19th webinar of the Virtual Medical Consilium

9.12.2022

How to interpret discordant results from genotypic and phenotypic DST tests for DR-TB diagnosis

World Health Organization
REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

Dr Claudio Köser (cuk21@cam.ac.uk)

Initial remarks

World Health Organization
REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

- My thanks Dr Soudeh Ehsani.
- Key terms and abbreviations are underlined throughout this presentation.
- Please do not hesitate to ask questions (today or later via email).

Background

- Although genotypic drug-susceptibility testing (gDST) assays revolutionized DST, their introduction has inevitably resulted in some discordant results, which had been rare previously as a clinician typically only ever received a single phenotypic DST (pDST).
- We will discuss the different types of discordances between:
  - gDST and pDST results.
  - Different gDST results (for the same or different assays).
  - Different pDST results (mostly the BACTEC mycobacterial growth indicator tube (MGIT) system because this is the most widely used method).

Overview

World Health Organization
REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

- Background to drug-resistant TB, pDST and gDST.
- Discussion of three main classes of errors that can be caused by human factors or inherent limitations of the assay:
  - Random.
  - Systematic.
  - Cut-off.

Background: drug-resistant TB

World Health Organization
REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

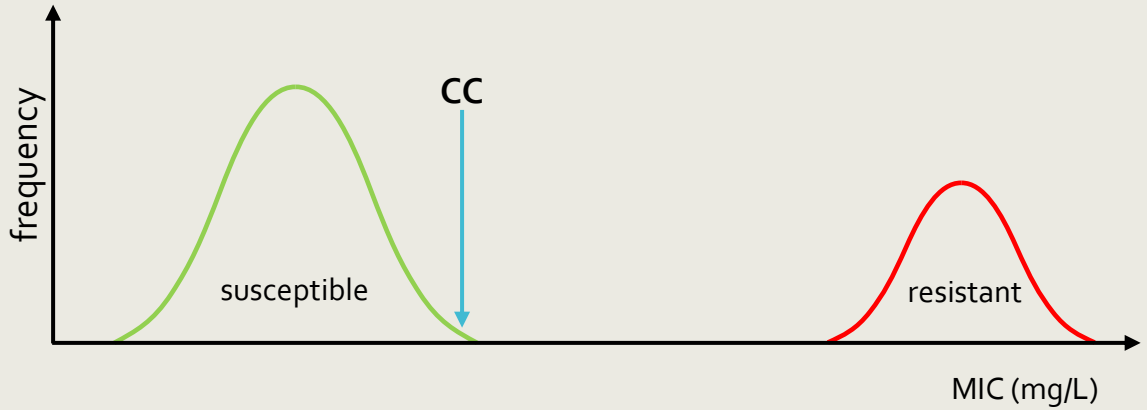
- Four main classes of resistances are recognized:
  - Isoniazid-resistant but rifampicin-susceptible TB: 8% of TB globally.
  - Rifampicin-resistant (RR) TB or multidrug-resistant (MDR) TB, which is resistant to both rifampicin and isoniazid: 3-4% of new and 18-21% of previously treated cases globally compared. Rates are substantially higher in countries of the former Soviet Union (e.g. 38% of new and 69% of previously treated cases in the Russian Federation).
  - Pre-extensively drug-resistant (pre-XDR) TB: MDR/-RR-TB and resistance to levofloxacin or moxifloxacin.
  - Extensively drug-resistant (XDR) TB: pre-XDR and resistance to bedaquiline or linezolid.

gDST assays

- Most widely used WHO-endorsed assays in WHO European region:
  - Cepheid Xpert MTB/RIF (Xpert) or Xpert MTB/RIF Ultra (Ultra) – for rifampicin (*rpoB*).
  - Hain Lifescience GenoType MTBDR*plus* VER 2.0 (FL-LPA) – for ethionamide/prothionamide (*inhA*), isoniazid (*inhA* and *katG*) and rifampicin (*rpoB*).
  - Hain Lifescience GenoType MTBDR*sl* VER 2.0 (SL-LPA) – for fluoroquinolones (*gyrA* and *gyrB*) and second-line injectable drugs (*eis* and *rrs*).
  - Cepheid Xpert MTB/XDR (XDR) – for ethionamide/prothionamide (*inhA*), isoniazid (*ahpC*, *fabG1*, *inhA* and *katG*), fluoroquinolones (*gyrA* and *gyrB*) and second-line injectable drugs (*eis* and *rrs*).

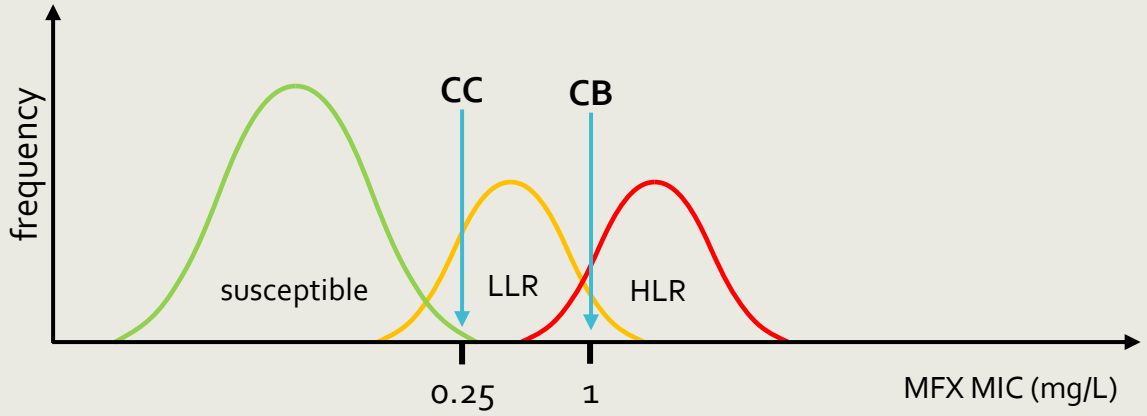
How does pDST work?

- A positive culture is required that is exposed to the drug in question.
- If the minimum inhibitory concentration (MIC) is higher than the critical concentration (CC), a strain is resistant.
- In practice, only the CC is tested in most labs globally.



How does pDST work?

- Moxifloxacin resistance is additionally stratified into low-level resistance (LLR) and high-level resistance (HLR) using the clinical breakpoint (CB). LLR and HLR are an exclusion criterion for shorter moxifloxacin-containing regimen, but high-dose moxifloxacin can be used as part of long individualized regimen for LLR strains.



- Isoniazid resistance is also stratified into LLR and HLR genotypically but an equivalent CB for pDST has not been endorsed by WHO.

How do mutations confer resistance?

- Many different types of mechanisms are possible (e.g. reduced activation or increased inactivation of drug).
- Resistance is regarded to be clinically relevant if at least 1% of the bacterial population is mutated (10% for pyrazinamide) – critical proportion.
- Heteroresistance occurs if the frequency of the resistant population is ≥1% but is below 100%, which may be due to:
  - An originally susceptible strain that is developing resistance.
  - A mixed infection with strains with different susceptibilities to the same drug.

How do gDST assays work?

World Health Organization
REGIONAL OFFICE FOR Europe

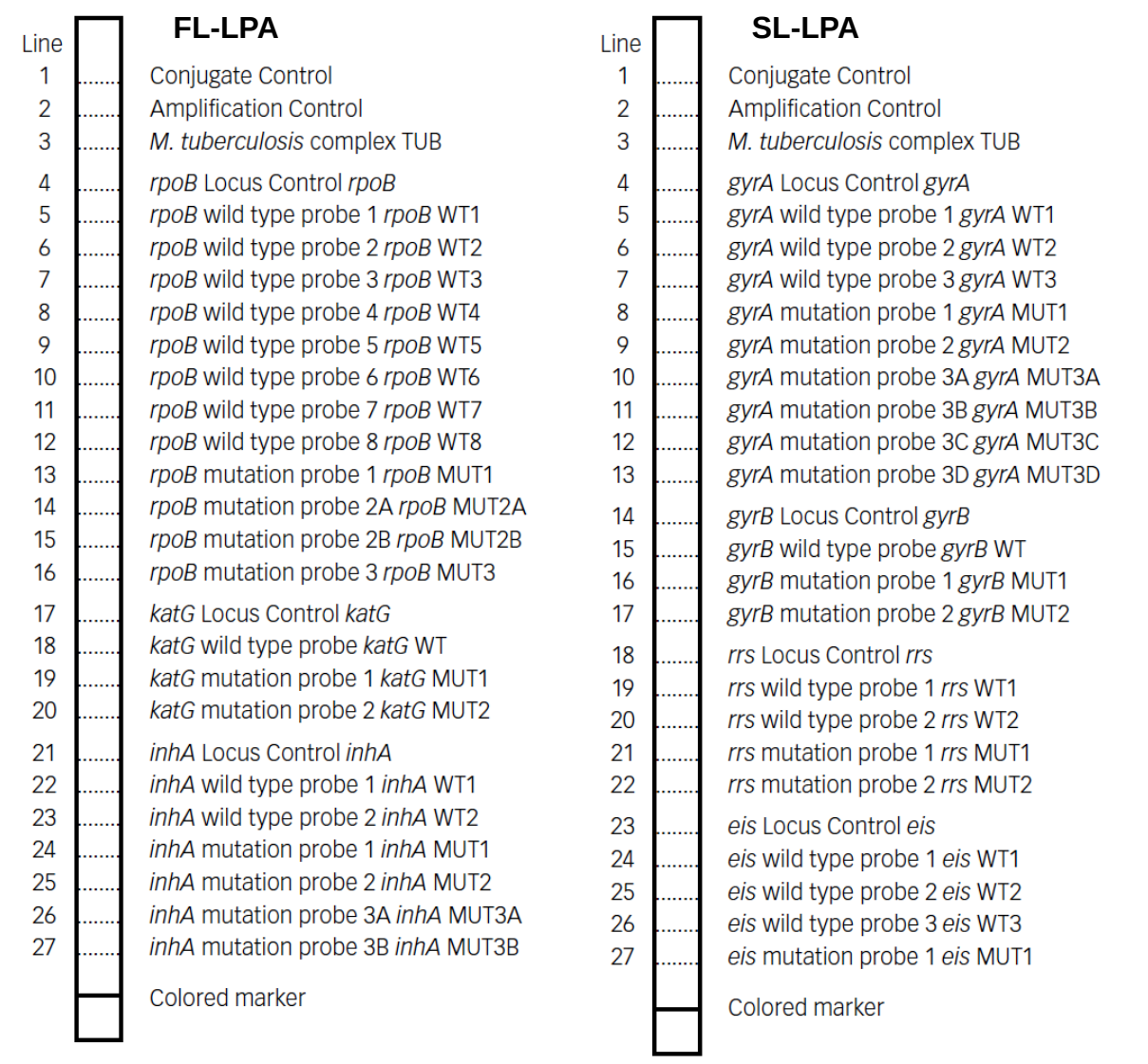
European Laboratory Initiative 2022

- DNA is extracted and amplified. Mutations are then detected or inferred using different techniques.
- Unlike pDST, gDST can be done directly from primary samples, provided that the bacillary load is sufficiently high (i.e. typically smear-positive samples).

How do gDST assays infer or detect resistance mutations?

World Health Organization
REGIONAL OFFICE FOR Europe

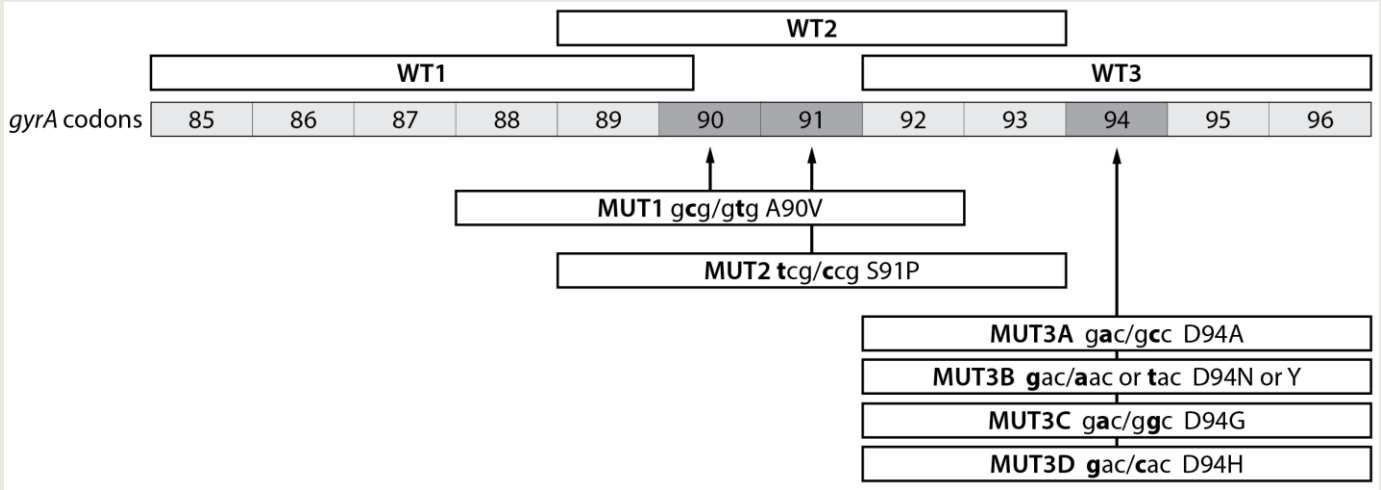
European Laboratory Initiative 2022



How do gDST assays infer or detect resistance mutations?

World Health Organization
REGIONAL OFFICE FOR Europe
European Laboratory Initiative 2022

- Mutant (MUT) probes directly detect resistance mutations by binding to those mutations.
- Wild-type (WT) indirectly infer the presence of a resistance mutation by not binding to the WT sequence, but other neutral mutations that do not confer resistance can also prevent the binding of WT probes, resulting in systematic false-resistant results.
- LPAs use both types of probes (MUT and WT), as shown here for the *gyrA* gene for the SL-LPA:



Reasons behind discordant results

World Health Organization
REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

- Errors due to human factors.
- Inherent limitations of the assays.

Human error can be minimized by a no-blame culture of continuous improvement

- Human errors cannot be eliminated fully but can be minimized by implementing and continuously monitoring a series of measures:
  - Make organizational and technical improvements through the implementation of quality management and quality assurance systems.
  - Foster a no-blame culture, given that operating problems are often complex and usually not the fault of a single person.
  - Recognize that learning from mistakes is an ongoing process.

Human error can be minimized by a no-blame culture of continuous improvement

World Health Organization
REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

RESEARCH ARTICLE

Introduction of quality management in a National Reference Laboratory in Germany

Susanne Homolka[1,2☯*], **Julia Zallet**[1,2☯], **Heidi Albert**[3], **Anne-Kathrin Witt**[1], **Katharina Kranzer**[1,4]

1 National Reference Laboratory for Mycobacteria, Research Center Borstel, Borstel, Germany, **2** Molecular and Experimental Mycobacteriology, Research Center Borstel, Borstel, Germany, **3** Foundation for Innovative New Diagnostics (FIND) South Africa, Cape Town, South Africa, **4** Clinical Research Department, London School of Hygiene and Tropical Medicine, London, England, United Kingdom

☯ These authors contributed equally to this work.
* shomolka@fz-borstel.de

Citation: Homolka S, Zallet J, Albert H, Witt A-K, Kranzer K (2019) Introduction of quality management in a National Reference Laboratory in Germany. PLoS ONE 14(10): e0222925. https://doi.org/10.1371/journal.pone.0222925

What are the three main classes of errors?

World Health Organization
REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

INT J TUBERC LUNG DIS 25(10):791–794

http://dx.doi.org/10.5588/ijtld.21.0428

EDITORIAL

Guidance is needed to mitigate the consequences of analytic errors during antimicrobial susceptibility testing for TB

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology®

COMMENTARY

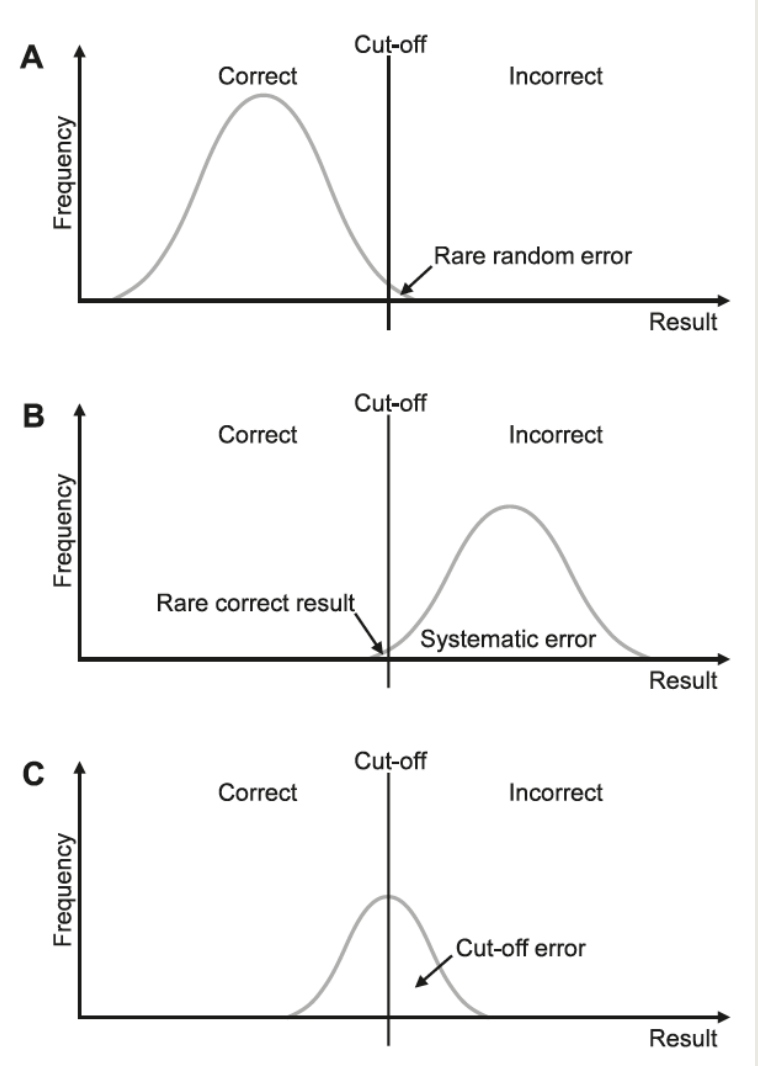
Check for updates

On the Consequences of Poorly Defined Breakpoints for Rifampin Susceptibility Testing of *Mycobacterium tuberculosis* Complex

Claudio U. Köser,[a] Sophia B. Georghiou,[b] Thomas Schön,[c,d,e] Max Salfinger[f]

What are the three main classes of errors?

- Radom errors:
- Systematic errors:
- Cut-off errors:

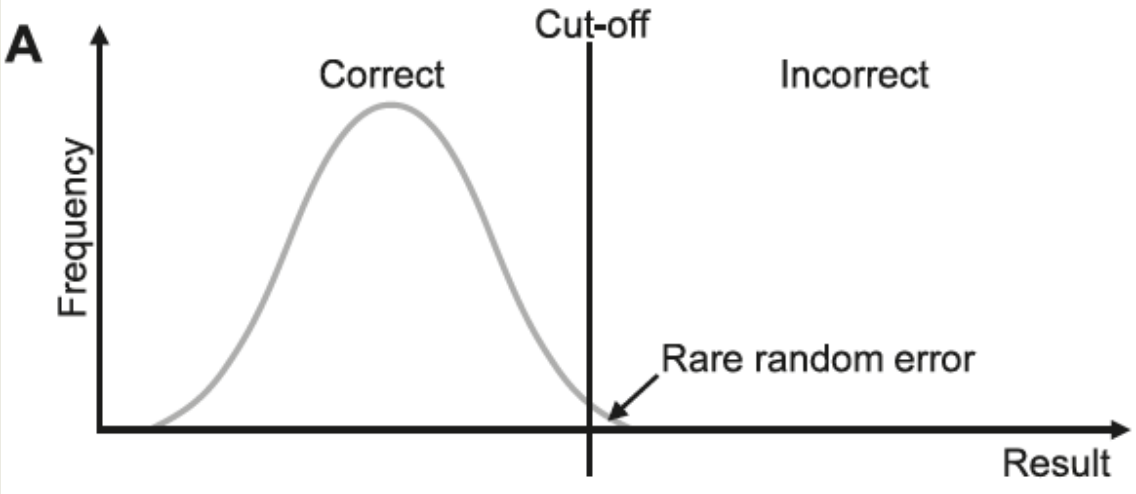


Random errors

- Even if DST is carried out in accordance with good laboratory practice, rare random errors (due to human, instrument or reagent errors) occur for all genotypic and phenotypic methods.
- Random errors occur even when proper validity and acceptability criteria are utilized (e.g. negative and positive controls, as well as process controls yielding expected results).
- Because these errors are random, the full spectrum of potential errors can occur, including indeterminate, false-susceptible or false-resistant results.

Random errors

- Because these errors are rare for well-designed and quality-controlled assays, the assay can simply be repeated utilizing the same primary specimen or cultured sample, with a high likelihood of yielding the correct result (i.e. truly random errors seldom occur consecutively).



Random errors

- As clinicians and microbiologists, you are naturally most interested in resistant DST results.
- However, any DST results (or, in fact, a test result of any kind) must be interpreted in the context of the prevalence (or likelihood/pre-test probability) of resistance given that this has a major impact on the positive predictive value (PPV), which corresponds to the likelihood of an initial resistant result being accurate:
  - $$\mathrm{PPV} = \frac{(\text{number of true-resitant results})}{(\text{number of true-resistant results}) + (\text{number of false-resitant results})}$$

Random errors

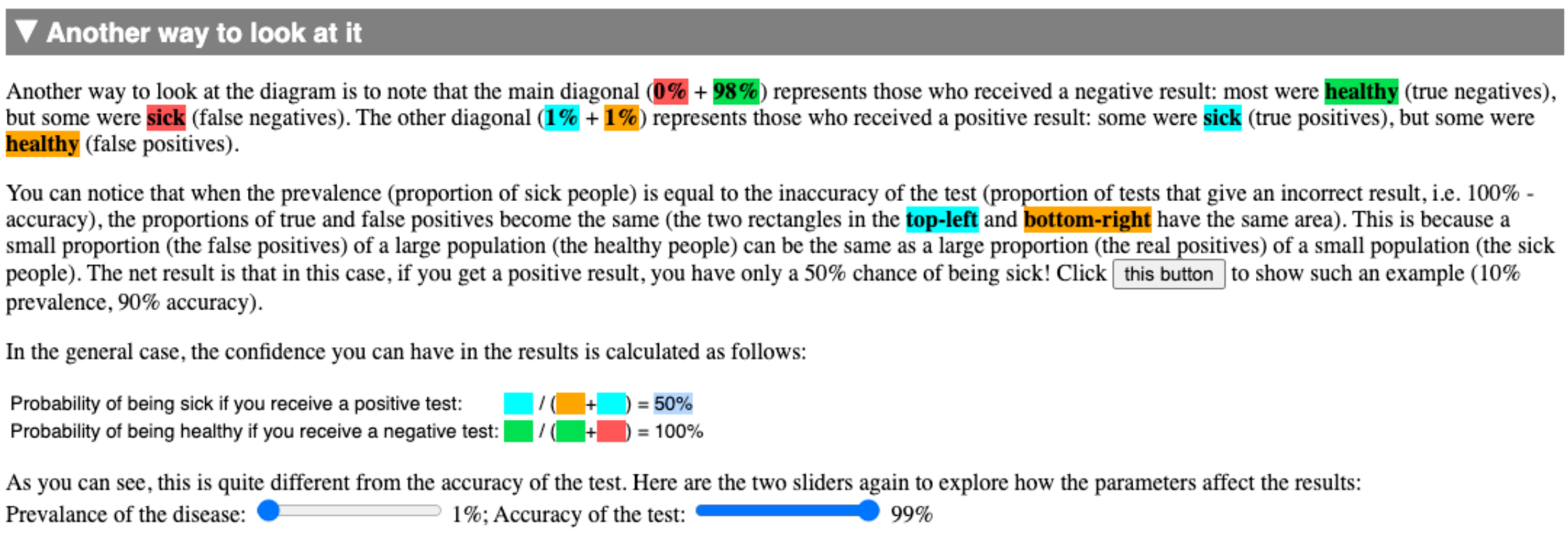
- Let us explore the effect of prevalence of resistance on the PPV for a DST assay that is 99% accurate for 1,000 strains tested:

Prevalence	Resistant strains	Susceptible strains	True-resistant results[1]	False-resistant results[2]	PPV[3]
0%	0	1,000	0	10	0.0%
1%	10	990	9.9	9.9	50.0%
5%	50	950	49.5	9.5	83.9%
10%	100	900	99	9	91.7%
12%	120	880	118.8	8.8	93.1%
20%	200	800	198	8	96.1%
25%	250	750	247.5	7.5	97.1%

[1] = (resistant strains)*99%
[2] = (susceptible strains)*1%
[3] Formula on previous slide.

Random errors

- You can explore this on your own at https://www.lri.fr/~mbl/COVID19/bayes.html:



▼ Another way to look at it

Another way to look at the diagram is to note that the main diagonal (0% + 98%) represents those who received a negative result: most were healthy (true negatives), but some were sick (false negatives). The other diagonal (1% + 1%) represents those who received a positive result: some were sick (true positives), but some were healthy (false positives).

You can notice that when the prevalence (proportion of sick people) is equal to the inaccuracy of the test (proportion of tests that give an incorrect result, i.e. 100% - accuracy), the proportions of true and false positives become the same (the two rectangles in the top-left and bottom-right have the same area). This is because a small proportion (the false positives) of a large population (the healthy people) can be the same as a large proportion (the real positives) of a small population (the sick people). The net result is that in this case, if you get a positive result, you have only a 50% chance of being sick! Click this button to show such an example (10% prevalence, 90% accuracy).

In the general case, the confidence you can have in the results is calculated as follows:

Probability of being sick if you receive a positive test: / (+) = 50%
Probability of being healthy if you receive a negative test: / (+) = 100%

As you can see, this is quite different from the accuracy of the test. Here are the two sliders again to explore how the parameters affect the results:
Prevalance of the disease: 1%; Accuracy of the test: 99%

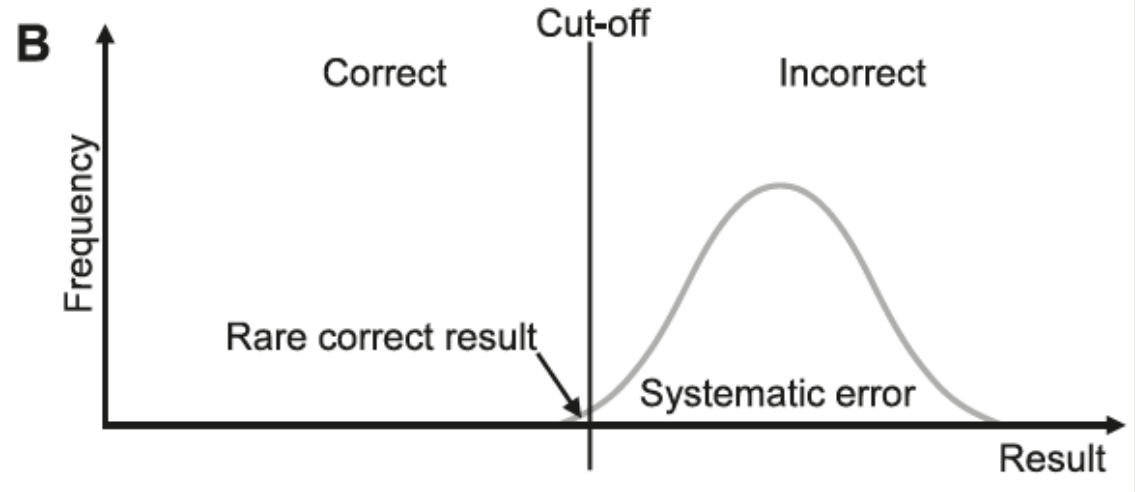
Random errors

- This is why WHO emphasizes that only patients with MDR-TB risk factors should be started on MDR-TB treatment based on a single rifampicin resistant gDST result (i.e. previously treated patients, including those who had been lost to follow-up, relapsed or failed a treatment regimen; non-converters (smear-positive at end of intensive phase); MDR-TB contacts; and any other groups at risk for MDR-TB identified in the country). See: https://apps.who.int/iris/rest/bitstreams/1354706/retrieve
- Crucially, you can correct a random false-resistant result by repeating DST using the same assay as you are unlikely to get two false-resistant results in a row.

Prevalence	Resistant strains	Susceptible strains	True-resistant results[1]	False-resistant results[2]	PPV[3]
1%	10	990	9.9	9.9	50%
10%	100	900	99	9	91.7%
20%	200	800	198	8	96.1%

Systematic errors

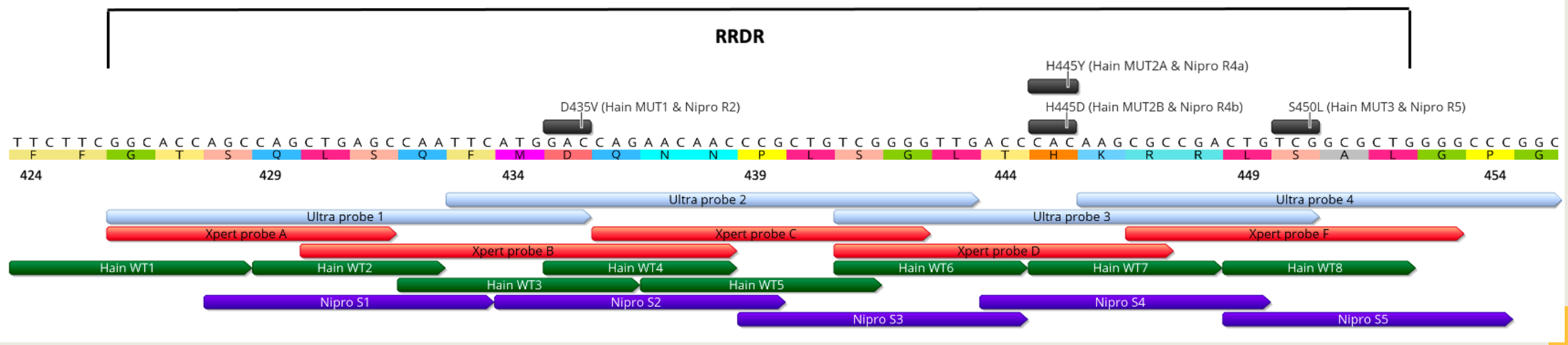
- Unlike rare random errors, completely systematic errors are due to inherent limitations of the assay or procedure and cannot be eliminated by repeat testing (i.e. incorrect results are always obtained, except in rare instances when the result is falsely interpreted as correct because of a random error).



Examples of systematic errors

World Health Organization
REGIONAL OFFICE FOR Europe
European Laboratory Initiative 2022

- gDST false-susceptible results:
  - Not all resistance genes are interrogated (e.g. Cepheid XDR covers *fabG1* for isoniazid, unlike FL-LPA).
  - Even if a gene is analyzed, different assays cover slightly different parts of that gene, as illustrated here for the rifampicin resistance determining region (RRDR):



Examples of systematic errors

- gDST false-susceptible results:
  - The frequency of the resistance mutation in the sample is below the limit of detection (LoD) of the assay but ≥1%, the critical proportion used for pDST. The precise LoDs for heteroresistance depend on the mutation but are approximately):
    - 20–70% for Cepheid Ultra.
    - 5–10% for MUT probes compared with >95% for WT probes for Hain LPAs.

Examples of systematic errors

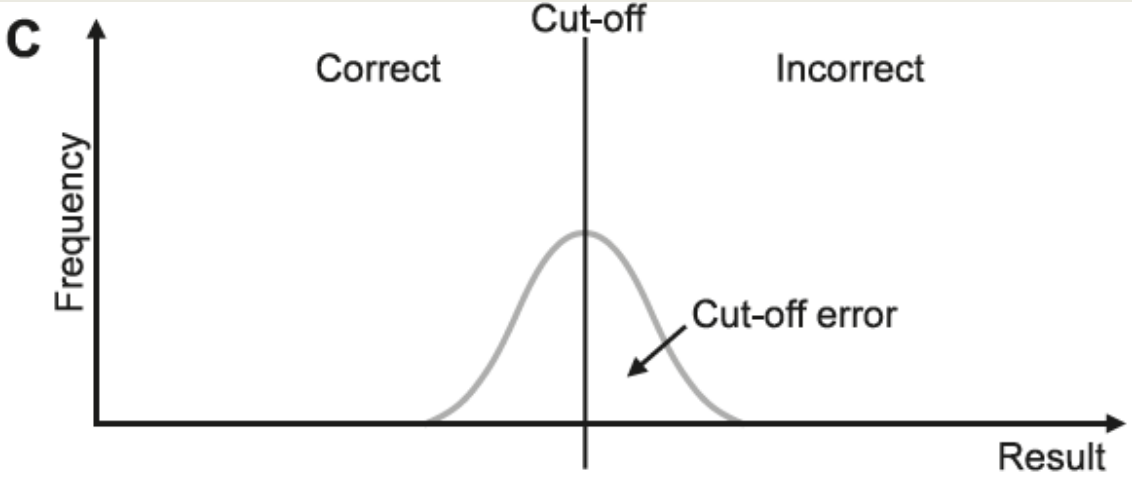
- gDST false-resistant results:
  - Neutral mutations in regions targeted by WT probes (i.e. inferred results). These are usually rare but can be frequent locally relative to the frequency of resistance (e.g. 7% of MDR-TB in Colombia have a mutation that causes false-resistance to fluoroquinolones).

Other factors for systematic differences

- Culture may not be representative of different strains present in primary specimen.
- Contamination of a susceptible culture with a resistant strain.

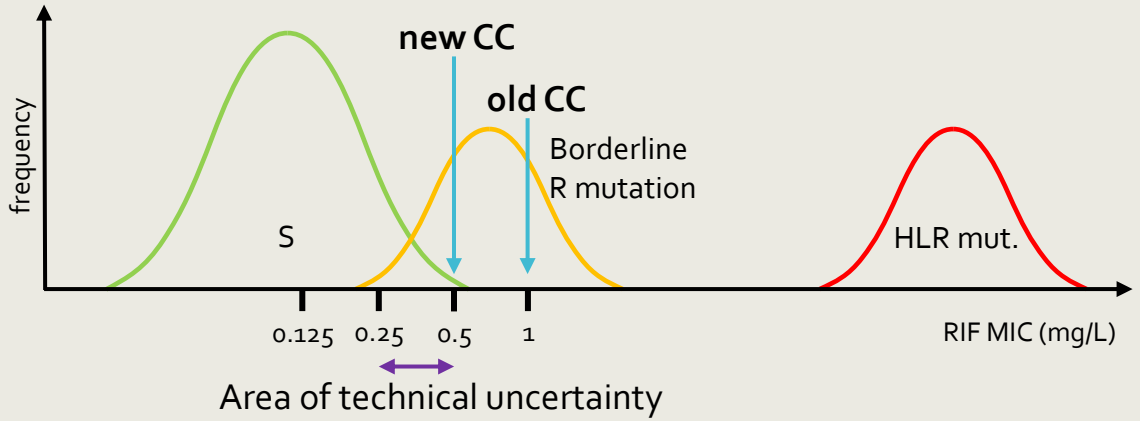
Cut-off errors

- This class of error is systematic in the sense that it only occurs under specific circumstances determined by the cut-off for a particular measurand of the assay. However, unlike the previously discussed systematic errors, they have a large random component (i.e. there is merely a higher likelihood of obtaining an incorrect result, as opposed to near certainty for completely systematic errors).



Examples of cut-off errors

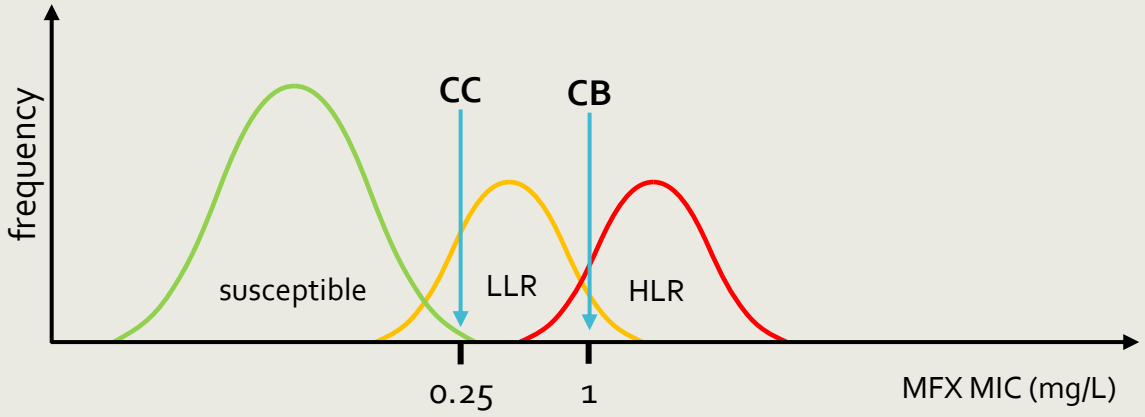
- pDST is not a reliable confirmatory test for borderline rifampicin mutations (*rpoB* L430P, D435Y, H445L, H445N, H445S, L452P, and I491F). Therefore, WHO has endorsed that the detection of these mutation by sequencing overrules a susceptible pDST result.



Mutation type	Misclassified as susceptible at old MGIT CC	Misclassified as susceptible at new MGIT CC
S450 (S531)	0% (95% CI 0-1%)	0% (95% CI 0-1%)
borderline RRDR	74% (95% CI 67-81%)	53% (95% CI 45-61%)
borderline I491F (I572F)	100% (95% CI 69-100%)	80% (95% CI 44-97%)

Examples of cut-off errors

- Moxifloxacin CC is not a reliable method to confirm LLR (e.g. *gyrA* A90V) and CB is not reliable for HLR (e.g. *gyrA* D94G):



- Levofloxacin CC provides a better resolution between susceptible strains and LLR mutants (e.g. it is not uncommon for *gyrA* A90V to test phenotypically resistant to levofloxacin but phenotypically susceptible to moxifloxacin).

Conclusion

- The performance of DST in one setting cannot necessarily be extrapolated to even a nearby setting as strains may differ.
- Set up a system to rapidly spot discordant results and monitor their frequency over time. You are the experts of your country and need to look out for unusual results.
- Consider whether your discordance is likely due to a random, systematic or cut-off error and select the right confirmatory approach if available. You cannot investigate all types of errors (e.g. the only reliable confirmatory test for a borderline rifampicin mutation is sequencing), but you can address many reasons for discordance (e.g. random gDST errors) and exclude some systematic errors (e.g. if inferred gDST results are caused by mutations that confer large MIC increases, for which pDST is a reliable confirmatory test).

Conclusion

World Health Organization
REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

- Reach out to your SRL or ELI experts via the forum of the OpenWHO ELI course
  - English course: https://openwho.org/courses/multi-drug-resistant-tb
  - Russian course: https://openwho.org/courses/multi-drug-resistant-tb-RU

World Health Organization

REGIONAL OFFICE FOR Europe

European Laboratory Initiative 2022

Thank you very much for your attention.