number of reads (approximately 3 to 10%) containing this SNP. Their significance is unknown but likely represents low level contamination with DNA from a closely related species.

multidrug-resistant tuberculosis (MDR-TB) isolate due to an error in the analysis pipeline, thus scoring  $< 75$  points. In 2018, the number of laboratories receiving the proficiency certificate went down to 10 (83.3%) of 12, and in 2019 and 2020, 100% of participating laboratories received the proficiency certificate.

**Detection and reporting of resistance-associated genotypes.** Overall, during the 4 years EQA time frame, a total of 55 samples, including 15 FASTQ files, three of which synthetic, were provided to the participating laboratories. The four panels included a total of 32 fully susceptible isolates that were correctly characterized by all laboratories; these fully susceptible isolates represented 428 (100%) correct calls.

Single nucleotide polymorphism (SNP) mutations associated with resistance in pure isolates were not always correctly detected: 27 (15%) of 178 resistance-associated calls were missed or incorrectly interpreted (Table 1). In the 2017 panel, an *inhA* c-15t mutation was missed by a single laboratory, while six laboratories failed to report an *embB* M306V mutation. In addition, the panel included a *Mycobacterium canettii* isolate intrinsically resistant to pyrazinamide due to a M117T mutation in the *panD* gene (15), which was reported by only 3 (23%) of the 13 participating laboratories. In 2018, four laboratories failed to report a *pncA* A46V mutation associated with pyrazinamide resistance, and in 2019, six laboratories did not report a silent *fabG* L203L mutation previously reported to be associated with isoniazid resistance (16, 17).

As expected, mutations associated with an only moderately raised MIC and or a weak association to phenotypic drug resistance (18) were more frequently mischaracterized: 17 (37%) of 46 calls of this type were missed or incorrectly interpreted (Table 2). In 2018, an isolate with a double mutation in the *inhA* gene was included; while the c-15t in the promoter region was correctly identified by all 12 laboratories, the I21T in the coding region was reported by 10 (83%) laboratories. In 2020, a synthetic FASTQ file with mutations associated with an increased MIC to bedaquiline and clofazimine (i.e., position 2 GTG-GCG in Rv0678 loss of start codon [19]) and delamanid (i.e., position 3987106 G > A Trp88Stop in Rv3547 [19]) was included in the panel. Mutations in Rv0678 and Rv3547 were correctly reported by four (33%) and seven (58.3%) laboratories, respectively.

Of the 26 frame shifts/insertions and deletions associated with resistance, 17 (65.4%) were accurately reported (Table 1). In 2018, a large deletion in the *pncA* gene (223419 to 229307) was included and was detected and correctly interpreted by 7 (58%) of 12 laboratories. In 2019, a single base pair insertion in the *katG* gene was included and correctly detected and interpreted by 10 (71%) of 14 laboratories. In 2020, a deletion in the *embB* gene with unknown significance for the ethambutol MIC and a double single base pair deletion in the *katG* gene were both reported by 4 (50%) of 14 laboratories.

**TABLE 2** Criteria used for scoring the WGS external quality assessment<sup>a</sup>

Error	Points lost
Incorrect drug resistance-associated call (resistance versus susceptibility)	10
Unreported low-confidence resistance mutation or resistance predicted at the incorrect level (high versus intermediate versus low)	5
Failure to report an MDR-TB (yr 2017) <sup>a</sup>	25
Failure to detect a mixed genotype	10
Failure to detect rifampicin resistance, 50% or more, in a mixed sample	10
Incorrect genotype (sub/species) assigned to an isolate	10
Incorrect assignment of an isolate to a genetic cluster	6

<sup>a</sup>MDR-TB, multidrug-resistant tuberculosis; WGS, whole-genome sequencing.

<sup>b</sup>This heavy penalty for missing an MDR-TB was not included in the later panels as more complex isolates mixed with other strains with different sensitivities or less common resistance mutations were included in these panels. Miscalling individual drugs was still penalized with 10 points when a high-confidence mutation was missed.

In the 2018 EQA panel, a sample with mixed resistant loci at a 50:50 ratio was correctly detected by 10 (83.3%) of 12 laboratories. Conversely, in the 2019 EQA panel, only 7 (58%) of 12 laboratories correctly reported the resistant mutations of a mixed sample with resistant loci at a 25% frequency, even if the presence of a mixed genotype was reported by the majority (85.7%) of participating laboratories. In 2020, of 11 (78.6%) laboratories able to analyze the 3 synthetic FASTQ files included in the panel, 8 (72.7%) and 7 (63.6%) laboratories, respectively, correctly detected the resistant mutations present at 60% frequency and the *pncA* deletion at 40% frequency (Table 1).

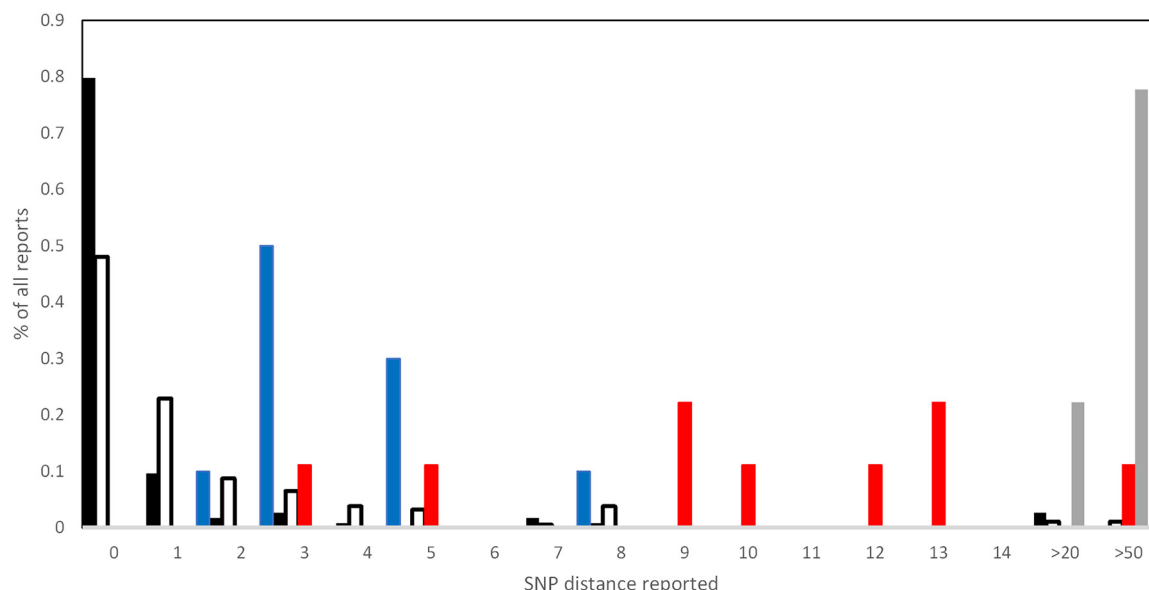
**Identification of lineages.** Identification to the subspecies level was highly accurate in all panels. All laboratories were able to correctly assign the Coll (4) type or genotypic designation to the pure samples included. The naming conventions were consistent within each laboratory but not completely consistent between laboratories. Detection of the mixed genotypes was not required to score full marks; reporting a mixed genotype was sufficient. Nevertheless, in 2019, 9 (64.3%) of 14 laboratories correctly identified the constituents and proportions of the 75:25 mixture of Euro-American (LAM) lineage 4.3.4.2 and East Asian (Beijing) lineage 2.2.1.

**Relatedness analysis. (i) Duplicate isolates.** Over the 4 years, 10 duplicate isolates were included in the panels. Laboratories were expected to cluster these isolates and indicate they were very closely related, suggesting either possible transmission or laboratory contamination. There were a total of 133 duplicate calls expected, of which 131 (98.5%) were correctly identified as closely related. The single pair of identical isolates missed by one laboratory was due to a clerical error. A precise SNP distance was reported for 114 of these duplicate calls, of which 91 (80%) were reported to have zero SNP distance. The distribution of the SNP distances reported is summarized Fig. 1.

**(ii) Epidemiologically clustered but nonidentical samples.** Over the 4 years, 230 relationships between independent epidemiologically linked isolates were included in the panels. Among those, 220 (95.7%) were correctly reported by the laboratories as potentially epidemiologically linked. The 10 missed links were due to a failure to correctly cluster a DNA sample with a FASTQ file. For 170 (77.3%) of the identified links, a precise SNP distance was provided, including three links reported as likely to be epidemiologically linked but with a SNPs distance of 20 (Fig. 1). In 2019, a sample closely related to two other clustered isolates was included in the panel with an estimated SNP distance of 12 SNPs (based on the organizing laboratories analytical pipeline). The SNP distances reported for this sample are indicated in Fig. 1.

Two synthetic FASTQ files in the 2020 panel differed by three SNPs. The two samples were correctly identified by all laboratories as clustered, and 10 (71%) laboratories reported an exact SNPs distance between these isolates which ranged from 1 to 7 SNPs (blue bars in fig. 1).

Overall, there were 18 incorrect epidemiologically linked samples reported; 1 was the result of a clerical error already mentioned above, while the remaining 17 were the result of unrelated isolates being incorrectly joined by the clustering algorithm. This was due to the



**FIG 1** Frequency of scored single nucleotide polymorphism (SNP) distances reported for duplicate samples (black bars), epidemiologically linked samples (empty bars), closely related (approximately 12 SNP distance, red bars), synthetic FASTQ pair (four SNPs and one 3-bp insertion, blue bars), and a genetically close but nonclustered sample (approximately 50-bp distance, gray bars). For the black, empty, and blue bars, a recent epidemiological link should not be ruled out; for the red bars, an epidemiological link is possible but unlikely; and for the gray bars, a recent epidemiological link should be ruled out.

presence of a DNA sample containing a mixture of two isolates belonging to different clusters in both the 2018 and 2019 panels. In 2017 and 2020, a genetically related but not clustered sample with an approximate distance of 50 SNPs was included in the panels. In 2017, none of the 12 laboratories considered this sample epidemiologically linked with any other sample in the panel, while in 2020, one of 14 laboratories indicated this sample as closely related to a cluster. The laboratories that reported an SNP distance for these pairs estimated a distance of between 23 and 81 SNPs, with an average of 64 SNPs (gray bars in Fig. 1).

## DISCUSSION

Of the 32 EU/EEA TB NRLs invited to participate, almost 50% (between 12 and 15) joined this initiative. They represent the current EU/EEA laboratories routinely using WGS or in the process of implementing this technology to characterize *M. tuberculosis* isolates. A rapid switch to the use of *M. tuberculosis* WGS is under way, but as this methodology is not yet fully standardized, the development of an external quality assurance system is critical to assess and compare laboratories' proficiency. *M. tuberculosis* WGS-based genotyping has been shown to be superior to MIRU-VNTR-based genotyping with respect to resolution, reproducibility, and cluster detection (8, 12, 13), and the use of WGS for prediction of resistance and susceptibility to anti-TB drugs shows great promise, even if the precise resistance mechanisms for new and repurposed drugs remain uncertain. In addition, laboratories value *M. tuberculosis* WGS as a potential cost-saving opportunity, for instance by reducing the need for phenotypic drug susceptibility testing (DST) in the presence of fully susceptible isolates.

Importantly, the recent publication by the World Health Organization (WHO) of a curated list of *M. tuberculosis* mutations and their associated link with resistance/susceptibility provides a basis for future development of *M. tuberculosis* WGS, including its use as the reference method for specific drugs and EQAs. This publication represents a significant development with respect to the standardization and acceptance of this methodology for *M. tuberculosis* complex drug susceptibility prediction (20).

**Genotyping.** Overall, the results of WGS-based mycobacterial genotyping were more accurate compared to the ones previously observed in EQA rounds based on MIRU-VNTR typing (12). Nonetheless, some laboratories did report structurally high

SNP distances for closely related isolates, even if the isolates were correctly reported as potentially related. This is likely due to incomplete filtering of noise in the SNP-calling algorithm, often caused by not excluding all poorly mapped reads, used by different laboratories, which makes the comparison of precise SNP distances provided by the EQA participants potentially misleading (21).

The presence of SNP-calling noise was further confirmed by the analysis of the duplicate isolates included in the EQA panel. In this case, only 67% of the links were identified with a zero SNP distance (Table 1; Fig. 1). Thus, low numbers of SNPs should be treated with caution when investigating transmission chains, as in many pipelines, a proportion of these SNPs may be due to artifacts (Fig. 1). In addition, the inclusion of two synthetic FASTQ files generated from an edited reference sequence (22) and differing by three single SNPs, one double nucleotide change, and one 3-bp insertion (sample 12 and 13 in Table 3) were reported by participating laboratories as differing by a number of SNPs ranging from one to seven, which is in line with previous findings (21). These minor differences between pipelines, although having only minor effects when detecting potentially epidemiologically linked isolates, may have implications when inferring transmission chains as the natural accumulation of mutations in *M. tuberculosis* is extremely slow (9).

The inclusion in the panel of samples containing a mixed genotype between isolates belonging to two different clusters resulted in over clustering when the analysis was performed by core genome multilocus sequence typing (cg-MLST). In this case, two unrelated clusters were combined in one cluster with the mixed genotypes isolates linking the two unrelated clusters together. This was largely a consequence of lack of familiarity with the characteristics of the specific clustering algorithms used, as in all cases the laboratories correctly identified the different genotypes (Coll clades) of the artificially clustered isolates. The naming conventions were consistent within each laboratory but not completely consistent between laboratories; this problem related to the complex publication history of these names and was not encountered for the Coll system (4). For this reason, for communication between laboratories, the Coll classification system is favored and will be required in future rounds of this EQA.

Given the lack of a standardization between different analytical pipelines, the inter-laboratory comparison of WGS genotypic results is not fully accurate. Therefore, the primary way to check whether a cluster identified by one laboratory is linked to a cluster identified by another laboratory is by sharing strains, DNA, or FASTQ files between the two sites and having them analyzed by a single pipeline.

This is a critical issue that was partially addressed within the EuSeqMyTB initiative (14, 23) by generating a central WGS data repository in which all sequence data of rifampicin-resistant/multidrug-resistant *M. tuberculosis* strains were analyzed using a common pipeline. Importantly, the lessons learned from this pilot initiative served as a base for the development by the European Centre for Disease Prevention and Control (ECDC) of a WGS-based surveillance system for *M. tuberculosis* within the newly established EpiPulse platform (<https://www.ecdc.europa.eu/en/news-events/launch-epipulse-new-portal-strengthen-prevention-and-control-infectious-diseases>).

**Prediction of drug susceptibility.** The use of WGS for the prediction of drug susceptibility was highly accurate, as all laboratories over the 4 years of this study correctly identified all fully susceptible strains. Similarly, laboratories showed good performance in the detection of mutations associated with drug resistance. It should be noted that these EQA rounds were performed prior to the publication of the WHO catalogue (20). The most common errors were related to: (i) intrinsic limitations of the analytical pipeline used (e.g., unable to detect large gene deletions); (ii) interpretation and reporting of specific resistance-associated mutations present in genes not routinely screened by laboratories (e.g., mutation M117T in the *panD* gene conferring PZA resistance); and (iii) clerical errors. The second type of error is likely overemphasized in this exercise, as in order to challenge laboratories, rarely encountered mutations were included, and









during this exercise, no standard curated list of potentially resistance-associated was available.

All mutations with strong association to phenotypic drug resistance were expected to be reported and the impact on the associated phenotypic resistance was expected to be indicated. Mutations with a moderate or weak association to phenotypic drug resistance were expected to be identified, and further investigations, for example by phenotypic testing, were expected to be suggested. In 2017, one laboratory failed to report a high-confidence mutation for isoniazid resistance located in the gene upstream region. In this case, the specific pipeline used correctly detected the mutation, but it was, however, missed by the user because only the gene-coding regions were included the screening scheme.

Mutations in genes of interest with unknown or weak association with resistance were obviously more challenging (18). The *fabG1* L203L mutation (16, 17) included in 2019 was detected and reported as significant by 8 of 14 laboratories but correctly linked to isoniazid (INH) resistance by only 7 laboratories. When this omission was communicated, one laboratory, which missed the phenotype associated with this mutation, indicated that they have previously seen this mutation but did not link it with an INH-resistant phenotype. This raises an interesting reporting problem, as three laboratories indicated they would not report an undocumented mutation with unknown phenotypic implications to their clinical colleagues. Reporting standards are an ongoing discussion and future EQA panels should reflect these discussions. The recently published list of mutations by the WHO addresses this critical issue (20) and will be used to inform the scoring of this EQA in the following rounds. Nonetheless, even before this list became available, many laboratories were able to identify and even interpret novel mutations. For example, the single base pair deletion/insertion in samples 6 and 10 in the 2020 panel was reported, with 50% of laboratories suggesting confirmation of the resistant phenotype by MIC determination and drug susceptibility testing. This emphasizes the great potential of WGS for MIC prediction as the knowledge base builds (24).

In conclusion, this EQA activity was designed to assess and challenge *M. tuberculosis* WGS pipeline results, interpretation, and reporting—and we believe helped improve—the quality of *M. tuberculosis* WGS analysis in the EU/EEA particularly with respect to the prediction of drug resistance/susceptibility. Intralaboratory genotyping capacity was also assessed, and with the use of the shared FASTQ files, an assessment of interlaboratory reproducibility was also attempted. Overall, WGS-based genotyping EQA results were more reliable and accurate than those based on the previously used reference method, MIRU-VNTR. Even if for now we limited the evaluation to the laboratory capacity to identify likely related strains rather than directly comparing SNPs distance between isolates, we believe that future EQA rounds should look more carefully at the different performance of individual pipelines and help identifying criteria for standardization.

**Study limitations.** We recognize that the WGS EQA scheme described in this paper has some limitations. We assessed the laboratories' capacity to analyze WGS data with respect to drug resistance prediction and relatedness analysis but have not directly evaluated the quality of the generated raw sequence data nor compared the different analytical pipelines or the performance of the various sequencing technologies used (i.e., Illumina platforms and IonTorrent). Although only one laboratory used the IonTorrent platform, we did not observe any impact on the laboratory capacity to detect the resistance-associated mutations or on its capacity to identify likely related clones.

Although the quality of the raw sequence data generated by the different laboratories was not assessed, we did include in the EQA panel five FASTQ files, which would have indirectly indicated if the quality of the raw sequence data was the primary source of errors. This did not appear to be the case. However, the inclusion of these files did highlight some difficulties of some participants to effectively cluster genetically closely related isolates when the WGS sequence was not generated locally. This was likely due to inexperience or possibly by the lack of standardized protocols for the processing and downstream analysis of the shared FASTQ files compared to the ones

generated locally. If this is indeed the case, we expect it to be resolved in future EQA rounds.

Second, during the study period (2017 to 2020), there was no single globally endorsed list of resistance-associated mutations for *M. tuberculosis* complex to guide data interpretation. However, the recent publication by WHO of the catalogue of mutations in *M. tuberculosis* complex and their association with drug resistance (20) resolves this limitation and provides a standard to be used in future EQA rounds.

Third, although more rapid (or slightly delayed) reporting of EQA results was not scored, we acknowledge the importance of providing WGS results within short turnaround times. Therefore, we will discuss the possibility of scoring this indicator in future EQA schemes.

In addition, given the increasing relevance of nontuberculous mycobacteria (NTMs) infections within our network and globally, we will consider the inclusion of more NTM species within future EQA panels. Finally, we acknowledge that fact that this EQA activity is not ISO accredited.

## MATERIALS AND METHODS

**Participants.** National TB reference laboratories in the EU/EEA routinely performing WGS genotyping of *M. tuberculosis* isolates or in the process of establishing this service were invited to enroll in the EQA scheme for WGS. A total of 13, 12, 14, and 15 laboratories within the ERLTB-Net-2 participated in 2017, 2018, 2019, and 2020, respectively. The National Institute for Public Health and the Environment (RIVM) coordinated test panels preparation, distribution, data analysis, and laboratory certification (Fig. S1).

**EQA panel composition.** The four test panels prepared in 2017, 2018, 2019, and 2020 consisted of ten high molecular weight genomic DNA from mycobacterial isolates, supplemented, in year 2018 onwards, with five FASTQ files generated in the organizing laboratory. Then, in 2020, three FASTQ files were replaced by synthetic files to allow for additional flexibility with the inclusion of specific variants not represented in the isolate collection of the organizing institution (Table 3). The panels prepared at the RIVM were distributed as noninfectious inactivated DNA to participating laboratories. The laboratories were asked to perform WGS typing using their standard methodology and to report the results to the RIVM.

Reference results were based on the in-house analysis of WGS sequence data obtained in the RIVM (10) supplemented with phenotypic-based DST data routinely obtained from the selected isolates. As the majority of *M. tuberculosis* isolates in the EU/EEA are phenotypically sensitive to first line antimycobacterial drugs, all panels included at least 50% sensitive isolates, providing the possibility to assess the specificity of drug resistance prediction (i.e., detection of false-positive resistance calls). The remaining *M. tuberculosis* strains were mono- or multidrug-resistant isolates carrying a mix of frequently encountered and rare mutations associated/interim-associated with drug resistance. Strains with rarely encountered mutations in genes associated with drug resistance but with uncertain or no documented association with drug resistance were also included.

All panels contained either two or three duplicate samples consisting of exactly the same DNA. In 2018 and 2019, duplicates consisted of extracted DNA and data obtained by the organizing laboratory shared in the form of FASTQ files. Genetically close isolates ( $\geq 12$  to 50 SNPs apart) not considered to be part of a recent cluster, according to the RIVM pipeline (10), were also included in two panels.

The panels also included DNA from different members of the *M. tuberculosis* complex, as well as other mycobacterial species that were expected to be identified as nontuberculous isolates. The laboratories were asked to define the bacteria at the species/subspecies level.

Finally, panels included artificial mixtures of DNA from two different *M. tuberculosis* complex strains at specific ratio to challenge the laboratories WGS analysis pipeline by mimicking mixed infections or potential laboratory cross-contaminations. The composition of the different EQA panels from 2017 to 2020 is outlined in Table 3.

In 2020, three synthetic FASTQ files (200-bp reads) were included in the panel. They were generated from a reference *M. tuberculosis* reference FASTA (AE000516.2 CDC1551 complete genome [25]) file by CuReSim, a customized tool using the default options to generate synthetic new-generation sequencing reads (22). Mutations were manually introduced into the reference sequence and mixed alleles created by combining synthetic reads generated from different input files by CuReSim and combining them into a single file in the desired ratio. Table 4 reports the specific mutations introduced into the three synthetic FASTQ files in 2020 and their respective frequency. Of note, the results from the analysis of the synthetic FASTQ files did not contribute to the overall EQA assessment score due to their novelty but still reported to the participating laboratories for information.

**EQA panel preparation.** DNA was extracted using the CTAB method (26). The DNA was prepared exactly as previously described for MIRU-VNTR EQA panels (12). Briefly, DNA concentration was measured with the Nanodrop ND-1000 system (Thermo Scientific, Wilmington, DE). For the preparation of test panels, the DNA samples were diluted to a final concentration of 100 ng/ $\mu$ L. The test panels were stored at 4°C and shipped to registered laboratories at room temperature by a courier service. FASTQ files were made available for download from a link provided by email to laboratories who agreed to participate.

**TABLE 4** Mutations introduced into genome sequence AE000516.2 CDC1551 in the synthetic FASTQ prepared for the 2020 panel

Year	Sample	Gene	Mutation introduced (position in reference genome)	Frequency (%)	Interpretation
2020	11	<i>rpoB</i>	Ser450Leu C > T (761155)	60	Rifampicin-resistant
		<i>katG</i>	Ser315Thr G > C (2155168)	60	Isoniazid-resistant
		<i>pncA</i>	14-bp deletion (2289219 to 2289232)	40	Pyrazinamide-resistant
		<i>nirB</i>	Lys686Lys A > G (304923)	100	NA
2020	12	<i>rpoB</i>	Ser450Leu C > T (761155)	100	Rifampicin-resistant
		<i>katG</i>	Ser315Thr G > C (2155168)	100	Isoniazid-resistant
		<i>pncA</i>	14-bp deletion (2289219 to 2289232)	100	Pyrazinamide-resistant
		<i>nirB</i>	Lys686Asn A > G>C (304923)	100	NA
2020	13	<i>rpoB</i>	Ser450Leu C > T (761155)	100	Rifampicin-resistant
		<i>katG</i>	Ser315Thr G > C (2155168)	100	Isoniazid-resistant
		<i>pncA</i>	14-bp deletion (2289219 to 2289232)	100	Pyrazinamide-resistant
		<i>ddn</i> (Rv3547)	Trp88Stop G > A (3987106)	100	Possible delamanid resistance (27)
		<i>Rv0678</i>	Met1Ala T > C (778991), loss of start codon	100	Possible bedaquiline resistance (19)
		<i>nirB</i>	Lys686Asn A > G>C (304923)	100	NA
		<i>Rv1985C</i>	Pro34Pro G > C (2229801)	100	NA
		<i>Rv2845c</i> ( <i>lipQ</i> )	Arg60Leu C > A (2793810)	100	NA
		<i>glpK</i>	Insertion CCC (4139190)	100	NA
		<i>Rv1592c</i>	Ile322Leu T > G (1792777) and Glu321Asp T > A (1792778)	100	NA

**EQA results assessment and reporting.** Laboratories were expected to report the methodology used, any mutations regarded as significant with respect to drug resistance, and the genotype of the isolates, as well as to indicate isolates they regarded as genetically closely related (potentially clustered). Finally, laboratories were requested to provide an interpretation of these results based on their standard reporting of WGS sequence results for routine isolates. Duplicates and isolates belonging to a molecular cluster, defined by a six SNPs distance threshold, were correctly scored if reported as closely related. The reporting of the exact genetic SNPs distance between the different isolates was not mandatory but recorded when provided. The RIVM compared the results from each laboratory to those expected and reported the individual results to each of the laboratories. Laboratories were given the option to query any discrepancies.

An initial conclusion on the submitted EQA results was communicated to each laboratory. Participants were able to query specific results or ask for clarifications to the organizing laboratory. Points were deducted if isolates were not correctly identified or if drug resistance-associated mutations were not reported or correctly interpreted. The criteria used for scoring the EQA results slightly varied over the years (Table 2) to reflect the changes of the EQA panel composition. An EQA WGS proficiency certificate was granted to laboratories that scored at least 75 of 100 points.

**Data availability.** All WGS data generated at RIVM for this study are deposited as FASTQ files in the Sequence Read Archive of the National Center for Biotechnology Information (accession number [PRJNA896516](https://www.ncbi.nlm.nih.gov/sra/PRJNA896516)).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, PDF file, 0.1 MB.

## ACKNOWLEDGMENTS

This study has received funding from the European Centre for Disease Prevention and Control (ECDC) as part of the ECDC/GRANT/2018/001. The funding body had no role in the study design or data analysis and this publication reflects only the authors' views. ECDC is not responsible for any use as a result of the information it contains.

We thank Alejandra Hernandez Segura for helping to generate the synthetic files in 2020.

The coauthors acknowledge essential input from the ERLTB-Net project manager Yen Holicka. The ERLTB-Net-2/WGS EQA participating laboratories are listed here in alphabetical order:

Kathrin Lippert, Alexander Indra, AGES—Institute for Medical Microbiology and Hygiene, Department for Mycobacteriology and Clinical Molecular Biology, Vienna, Austria. Vanessa Mathys, Karine Soetaert, Reference Laboratory Tuberculosis and Mycobacteria Brussels, Belgium. Erik Michael Rasmussen, International Reference Laboratory of Mycobacteriology, Statens Serum Institut, Copenhagen, Denmark. Marjo Haanperä, Mycobacterial Reference Laboratory, Expert Microbiology Unit, Finnish Institute for Health and Welfare, Helsinki, Finland. Florence Morel, National Reference Centre for Mycobacteria, Pitié-Salpêtrière

Hospital, APHP, Sorbonne Université, France. Matthias Merker, Research Center Borstel, Leibniz Lung Center, Borstel, Germany. Emma Roycroft, Margaret Fitzgibbon, Irish Mycobacteria Reference Laboratory, St James's Hospital, Dublin/Department of Clinical Microbiology, School of Medicine, University of Dublin, Trinity College, Dublin, Ireland. Andrea Maurizio Cabibbe, Emerging Bacterial Pathogens Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Milan, Italy. Monique Perrin, Service MycoBac, Laboratoire National de Santé, Dudelange, Luxembourg. Eldholm Vegard, Ola Brønstad, Division of Infection Control, Norwegian Institute of Public Health, Oslo, Norway. Rita Macedo, National Reference Laboratory for Mycobacteria, National Institute of Health "Dr. Ricardo Jorge," Lisbon, Portugal. Laura Herrera-Leon, National Reference Laboratory for Mycobacterial, National Centre for Microbiology, Instituto de Salud Carlos III, Madrid, Spain. Sofia Samper, Laboratorio de Investigación Molecular, Instituto Aragonés de Ciencias de la Salud, CIBER Enfermedades Respiratorias, Zaragoza, Spain. Mikael Mansjö, Ramona Groenheit, National Reference Laboratory for Tuberculosis, Public Health Agency of Sweden, Solna, Sweden. Lucy Taylor, Priti Rathod, National *Mycobacterium* Reference Service South, London, UK, and North and Central, Birmingham, UK.

## REFERENCES

1. Feuerriegel S, Schleusener V, Beckert P, Kohl TA, Miotto P, Cirillo DM, Cabibbe AM, Niemann S, Fellenberg K. 2015. PhyResSE: a web tool delineating *Mycobacterium tuberculosis* antibiotic resistance and lineage from whole-genome sequencing data. *J Clin Microbiol* 53:1908–1914. <https://doi.org/10.1128/JCM.00025-15>.
2. Nikolayevskyy V, Niemann S, Anthony R, van Soolingen D, Tagliani E, Ködmön C, van der Werf MJ, Cirillo DM. 2019. Role and value of whole genome sequencing in studying tuberculosis transmission. *Clin Microbiol Infect* 25:1377–1382. <https://doi.org/10.1016/j.cmi.2019.03.022>.
3. Walker TM, Kohl TA, Omar SV, Hedge J, Elias CDO, Bradley P, Iqbal Z, Feuerriegel S, Niehaus KE, Wilson DJ, Clifton DA, Kapatai G, Ip CLC, Bowden R, Drobniewski FA, Allix-Béguec C, Gaudin C, Parkhill J, Diel R, Supply P, Crook DW, Smith EG, Walker AS, Ismail N, Niemann S, Peto TEA, Modernizing Medical Microbiology (MMM) Informatics Group. 2015. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect Dis* 15:1193–1202. [https://doi.org/10.1016/S1473-3099\(15\)00062-6](https://doi.org/10.1016/S1473-3099(15)00062-6).
4. Coll F, McNerney R, Guerra-Assuncao JA, Glynn JR, Perdigao J, Viveiros M, Portugal I, Pain A, Martin N, Clark TG. 2014. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nat Commun* 5:4812. <https://doi.org/10.1038/ncomms5812>.
5. Jajou R, van der Laan T, de Zwaan R, Kamst M, Mulder A, de Neeling A, Anthony RM, van Soolingen D. 2019. WGS more accurately predicts susceptibility of *Mycobacterium tuberculosis* to first-line drugs than phenotypic testing. *J Antimicrob Chemother* 74:2605–2616. <https://doi.org/10.1093/jac/dkz215>.
6. Revez J, Espinosa L, Albiger B, Leitmeyer KC, Struelens MJ, ECDC National Microbiology Focal Points and Experts Group. 2017. Survey on the use of whole-genome sequencing for infectious diseases surveillance: rapid expansion of European national capacities, 2015–2016. *Front Public Health* 5:347. <https://doi.org/10.3389/fpubh.2017.00347>.
7. Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, Farhat MR, Guthrie JL, Laukens K, Miotto P, Ofori-Anyinam B, Dreyer V, Supply P, Suresh A, Utpatel C, van Soolingen D, Zhou Y, Ashton PM, Brites D, Cabibbe AM, de Jong BC, de Vos M, Menardo F, Gagneux S, Gao Q, Heupink TH, Liu Q, Loiseau C, Rigouts L, Rodwell TC, Tagliani E, Walker TM, Warren RM, Zhao Y, Zignol M, Schito M, Gardy J, Cirillo DM, Niemann S, Comas I, Van Rie A. 2019. Whole genome sequencing of *Mycobacterium tuberculosis*: current standards and open issues. *Nat Rev Microbiol* 17: 533–545. <https://doi.org/10.1038/s41579-019-0214-5>.
8. Jajou R, Neeling A, Hunen RV, Vries G, Schimmel H, Mulder A, Anthony R, van der Hoek W, van Soolingen D. 2018. Epidemiological links between tuberculosis cases identified twice as efficiently by whole genome sequencing than conventional molecular typing: a population-based study. *PLoS One* 13:e0195413. <https://doi.org/10.1371/journal.pone.0195413>.
9. Meehan CJ, Moris P, Kohl TA, Pečerska J, Akter S, Merker M, Utpatel C, Beckert P, Gehre F, Lempens P, Stadler T, Kaswa MK, Kühnert D, Niemann S, de Jong BC. 2018. The relationship between transmission time and clustering methods in *Mycobacterium tuberculosis* epidemiology. *EBioMedicine* 37:410–416. <https://doi.org/10.1016/j.ebiom.2018.10.013>.
10. Jajou R, Kohl TA, Walker T, Norman A, Cirillo DM, Tagliani E, Niemann S, de Neeling A, Lillebaek T, Anthony RM, van Soolingen D. 2019. Towards standardisation: comparison of five whole genome sequencing (WGS) analysis pipelines for detection of epidemiologically linked tuberculosis cases. *Eurosurveillance* 24:1900130. <https://doi.org/10.2807/1560-7917.ES.2019.24.50.1900130>.
11. Nikolayevskyy V, Hillemann D, Richter E, Ahmed N, Van Der Werf MJ, Kodmon C, Drobniewski F, Ruesch-Gerdes S, ERLTB-Net Network. 2016. External quality assessment for tuberculosis diagnosis and drug resistance in the European Union: a five year multicentre implementation study. *PLoS One* 11:e0152926. <https://doi.org/10.1371/journal.pone.0152926>.
12. de Beer JL, Kremer K, Kodmon C, Supply P, van Soolingen D, Global Network for the Molecular Surveillance of Tuberculosis 2009. 2012. First worldwide proficiency study on variable-number tandem-repeat typing of *Mycobacterium tuberculosis* complex strains. *J Clin Microbiol* 50: 662–669. <https://doi.org/10.1128/JCM.00607-11>.
13. de Beer JL, Kodmon C, van der Werf MJ, van Ingen J, van Soolingen D, Participants EM-TMSP. 2014. Molecular surveillance of multi- and extensively drug-resistant tuberculosis transmission in the European Union from 2003 to 2011. *Euro Surveillance* 19:20742.
14. Tagliani E, Anthony R, Kohl TA, De Neeling A, Nikolayevskyy V, Ködmön C, Maurer FP, Niemann S, van Soolingen D, van der Werf MJ, Cirillo DM. 2017. Use of a whole genome sequencing-based approach for *Mycobacterium tuberculosis* surveillance in Europe in 2017–2019: an ECDC pilot study. *Eur Respir J* 57:2002272.
15. Zhang S, Chen J, Shi W, Liu W, Zhang W, Zhang Y. 2013. Mutations in *panD* encoding aspartate decarboxylase are associated with pyrazinamide resistance in *Mycobacterium tuberculosis*. *Emerg Microbes Infect* 2: 1–5. <https://doi.org/10.1038/emi.2013.38>.
16. Ando H, Miyoshi-Akiyama T, Watanabe S, Kirikae T. 2014. A silent mutation in *mabA* confers isoniazid resistance on *Mycobacterium tuberculosis*. *Mol Microbiol* 91:538–547. <https://doi.org/10.1111/mmi.12476>.
17. Torres JN, Paul LV, Rodwell TC, Victor TC, Amallraja AM, Elghraoui A, Goodmanson AP, Ramirez-Busby SM, Chawla A, Zadorozhny V, Streicher EM, Sirgel FA, Catanzaro D, Rodrigues C, Gler MT, Crudu V, Catanzaro A, Valafar F. 2015. Novel *katG* mutations causing isoniazid resistance in clinical *M. tuberculosis* isolates. *Emerg Microbes Infect* 4:1–9. <https://doi.org/10.1038/emi.2015.42>.
18. Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, Hanna D, Kim PS, Liwski R, Zignol M, Gilpin C, Niemann S, Denkinger CM, Fleming J, Warren RM, Crook D, Posey J, Gagneux S, Hoffner S, Rodrigues C, Comas I, Engelthaler DM, Murray M, Alland D, Rigouts L, Lange C,

- Dheda K, Hasan R, Ranganathan UDK, McNerney R, Ezewudo M, Cirillo DM, Schito M, Köser CU, Rodwell TC. 2017. A standardised method for interpreting the association between mutations and phenotypic drug resistance in *Mycobacterium tuberculosis*. *Eur Respir J* 50:1701354. <https://doi.org/10.1183/13993003.01354-2017>.
19. Schena E, Nedialkova L, Borroni E, Battaglia S, Cabibbe AM, Niemann S, Utpatel C, Merker M, Trovato A, Hofmann-Thiel S, Hoffmann H, Cirillo DM. 2016. Delamanid susceptibility testing of *Mycobacterium tuberculosis* using the resazurin microtitre assay and the BACTEC™ MGT™ 960 system. *J Antimicrob Chemother* 71:1532–1539. <https://doi.org/10.1093/jac/dkw044>.
20. World Health Organization. 2021. Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance. World Health Organization, Geneva, Switzerland. License CC BY-NC-SA 3.0 IGO.
21. Walter KS, Colijn C, Cohen T, Mathema B, Liu Q, Bowers J, Engelthaler DM, Narechania A, Croda J, Andrews JR. 2019. Genomic variant identification methods alter *Mycobacterium tuberculosis* transmission inference. *bioRxiv*. <https://doi.org/10.1101/733642>.
22. Caboche S, Audebert C, Lemoine Y, Hot D. 2014. Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. *BMC Genomics* 15:264. <https://doi.org/10.1186/1471-2164-15-264>.
23. Tagliani E, Cirillo DM, Ködmön C, van der Werf MJ, Anthony R, van Soolingen D, Niemann S, Nikolayevskyy V, EUSeqMyTB Consortium. 2018. EUSeqMyTB to set standards and build capacity for whole genome sequencing for tuberculosis in the EU. *Lancet Infect Dis* 18:377. [https://doi.org/10.1016/S1473-3099\(18\)30132-4](https://doi.org/10.1016/S1473-3099(18)30132-4).
24. Holicka Y, Tagliani E, Cirillo DM, Nikolayevskyy V. 2022. Utility of the whole genome sequencing based methodologies in routine European tuberculosis reference laboratory network setting. *Tuberculosis* 134:102185. <https://doi.org/10.1016/j.tube.2022.102185>.
25. Fleischmann RD, Alland D, Eisen JA, Carpenter L, White O, Peterson J, DeBoy R, Dodson R, Gwinn M, Haft D, Hickey E, Kolonay LF, Nelson WC, Umayam LA, Ermolaeva M, Salzberg SL, Delcher A, Utterback T, Weidman J, Khouri H, Gill J, Mikula A, Bishai W, Jacobs WR, Jr, Venter JC, Fraser CM. 2002. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J Bacteriol* 184:5479–5490. <https://doi.org/10.1128/JB.184.19.5479-5490.2002>.
26. Kremer K, van Soolingen D, Frothingham R, Haas WH, Hermans PW, Martín C, Palittapongarnpim P, Plikaytis BB, Riley LW, Yakus MA, Musser JM, van Embden JD. 1999. Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol* 37:2607–2618. <https://doi.org/10.1128/JCM.37.8.2607-2618.1999>.
27. Somoskovi A, Bruderer V, Hömke R, Bloemberg GV, Böttger EC. 2015. A mutation associated with clofazimine and bedaquiline cross-resistance in MDR-TB following bedaquiline treatment. *Eur Respir J* 45:554–557. <https://doi.org/10.1183/09031936.00142914>.