Author's response to reviews

Title: A commercial line probe assay for the rapid detection of rifampicin resistance in Mycobacterium tuberculosis: a systematic review and meta-analysis

Authors:

Maureen Morgan (morganm@ohsu.edu)
Shriprakash Kalantri (sp_kalantri@rediffmail.com)
Laura Flores (floresll@uclink.berkeley.edu)
Madhukar Pai (madhupai@berkeley.edu)

Version: 3 Date: 1 July 2005

Author's response to reviews: see over
Point-by-point response to reviewer’s comments

A commercial line probe assay for the rapid detection of rifampicin resistance in Mycobacterium tuberculosis: a systematic review and meta-analysis

Reviewer 1 (Marcel A Behr)

General

The authors have presented a systematic review that indicates that the commercial line probe assay for rapid detection of rifampin resistance appears to have very good operating characteristics. Therefore, what remains to be determined is whether this assay should be used in clinical settings, for which there appears to be no data at present. It is therefore unclear what take-home message can be derived from this work, other than that further studies are needed. There are two points that I think might bear greater consideration, both to understand the existing data, and to position the potential utility of this test.

Major issues:

1) Bacterial variability. Since different strains in different countries have been associated with different rpoB mutations, I think it is important to determine to what degree the primary studies accounted for this issue. For instance, if one applied a single test for a single rpoB polymorphism during the W strain outbreak in New York, one might easily achieve very high sensitivities for predicting phenotypic rifampin resistance. But does this really represent a denominator or > 200 when you are essentially re-sampling the exact same clone out of > 200 patients affected by the same strain? It would be instructive to learn how many different strains have been tested by this method in order to get a sense of whether this tool will apply equally in different epidemiologic settings. This is an excellent point, and we agree that geographic variation in resistance conferring mutations has the potential to result in variability of the accuracy of LiPA in detecting rifampicin resistance. However, the basis of developing INNO-LiPA was the extensive data to support that greater than 95% of the rifampicin resistant strains have mutations within the region of the rpoB gene examined by the test so that despite regional variation, the vast majority of resistant strains are captured. The studies included in this meta-analysis were conducted in 12 countries on 5 different continents. Many of the studies use specimens collected from multiple countries and/or multiple regions within one country. A high degree of accuracy is retained across studies despite this broad geographic variation in specimen origin, strongly supporting the accuracy of using mutations within the hot spot region of the rpoB gene as an indicator of rifampicin resistance and therefore the robustness of the accuracy of LiPA across geographic regions.

2) Laboratory issues, including cost. While the paper uses “samples” as the unit of analysis, it should be noted that microbiology labs are supposed to receive 2-3 samples per patients at the time of diagnosis, and then further samples to test for culture conversion (or failure, as the case may be). It would be useful to describe whether the studies tested different samples from the same patient, to look for concordance, or whether the studies simply chose one sample per patient, which is likely the case in these sample-based studies.
It’s implied in the majority of studies that each sample is from a different patient, though one study (Gamboa et al) specifies that they use multiple samples from the same patient, each one being a different type of sample (ie sputum, csf, urine, etc).

The next issue is whether one can really estimate the specificity of a test from such non-representative studies. In terms of applying this test on clinical samples, it must be noted that for every 100 sputum samples coming to a lab, only 10 (at most) will be positive for M. tuberculosis. Then, of these 10 positive for M. tuberculosis, it would be remarkable if as many as 1 would be rifampin resistant. Therefore, applying this test directly to sputum would require well over 100 tests for every positive result. To correctly determine the specificity in regular laboratory conditions, it would be important to apply this test on such a ratio of samples, not on a series where 2/3 are resistant. Likewise, if this test were to be used on positive cultures, it might be important to reflect on the costs, based on the estimate provided of $116 per sample. For a laboratory with 10,000 sputum samples per year, there might be 500 positive cultures in a year, representing something like 150-200 distinct patients (multiple samples from a patient). Of these patients, one can then determine the cost based on different prevalence estimates, e.g. 2% rifampin-resistance would yield 3 patients per year. Testing these 500 positive cultures would cost ~$50,000 per year, to advance the detection of rifampin resistance in 3 patients by the interval between a LIPA and a phenotypic result (1-2 weeks?). Perhaps this issue is therefore relevant not just in certain regions of the world but in most labs that have a budget to defend.

This is an extremely important point, and one with which we agree whole heartedly. This is precisely why we feel that in order to examine the clinical usefulness of LiPA (financial feasibility being an essential component of clinical usefulness), future studies should focus on applying direct clinical specimens from patients judiciously selected by experienced clinicians as having a high suspicion of MDR-TB (smear-positive patients with treatment failure or relapse from high incidence areas, previously treated patients, etc) since this is how the test would actually be used. It is in this setting that the test has potential utility, since it would eliminate the need for culture and would therefore decrease the time to definitive diagnosis by up to 2 months when compared with phenotypic results. By using judicious clinical guidelines rather than testing every sputum sample that reaches a laboratory would not only raise the pretest probability, but is likely to be the only way to make LiPA affordable.

To address these points, we have made the following modifications:

For example, if the baseline prevalence of rifampicin resistance is 1%, a positive test would translate into a positive predictive value of only 66%, i.e. one false positive test for every two true positives. As with any diagnostic test, if used judiciously (ie in patients suspected of having MDR-TB, thereby raising the pretest probability) the accuracy of LiPA could be maintained even in low prevalence regions. (page 11, paragraph 1)

The cost of the commercial LiPA kit is $45 per sample tested. When additional costs for import and transport are taken into account, the actual cost per sample is as high as $116 [27]. Though this may be prohibitively expensive for routine use in the regions of the world with the highest prevalence and incidence of TB and MDR-TB, judicious use of LiPA for patients with a high likelihood of MDR-TB (for example, smear-positive patients with treatment failure or relapse from high incidence areas and/or previously treated patients) may be possible, particularly when weighed against the costs of undetected drug resistant TB. (page 12, paragraph 1)
Study design should include selection of sputum samples from patients suspected of having MDR-TB (ie patients with treatment failure or relapse from high incidence areas and/or previously treated patients). (page 13, paragraph 2)

Finally, studies are needed to establish the cost benefit advantages of LiPA over conventional DST. (page 13, paragraph 3)

Minor points:

Page 4. there are 3 references provided to state that 95% of isolates have mutations in a 81-bp hot-spot. First, the literature on this is quite dense, perhaps it would be better to cite a review on this, rather than primary references.

Second, one of the references pertains to E. coli. This has been removed.

Third, it might be worth checking the literature for reports of rifampin resistance that do not have mutations in this hot-spot.
The majority of the rifampicin resistant specimens incorrectly categorized as rifampicin sensitive by INNO-LiPA in the included studies have mutations outside the hot-spot region examined by the test. The fact that the sensitivity and specificity are so high indicates that currently, mutations within the hot-spot region do indeed account for the vast majority of resistance conferring mutations (which is consistent with the literature) and are therefore an accurate marker of rifampicin resistance.

Methods: What does resolved by discussion mean?
We have made the following modifications: Differences between reviewers were reconciled by consensus, and the full text of all relevant studies was evaluated. (page 6, paragraph 1)
Discrepancies between reviewers were reconciled by consensus. (page 6, paragraph 3)

Methods: By a minimum of 10 samples, do the authors require that the samples come from different patients, and is there any requirement that the samples have different genotypes (to avoid repeat testing of the same clone)?
We do not specifically require 10 samples from 10 different patients with 10 different genotypes. Of the studies included in this meta-analysis, only one (Gamboa et al) specifies that they used multiple specimens from the same patients (59 specimens from 31 patients). The majority of remaining studies use specimens collected from multiple continents and/or countries and/or multiple regions within one country, so the implication is that they are including specimens from different patients. In terms of genotypes, each study reports multiple genotypes among the specimens tested, but many of the rifampicin resistant specimens contain one of the four most common resistance conferring mutations (which are the four specific mutations specified by INNO-LiPA).

Discussion: What is meant by regions of the world most in need of rapid detection of RIF-resistant TB? Is the US such a region, based on the experience of the 1990’s?
We have modified the sentence as follows: Though this may be prohibitively expensive for routine use in the regions of the world with the highest prevalence and incidence of TB and MDR-TB, judicious use of LiPA for patients with a high likelihood of MDR-TB (for example, smear-positive patients with treatment failure or relapse from high incidence areas and/or previously treated patients) may be possible, particularly when weighed against the costs of undetected drug resistant TB. (page 12, paragraph 1)

Discussion: “This prevalence of 67% may different from the prevalence seen in …many settings”. I am not aware of any clinical setting where 2/3 of M. tuberculosis isolates are rifampin resistant. I think it would be better to provide a reference or two on the prevalence that one observes in studies in relevant settings, where bacteria were isolated and routinely tested, and then state more definitively that this prevalence is not consistent with routine clinical practice. We have made the following modification: This prevalence of 67% differs significantly from the prevalence of MDR-TB seen in routine clinical practice settings, even in high prevalence regions such as Estonia (14.1%), Henan Province in China (10.8%), Latvia (9%), and the Russian oblasts of Ivanovo (9%) and Tomsk (6.5%). (page 12, paragraph 3)

Discussion: In terms of additional studies, do the authors advocate consecutive selection of patients or consecutive selection of sputum samples? The latter is clearly more conducive to routine laboratory work-flow, so if they suggest the former, it might be useful to determine the rationale. Is it worth including a point of demonstrating the operating characteristics across a broad panel of genotypically different strains? We have made the following modification: Study design should include selection of sputum samples from patients suspected of having MDR-TB (ie patients with treatment failure or relapse from high incidence areas and/or previously treated patients). (page 13, paragraph 2)

Reviewer 2 (Tim Brewer)

General

Overall it is a well done analysis. Page 3, conclusions: As noted throughout the study, the operating characteristics of the LiPA are very good, and consistent across all of the studies examined. What additional evidence do the authors feel is needed "before LiPA can be used to detect MDR-TB..."? Based on these data, it would seem appropriate to consider using LiPA any place where MDR-TB rates are sufficiently high and resources allow. It would seem that cost-effectiveness analyses, rather than more data on LiPA testing performance, would help laboratories and public health organizations decide whether to include LiPA in their testing tuberculosis diagnostic strategies. This is an excellent point, and we agree that cost-effectiveness analyses would be extremely useful. We also feel that, though adequate research has been conducted to demonstrate the accuracy of INNO-LiPA when used on culture isolates, rigorous testing in a typical clinical setting using specific criteria to determine patients with a high clinical suspicion of MDR-TB and testing INNO-LiPA directly on clinical specimens would help a great deal in establishing the actual usefulness of this test in rapidly diagnosing
rifampicin resistance (which would require the use of clinical specimens rather than culture isolates), and developing clear guidelines for its use.

Major Compulsory Revisions

None

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Page 2, Methods
I do not think there is a verb to "meta-analyze".
We have revised this phrase: “perform meta-analysis” (page 2, paragraph 2)

Page 3, paragraph 3
Though MDR-TB remains a major problem in many parts of the world, I did not think that the incidence of tuberculosis that multidrug-resistance was rising. Overall TB incidence is rising globally, but that may not be true for MDR-TB. The authors may want to confirm that the incidence of MDR-TB is rising, not just TB.
We have made the following modification: The prevalence of multidrug-resistant TB (MDR-TB), defined as resistance to at least rifampicin (RIF) and isoniazid (INH), is rising in a number of geographic regions. (page 3, paragraph 3)

Page 6, paragraph 3 and figure 1
The manuscript should include some information about the 6 non-English, non-Spanish studies excluded from the analysis. Did these studies report similar test results to studies that were included? If not, how did their results differ from those studies that were included? If these studies were markedly different from the included studies, their exclusion may bias the results of the meta-analysis. If their results are very similar, then their exclusion is less likely to bias the results.
We appreciate this very important point, and have reviewed each of the 6 abstracts and made the following modification: We excluded studies not available in English or Spanish language, which could introduce publication bias. However, a review of the abstracts of these papers suggests that the overall results are similar to the results in the included English and Spanish language studies. (page 12, paragraph 3)

Page 11, paragraph 4
Without a cost-effectiveness or cost-benefit analysis it is hard to know whether the use of LiPA is "prohibitively expensive". Treatment for MDR-TB with second line agents is much more expensive than treatment for drug-sensitive TB. If an area has the resources to treatment MDR-TB, it may be cost-effective to look for drug resistance even with LiPA at $116 per test (especially with selected testing based on epidemiologic criteria). If an area does not have the resources to treat MDR-TB, there is little point in looking for drug resistance even if the test was cheap. The issue of cost comes up again in the conclusion. The authors assume that because the test may cost $116, it is not practical for Eastern Europe, parts of India and China and elsewhere where MDR rates are high. This assumption ignores the costs of not rapidly diagnosing MDR-TB, which may outweigh those involved with testing.
We have made the following modifications:
Though this may be prohibitively expensive for routine use in the regions of the world with the highest prevalence and incidence of TB and MDR-TB, judicious use of LiPA for
patients with a high likelihood of MDR-TB (for example, smear-positive patients with
treatment failure or relapse from high incidence areas and/or previously treated patients)
may be possible, particularly when weighed against the costs of undetected drug
resistant TB.  

Finally, studies are needed to establish the cost benefit advantages of LiPA over
conventional DST.  

**Discretionary Revisions**

**Page 9, paragraph 5**
When DNA probes for the diagnosis of TB were first studied, similar results were seen.
Sensitivity was lower in clinical specimens, especially smear-negative one. There may
not be enough data, but it would be interesting to know if test characteristics vary by
sputum smear status.  
This would indeed be very interesting. Of the studies included in this meta-analysis,
however, only one (Gamboa et al) includes sputum smear status (sensitivity of LiPA for
smear + = 97.4%, smear - = 100%).
We made the following modification for implications for research: Indeterminate results,
the proportion of RIF-resistant specimens that meet MDR-TB criteria, patients’ sputum
smear status, and turnaround time for diagnosis should be reported.  

**Page 10, paragraph 4**
To describe a scenario, using "For example" or some similar term may be more
appropriate than "I will illustrate..." The reader does not know to which author the "I"
refers to.
We modified this sentence by changing “I will illustrate this by using a hypothetical
example” to “for example” (Page 10, paragraph 4)

This example, and the one that follows on page 11 assume that LiPA is used for all TB
isolates. In low prevalence areas, a more useful (and clinically realistic) scenario would
be to only use the LiPA when the index of suspicion for multiple drug-resistance is high.
Because the pretest probability is higher, test retains its ability to discriminate between
drug-sensitive and MDR-TB in a clinically meaningful way even in low prevalence
settings.
We have added to following sentence: As with any diagnostic test, if used judiciously (ie
in patients suspected of having MDR-TB, thereby raising the pretest probability) the
accuracy of LiPA could be maintained even in low prevalence regions. (page 11, paragraph 1)

**Page 12, paragraph 3**
It would be unusual, and very tragic, for a routine clinical setting to have an MDR-TB
prevalence of 67%.
In order to emphasize how unlikely this prevalence would be, we have made the
following modification: This prevalence of 67% differs significantly from the prevalence
of MDR-TB seen in routine clinical practice settings, even in high prevalence regions
such as Estonia (14.1%), Henan Province in China (10.8%), Latvia (9%), and the
Russian oblasts of Ivanovo (9%) and Tomsk(6.5%).  

(page 12, paragraph 3)