Bacteriophage assays for rifampicin resistance detection in *Mycobacterium tuberculosis*: updated meta-analysis

J. Minion, M. Pai

Department of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal, and Respiratory Epidemiology & Clinical Research Unit, Montreal Chest Institute, Montreal, Quebec, Canada

**OBJECTIVE:** To update a previously reported meta-analysis of evidence regarding the diagnostic accuracy and performance characteristics of commercial and non-commercial phage-based assays for the detection of rifampicin (RMP) resistant tuberculosis (TB).

**DESIGN AND OUTCOMES:** We conducted a systematic review and meta-analysis of test accuracy using bivariate random effects regression and hierarchical summary receiver operating characteristics (HSROC) analysis. Tests included the commercial FASTPlaque™ assays, luciferase reporter phage (LRP) assays, and in-house phage amplification tests. Sensitivity and specificity for RMP resistance were the main outcomes.

**RESULTS:** By updating previous literature searches, a total of 31 studies (with 3085 specimens) were included in this meta-analysis. Evaluations of commercial phage amplification assays yielded more variable estimates of sensitivity (range 81–100%) and specificity (range 73–100%) compared to evaluations of in-house amplification assays (sensitivity range 88–100%, specificity range 84–100%). LRP evaluations yielded the most consistent estimates of diagnostic accuracy, with seven of eight studies reporting 100% sensitivity and four of eight reporting 100% specificity. Estimates of accuracy failed to capture a major failing of the commercial assay, i.e., the rate of contaminated and indeterminate results. These ranged from 3% to 36% in studies looking at direct detection of RMP resistance from patient specimens (mean 20%).

**CONCLUSION:** Phage-based assays will require further development to maximise interpretable results and reduce technical failures. Once technical issues are resolved, impact on patient-important outcomes and cost-effectiveness need to be determined to inform policy for widespread use.

**KEY WORDS:** tuberculosis; bacteriophage assays; rifampicin; MDR-TB; diagnosis
also been developed using D29 phages, and are based on the same underlying principle.

The LRP assay is another in-house format. Although efforts were made to develop a commercial semi-automated version based on the Bronx box (Sequella Inc, Rockville, MD, USA), the company has now discontinued development (Alan Klein, Sequella Inc, personal communication).

In 2005, we performed a systematic review and meta-analysis synthesising the evidence available on the accuracy of phage-based assays for detecting RMP resistance. Evidence summarised in that review was largely restricted to the use of phage assays in the detection of RMP resistance in culture isolates, as data on direct testing were scarce. When used on culture isolates, these assays had high sensitivity, but variable and slightly lower specificity. However, further studies were needed to assess their performance when directly applied to sputum specimens. Since 2005, there have been several new studies looking at not only the accuracy of phage-based assays, but also addressing important test characteristics such as contamination rates and uninterpretable results. With the emergence of XDR-TB and ongoing initiatives (e.g., the EXPAND-TB programme funded by UNITAID) to roll-out rapid tests for drug resistance in high-burden settings, there is a need to update the previous meta-analysis and generate up-to-date evidence that will inform policy making.

OBJECTIVES

The objective of the study was to update a previously reported systematic review and meta-analysis of evidence regarding the diagnostic accuracy and performance characteristics of phage-based assays for the detection of RMP resistance in *M. tuberculosis*.

METHODS

Our review was performed according to a prespecified protocol. Descriptions of study eligibility, search strategy, study selection, data extraction and assessment of study quality are given in detail in the original review. Briefly, we included original clinical studies that evaluated the accuracy of phage-based assays (phage amplification or LRP) for the detection of RMP-resistant *M. tuberculosis* in clinical specimens or isolates. Both commercial tests and in-house assays were included, but separately analysed. With respect to study design, we included diagnostic studies (case-control or cross-sectional) that evaluated the accuracy (sensitivity and specificity) of phage-based assays against a reference standard. Any of the following conventional DST tests were accepted as reference standards: absolute concentration method, proportion method, resistance ratio method, and BACTEC 460 (BD Diagnostics, Sparks, MD, USA) or Mycobacteria Growth Indicator Tube 960 (MGIT 960; BD, USA) methods. Only direct comparisons between the index and reference test were included; results from the analysis of discordant results were excluded because post-hoc discrepant analysis can potentially introduce bias.

We excluded from this review all the studies that used phage-based assays for TB case detection (diagnosis). A previous systematic review has summarised the evidence on phage-based assays for case detection.

We searched the following electronic databases for primary studies and conference abstracts: PubMed, Web of Science, Embase, BIOSIS, Cochrane Library and Google Scholar. The original searches were updated in July 2009. We did not impose language restrictions in our searches. The keywords and search terms used included ‘tuberculosis’, ‘*Mycobacterium tuberculosis*’, ‘mycobacteria’, ‘bacteriophages’, ‘mycobacteriophage’, ‘phage’, ‘fastplaque’, ‘phage amplification’, ‘phage-based’, ‘bacteriophage-based’, ‘luciferase’, ‘sensitivity and specificity’, ‘accuracy’ and ‘predictive value’. We also contacted experts in the field and searched the reference lists from the primary studies and review articles. To identify all relevant studies of commercial assays, we contacted and obtained lists of studies from the commercial kit manufacturer.

We assessed the quality of the studies by using the following criteria: selection bias (consecutive or random sampling of patients/specimens vs. studies that used neither method), blinding (single/double blind vs. unblinded interpretation of phage test and reference standard results) and potential for verification bias (complete vs. partial/differential verification of index test results by reference standards).

Outcome measures

Data were extracted to construct 2 by 2 tables of true-positive, false-positive, false-negative and true-negative values. From these data, we calculated the sensitivity and specificity, where sensitivity refers to the proportion of reference-resistant isolates correctly identified by the phage-based assay and specificity refers to the proportion of reference-susceptible isolates correctly identified by the phage-based assay.

The number of contaminated or indeterminate specimens was extracted and presented as a percentage of the total specimens initially tested. Studies providing further evidence regarding contamination rates, rates of indeterminate results or other user-important characteristics were also assessed and are discussed narratively.

Analysis

Data were analysed using STATA/IC 11.0 (StatCorp, College Station, TX, USA). Forest plots visually displaying sensitivity and specificity estimates and their
95% confidence intervals (CIs) from each study were constructed using MetaDiSc software. These measures tend to be correlated and vary according to thresholds (explicit or implicit cut-off values determining positive vs. negative results). Hierarchical summary receiver operating characteristic (HSROC) curves were analysed to explore the influence of those thresholds and to visually display the relationship between sensitivity and specificity.

Accuracy measures were pooled using bivariate random effects regression models, using the user-written programme ‘metandi’ in STATA. Heterogeneity of accuracy estimates was assessed using the I² statistic. If fewer than four studies evaluated a given assay, their results were not included in the meta-analysis, in part because the bivariate random effects model does not converge with very small numbers of studies.

Subgroup analysis

Results are presented separately for studies evaluating commercial amplification assays, in-house amplification assays and luciferase reporter assays. Studies evaluating the commercial amplification assays were further divided into those that included one or more authors employed by the commercial manufacturer and those that did not. The developer of the commercial FASTPlaqueTB™ tests has recently developed an antibiotic supplement (nystatin-oxacillin-aztreonam [NOA]) to add to their assay for the purpose of inhibiting bacterial and fungal contamination and overgrowth. Evaluations of assays using the NOA supplement are included and discussed separately.

RESULTS

Study selection

Our previously published systematic review in 2005 included a total of 21 studies. Of these, 14 studies used phage amplification assays (including eight studies on the commercial FASTPlaque-TB kits) and seven studies used LRP assays. Our updated searches in July 2009 identified three new evaluations on the FASTPlaque-Response assay, six new in-house phage amplification studies, and one new LRP assay evaluation. In this updated review, the total number of studies is therefore 31 (with a total of 3085 specimens).

Characteristics of included studies

Tables 1–3 describe the studies included in this review. Table 1 shows those studies that evaluated commercial phage-based amplification assays. These included assessments of both FASTPlaqueTB-RIF and FASTPlaque-Response. In addition to the eight studies from our previous review, three new studies were identified, all of which evaluated the FASTPlaque-Response assay. Two of these studies reported results using the NOA supplement separately.

Table 1. Commercial amplification-based phage assays (n = 13)

<table>
<thead>
<tr>
<th>Study, year, reference</th>
<th>Country</th>
<th>Assay</th>
<th>Reference</th>
<th>Inoculum</th>
<th>Sample size</th>
<th>Contaminated or indeterminate</th>
<th>Agreement κ</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Sample</th>
<th>Sample size</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Contaminated or indeterminate % R/S</th>
<th>κ (95% CI)</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Sample</th>
<th>Sample size</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Contaminated or indeterminate % R/S</th>
<th>κ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albert, 2007a‡ South Africa</td>
<td>FASTPlaque-TB-RIF PM</td>
<td>Indirect</td>
<td>No</td>
<td>No</td>
<td>73/12</td>
<td>0</td>
<td>0.96 (0.88–0.99)</td>
<td>1.0 (0.74–1.0)</td>
<td>96 (0.87)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oguz, 200218 Turkey</td>
<td>FASTPlaque-TB-RIF PM</td>
<td>Indirect</td>
<td>NR</td>
<td>No</td>
<td>21/11</td>
<td>0</td>
<td>0.81 (0.58–0.95)</td>
<td>0.82 (0.48–0.98)</td>
<td>81 (0.60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Butt, 200420 Pakistan</td>
<td>FASTPlaque-TB-RIF BACTEC 460 Direct</td>
<td>Yes</td>
<td>Yes</td>
<td>28/11</td>
<td>3</td>
<td>0.86 (0.67–0.96)</td>
<td>0.73 (0.59–0.88)</td>
<td>82 (0.57)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Albert, 2007b23‡ South Africa</td>
<td>FASTPlaque-Response PM Direct</td>
<td>Yes</td>
<td>No</td>
<td>87/142</td>
<td>18</td>
<td>0.95 (0.89–0.99)</td>
<td>0.97 (0.83–0.98)</td>
<td>96 (0.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mole, 200724‡ South Africa</td>
<td>FASTPlaque-Response PM Direct</td>
<td>Yes</td>
<td>No</td>
<td>52/331</td>
<td>26</td>
<td>0.98 (0.91–1.0)</td>
<td>0.96 (0.93–0.99)</td>
<td>96 (0.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Direct inoculum refers to applying patient specimens to assay; indirect inoculum refers to applying a previously isolated strain of M. tuberculosis to assay.
† Added since the previous review.
‡ Supplement added: nystatin, oxacillin, aztreonam.
#R = number resistant; #S = number susceptible; CI = confidence interval; PM = proportion method; NR = not reported.
Table 2  In-house amplification-based phage assays (*n* = 12)

<table>
<thead>
<tr>
<th>Study, year, reference</th>
<th>Country</th>
<th>Assay</th>
<th>Reference</th>
<th>Inoculum*</th>
<th>Appropriate selection†</th>
<th>Blinding reported</th>
<th>Sample size, #R/#S</th>
<th>Contaminated or indeterminate %</th>
<th>Sensitivity % (95%CI)</th>
<th>Specificity % (95%CI)</th>
<th>Agreement with reference % (κ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilson, 199725</td>
<td>UK</td>
<td>In-house (D29)</td>
<td>AC + RR</td>
<td>Indirect</td>
<td>NR</td>
<td>No</td>
<td>9/37</td>
<td>0</td>
<td>1.0 (0.66–1.0)</td>
<td>0.95 (0.82–0.99)</td>
<td>96 (0.87)</td>
</tr>
<tr>
<td>Eltringham, 199926</td>
<td>UK</td>
<td>In-house (D29)</td>
<td>RR</td>
<td>Indirect</td>
<td>NR</td>
<td>No</td>
<td>31/46</td>
<td>15</td>
<td>1.0 (0.89–1.0)</td>
<td>1.0 (0.92–1.0)</td>
<td>100 (1.0)</td>
</tr>
<tr>
<td>McNerney, 200027</td>
<td>South Africa, UK</td>
<td>In-house (D29)</td>
<td>PM + BACTEC 460</td>
<td>Indirect</td>
<td>No</td>
<td>Yes</td>
<td>17/20</td>
<td>0</td>
<td>1.0 (0.81–1.0)</td>
<td>1.0 (0.83–1.0)</td>
<td>100 (1.0)</td>
</tr>
<tr>
<td>Gali, 200328</td>
<td>Spain</td>
<td>In-house (D29)</td>
<td>BACTEC 460</td>
<td>Indirect</td>
<td>No</td>
<td>Yes</td>
<td>18/71</td>
<td>0</td>
<td>1.0 (0.82–1.0)</td>
<td>1.0 (0.95–1.0)</td>
<td>100 (1.0)</td>
</tr>
<tr>
<td>Mani, 200329</td>
<td>India</td>
<td>In-house (D29)</td>
<td>AC</td>
<td>Indirect</td>
<td>No</td>
<td>No</td>
<td>101/100</td>
<td>0</td>
<td>0.97 (0.92–0.99)</td>
<td>0.84 (0.75–1.0)</td>
<td>91 (0.91)</td>
</tr>
<tr>
<td>Simboli, 200510</td>
<td>Argentina</td>
<td>In-house (D29)</td>
<td>PM</td>
<td>Indirect</td>
<td>Yes</td>
<td>Yes</td>
<td>42/87</td>
<td>7</td>
<td>1.0 (0.92–1.0)</td>
<td>0.99 (0.94–1.0)</td>
<td>99 (0.98)</td>
</tr>
<tr>
<td>Da Silva, 200611</td>
<td>Brazil</td>
<td>In-house (D29)</td>
<td>PM</td>
<td>Indirect</td>
<td>NR</td>
<td>No</td>
<td>8/10</td>
<td>0</td>
<td>0.88 (0.47–1.0)</td>
<td>1.0 (0.69–1.0)</td>
<td>94 (0.89)</td>
</tr>
<tr>
<td>Yzquirodo, 200612</td>
<td>Cuba</td>
<td>In-house (D29)</td>
<td>PM</td>
<td>Indirect</td>
<td>NR</td>
<td>No</td>
<td>46/57</td>
<td>0</td>
<td>0.98 (0.88–1.0)</td>
<td>1.0 (0.94–1.0)</td>
<td>99 (0.98)</td>
</tr>
<tr>
<td>Chauca, 200739</td>
<td>Peru</td>
<td>In-house (D29)</td>
<td>PM</td>
<td>Indirect</td>
<td>NR</td>
<td>No</td>
<td>43/50</td>
<td>0</td>
<td>1.0 (0.92–1.0)</td>
<td>0.98 (0.89–1.0)</td>
<td>99 (0.98)</td>
</tr>
<tr>
<td>Traore, 200740</td>
<td>Uganda</td>
<td>In-house (D29)</td>
<td>BACTEC 460</td>
<td>Indirect</td>
<td>NR</td>
<td>Yes</td>
<td>35/114</td>
<td>0</td>
<td>1.0 (0.9–1.0)</td>
<td>0.96 (0.91–0.99)</td>
<td>97 (0.93)</td>
</tr>
<tr>
<td>McNerney, 200741</td>
<td>UK</td>
<td>In-house (D29)</td>
<td>Sequencing</td>
<td>Indirect</td>
<td>NR</td>
<td>No</td>
<td>11/85</td>
<td>0</td>
<td>0.91 (0.59–1.0)</td>
<td>0.96 (0.90–0.99)</td>
<td>96 (0.81)</td>
</tr>
<tr>
<td>Ogwang, 200943‡</td>
<td>Uganda</td>
<td>In-house (D29)</td>
<td>BACTEC 460</td>
<td>Direct</td>
<td>Yes</td>
<td>Yes</td>
<td>16/60</td>
<td>27</td>
<td>0.94 (0.70–1.0)</td>
<td>0.95 (0.83–0.99)</td>
<td>95 (0.87)</td>
</tr>
</tbody>
</table>

* Direct inoculum refers to applying patient specimens to assay; indirect inoculum refers to applying a previously isolated strain of *M. tuberculosis* to assay.
† Studies were considered to have appropriate selection if it was clearly stated that patients/specimens were recruited in a consecutive or random manner.
‡ Added since the previous review.
#R = number resistant; #S = number susceptible; CI = confidence interval; AC = absolute concentration method; RR = resistance ratio method; NR = not reported; PM = agar proportion method.

Table 3  In-house luciferase reporter phage assays (*n* = 8)

<table>
<thead>
<tr>
<th>Author, year, reference</th>
<th>Country</th>
<th>Assay</th>
<th>Reference</th>
<th>Inoculum*</th>
<th>Appropriate selection†</th>
<th>Blinding reported</th>
<th>Sample size, #R/#S</th>
<th>Contaminated or indeterminate %</th>
<th>Sensitivity % (95%CI)</th>
<th>Specificity % (95%CI)</th>
<th>Agreement with reference % (κ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riska, 199937</td>
<td>USA</td>
<td>In-house LRP (Bronx box)</td>
<td>BACTEC 460</td>
<td>Indirect</td>
<td>No</td>
<td>No</td>
<td>10/17</td>
<td>0</td>
<td>1.0 (0.69–1.0)</td>
<td>0.94 (0.71–0.99)</td>
<td>96 (0.92)</td>
</tr>
<tr>
<td>Lu, 200038</td>
<td>China</td>
<td>In-house LRP (luminometry)</td>
<td>AC</td>
<td>Indirect</td>
<td>Yes</td>
<td>Yes</td>
<td>13/9</td>
<td>0</td>
<td>0.92 (0.64–0.99)</td>
<td>0.89 (0.52–0.99)</td>
<td>91 (0.81)</td>
</tr>
<tr>
<td>Banaee, 200119</td>
<td>Mexico</td>
<td>In-house LRP (luminometry)</td>
<td>BACTEC 460</td>
<td>Indirect</td>
<td>Yes</td>
<td>No</td>
<td>3/47</td>
<td>0</td>
<td>1.0 (0.29–1.0)</td>
<td>1.0 (0.93–1.0)</td>
<td>100 (1.0)</td>
</tr>
<tr>
<td>Banaee, 200340</td>
<td>Mexico</td>
<td>In-house LRP (luminometry)</td>
<td>BACTEC 460</td>
<td>Indirect</td>
<td>Yes</td>
<td>No</td>
<td>7/65</td>
<td>0</td>
<td>1.0 (0.59–1.0)</td>
<td>1.0 (0.95–1.0)</td>
<td>100 (1.0)</td>
</tr>
<tr>
<td>Hazbon, 200341</td>
<td>Colombia</td>
<td>In-house LRP (luminometry)</td>
<td>PM</td>
<td>Indirect</td>
<td>No</td>
<td>No</td>
<td>11/37</td>
<td>6</td>
<td>1.0 (0.72–1.0)</td>
<td>0.89 (0.75–0.97)</td>
<td>92 (0.79)</td>
</tr>
<tr>
<td>Hazbon, 200341</td>
<td>Colombia</td>
<td>In-house LRP (Bronx box)</td>
<td>PM</td>
<td>Indirect</td>
<td>No</td>
<td>No</td>
<td>10/34</td>
<td>14</td>
<td>1.0 (0.69–1.0)</td>
<td>0.94 (0.80–0.99)</td>
<td>95 (0.88)</td>
</tr>
<tr>
<td>Bardarov, 200342</td>
<td>Uganda</td>
<td>In-house LRP (luminometry)</td>
<td>BACTEC 460</td>
<td>Indirect</td>
<td>Yes</td>
<td>Yes</td>
<td>1/18</td>
<td>0</td>
<td>100 (0.03–100)</td>
<td>100 (0.82–100)</td>
<td>100 (1.0)</td>
</tr>
<tr>
<td>Banaee, 200831‡</td>
<td>South Africa</td>
<td>In-house LRP (luminometry)</td>
<td>BACTEC 460</td>
<td>Indirect</td>
<td>Yes</td>
<td>Yes</td>
<td>9/182</td>
<td>4</td>
<td>1.0 (0.66–1.0)</td>
<td>1.0 (0.98–1.0)</td>
<td>100 (1.0)</td>
</tr>
</tbody>
</table>

* Direct inoculum refers to applying patient specimens to assay; indirect inoculum refers to applying a previously isolated strain of *M. tuberculosis* to assay.
† Studies were considered to have appropriate selection if it was clearly stated that patients/specimens were recruited in a consecutive or random manner.
‡ Added since the previous review.
#R = number resistant; #S = number susceptible; CI = confidence interval; LRP = luciferase reporter phage; AC = absolute concentration method; PM = agar proportion method.
Figure 1 Forest plots of commercial phage amplification assays (FASTPlaque™ tests; n = 13). Point estimates of sensitivity and specificity from each study are shown as closed squares (for evaluations using indirect isolate inoculation) or open squares (for evaluations using direct specimen inoculation); size of the symbol is proportionate to the size of the study, solid lines represent 95% CIs. *Arms with NOA antibiotic supplement added. Estimates within the top brackets were from studies that included authors employed by the commercial manufacturer. CI = confidence interval; df = degree of freedom; NOA = nystatin-oxacillin-aztreonam.

Figure 2 Forest plots of in-house phage amplification assays (n = 12). Point estimates of sensitivity and specificity from each study are shown as closed squares (for evaluations using indirect isolate inoculation) or open squares (for evaluations using direct specimen inoculation); size of the symbol is proportionate to the size of the study, solid lines represent 95% CIs. CI = confidence interval; df = degree of freedom.
Seven of the commercial product evaluations used indirect specimen inoculation and four used direct patient specimens. Six of the studies evaluating the commercial assays included one or more authors employed by the manufacturer. Studies of in-house phage amplification assays are listed in Table 2. A total of 12 studies were identified, six of which were published since our last review. All used assays using the D29 phage and all but one used indirect culture isolates for inoculation. Table 3 shows the eight studies that evaluated LRP assays, one of which was added since our last review. All LRP evaluations used indirect specimen inoculation.

Accuracy estimates
Sensitivity and specificity estimates from each of the included studies are shown in Figures 1–3. Evaluations of commercial phage amplification assays yielded more variable estimates of sensitivity (range 81–100%) and specificity (range 73–100%) compared to evaluations of in-house amplification assays (sensitivity range 88–100%; specificity range 84–100%). The variability in evaluations of the commercial assay was predominantly a feature of studies that did not have an author employed by the manufacturer (Figure 1), and industry studies reported consistently higher test accuracy than independent studies. LRP evaluations yielded the most consistent estimates of diagnostic accuracy, with seven of eight studies reporting 100% sensitivity and four of eight reporting 100% specificity. The two studies that reported on arms using the NOA supplement to the FASTPlaque-Response assay did not have significantly different estimates of sensitivity and specificity (based on overlapping confidence intervals) from the arms that did not use the NOA supplement (Figure 1).

HSROC models with their respective summary points of pooled sensitivity and specificity are displayed in Figure 4. Using bivariate random effects regression, pooled estimates of sensitivity and specificity were calculated for commercial phage assays, in-house amplification assays and LRP assays (Table 4). Overall, LRP assays had the highest accuracy (sensitivity = 99.3%, specificity = 98.6%), with in-house phage amplification assays also performing well (sensitivity = 98.5%, specificity = 97.9%). Estimates from studies evaluating the commercial FASTPlaque kits were slightly lower (sensitivity = 95.5%, specificity = 95.0%); however, this difference was not statistically significant (based on overlapping confidence intervals) from the LRP and in-house assays. When studies evaluating the FASTPlaque products were separated based on whether direct or indirect specimens were used, those using indirect specimens had slightly higher sensitivity but lower specificity compared to those using direct specimens (Table 4). Studies that

Figure 3 Forest plots of luciferase reporter phage assays (n = 8). Point estimates of sensitivity and specificity from each study are shown as closed squares (for evaluations using indirect isolate inoculation); size of the symbol is proportionate to the size of the study; solid lines represent 95% CIs. CI = confidence interval; df = degree of freedom.
included authors employed by the commercial manufacturer reported more accurate results (sensitivity = 96.9%, specificity = 96.7%) compared to studies without industry-employed authors (sensitivity = 92.6%, specificity = 85.1%), with the difference in specificity reaching statistical significance (Table 4).

**Contamination rates**

The rate of specimens that were either contaminated or yielded indeterminate results ranged from 0% to 36% (mean 5.8%, 95%CI 2.2–8.9). This was primarily a problem for evaluations using direct patient specimens, where the rates of uninterpretable results ranged from 3% to 36% (mean 21.2%, 95%CI 9.3–33.0). By comparison, 18/27 arms from studies using indirect specimen inoculation did not report any uninterpretable results and the mean rate was only 2.1% (95%CI 0.5–3.7). The highest rates of uninterpretable results were seen in evaluations of the FASTPlaque tests; however only one of the evaluations of in-house amplification assays or LRP assays used direct specimen inoculation.

Two of the new FASTPlaque-Response studies also included arms evaluating the addition of the NOA supplement.22,24 The addition of NOA reduced the proportion of contaminated specimens from 16% to 5%,24 and from 1.4% to 0.5%22 in these studies.

**DISCUSSION**

In addition to the 21 studies reviewed previously,6 an additional 10 evaluations of phage-based assays detecting RMP resistance in M. tuberculosis were identified in this review. Six new studies presented results using in-house phage amplification assays,31–36 three evaluated the FASTPlaque-Response22–24 and one reported on an LRP test.43 Thus, in this updated review, data from 31 studies are included, making it the most comprehensive review to date.

Accuracy estimates varied between the types of test and the type of inocula used. There was a tendency for the in-house assay evaluations to yield more favourable results compared to those evaluating the commercial products. LRP studies reported the most consistently high estimates of sensitivity as well as specificity. The majority of studies that tested direct patient specimens were those evaluating the FASTPlaque-Response assay, and subgroup analysis showed these to have lower sensitivity but higher specificity compared to studies using indirect isolates.

Studies that had close involvement of the commercial manufacturer or were conducted entirely by the manufacturer had higher estimates of accuracy compared to those performed independently. This may be due to inexperience of the personnel performing the assays without the guidance of the manufacturer or

![Figure 4 A. HSROC of commercial phage amplification assays (FASTPlaque™ tests; n = 13). B. HSROC of in-house phage amplification assays (n = 12). C. HSROC of luciferase reporter phage assays (n = 8). Individual studies are shown as open circles whose size is proportionate to the size of the study. Summary point is shown as a closed square, representing sensitivity and specificity estimates, pooled using bivariate random effects regression. HSROC curve is truncated outside of the area for which data exist. HSROC = hierarchical summary receiver operating characteristic.](image)

**Table 4** Pooled estimates using bivariate random effects regression

<table>
<thead>
<tr>
<th>Test (no. of studies)</th>
<th>Pooled sensitivity (95%CI)</th>
<th>I² (P value)</th>
<th>Pooled specificity (95%CI)</th>
<th>I² (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial phage amplification (n = 13)</td>
<td>95.5 (92.2–97.4)</td>
<td>55.6 (0.008)</td>
<td>95.0 (91.3–97.2)</td>
<td>69.8 (0.0001)</td>
</tr>
<tr>
<td>Indirect only (n = 8)</td>
<td>95.7 (90.8–98.0)</td>
<td>65.8 (0.005)</td>
<td>94.1 (88.3–97.2)</td>
<td>70.5 (0.001)</td>
</tr>
<tr>
<td>Direct only (n = 5)</td>
<td>93.6 (88.0–96.7)</td>
<td>2.1 (0.4)</td>
<td>96.3 (91.6–98.4)</td>
<td>22.6 (0.25)</td>
</tr>
<tr>
<td>Industry (n = 8)</td>
<td>96.9 (93.1–98.7)</td>
<td>49.6 (0.05)</td>
<td>96.7 (95.1–97.8)</td>
<td>22.6 (0.25)</td>
</tr>
<tr>
<td>Non-industry (n = 5)</td>
<td>92.6 (83.5–96.8)</td>
<td>62.0 (0.03)</td>
<td>85.1 (72.3–92.6)</td>
<td>63.2 (0.03)</td>
</tr>
<tr>
<td>In-house phage amplification (n = 12)</td>
<td>98.5 (96.1–99.4)</td>
<td>14.3 (0.3)</td>
<td>97.9 (94.8–99.2)</td>
<td>73.9 (&lt;0.0001)</td>
</tr>
<tr>
<td>Luciferase reporter phage (n = 8)</td>
<td>99.3 (49.1–100)</td>
<td>0.0 (0.9)</td>
<td>98.6 (92.5–99.8)</td>
<td>71.2 (0.001)</td>
</tr>
</tbody>
</table>

CI = confidence interval.
inadequate training in their use of the test. This situation also applies when comparing the use of a test by its academic developers to its attempted reproduction by independent investigators. These types of issues and related problems in transferring a technology from its original research or development settings to more pragmatic, implementation-type settings are important to keep in mind.

An assay that performs well in the hands of the manufacturer or original developer, but does not achieve the same level of performance when applied by users under field conditions, can lead to overly optimistic estimates of accuracy and have potentially disastrous consequences if invested in by programmes without proper internal validation procedures. Studies evaluating diagnostics under programmatic conditions, independently of the manufacturer or industry sponsorship, provide a more useful type of evidence for decision- and policy-makers, and need to be encouraged by funding sources and academic journals alike. Unfortunately, there is evidence that nearly 40% of diagnostic studies in TB, human immunodeficiency virus (HIV) and malaria have some degree of industry involvement or support. 44

A concern that has been identified in multiple publications using direct patient specimens is the high rate of contamination. We found a mean rate of 21% uninterpretable results across studies using direct specimen inoculation, with estimates from programmatic evaluations being even higher. In response to these concerns, Biotec Labs, the manufacturer of the commercial phage-based amplification assays, has developed an antibiotic supplement (NOA) to help control bacterial and fungal overgrowth. Two studies included assessments of the FASTPlaque-Response assay with the NOA supplement added and found that uninterpretable results decreased by an average of 68%.

Although not using the phage technology for resistance testing, and thus not included in the primary analysis of this review, evaluations of phage assays for M. tuberculosis detection have also reported useful information regarding contamination rates. A previous systematic review on phage for detection was published in 2005, 9 but only three of the studies included reported rates of contamination.45-47 Contamination rates in these studies varied widely (2.5–40.4%) between studies and depending on whether antibiotic supplementation was used. Subsequent to this review, other studies have been published on the use of FASTPlaque assays for TB diagnosis. Notably, when using FASTPlaqueTB to diagnose smear-negative TB in an HIV-endemic population, Bonnet et al. reported that 46.8% (95/203 specimens) of their tests were uninterpretable due to contamination and a further five yielded indeterminate results. 48 Of those tests that were considered interpretable, some amount of contamination was detected in another 71, bringing the total rate of any contamination to 81.8%.

Due to a lack of consistent reporting, we were unable to compare the rate of culture contamination using solid or liquid detection media in the same laboratories evaluating phage-based assays. It was initially believed that phage-based assays using direct patient specimens would be feasible for implementation in low-resource settings, in laboratories without the biosafety infrastructure required for mycobacterial culture, as no amplified TB isolates were being used. However, if a laboratory does not have technical expertise and experience in implementing and performing routine TB cultures, it is unlikely that it will have the capability to perform quality-assured mycobacterial decontamination needed for the performance of direct phage-based assays. Until it is clearly demonstrated that phage-based assays are appropriate for more peripheral settings, the ability to perform quality-assured mycobacterial culture with acceptable rates of contamination should be considered a prerequisite to adopting phage-based assays.

The Foundation for Innovative New Diagnostics (FINN) had also planned extensive evaluations of the FASTPlaque-Response assay for the detection of RMP resistance in 2007. According to the FIND website, 49 during the initial phase of demonstration projects at two trial sites in South Africa, the FASTPlaque-Response test failed to meet required performance targets. As a consequence, FIND decided to discontinue all activities with the FASTPlaque assay until improvements or satisfactory alternatives were available.

Biotec Labs has made modifications in the protocol for the FASTPlaque assays, but reported that while modifications increased the sensitivity in smear-negative specimens, this was associated with high rates of false-positive results. 50 The authors from a study conducted in Kenya concluded that ‘further work is required to improve the specificity of the test using the modified protocol. Furthermore, the poor performance of the FASTPlaqueTB test in this study needs to be resolved before investigation of a modified protocol can be made’. 50

The high accuracy and low rates of uninterpretable results found in the evaluations of the in-house assays performed on M. tuberculosis isolates suggest that the principle of the tests is sound, and given the very rapid time to result of these assays, they may find a role in more centralised laboratories with established TB culture services. Molecular detection tests for the direct detection of drug resistance (i.e., line-probe assays) are also available commercially, and a meta-analysis reported pooled estimates of sensitivity and specificity of respectively 98.1% and 98.7%. 51 This is comparable to the accuracy achieved by the non-commercial phage-based assays reviewed here; however, line-probe assays showed equivalent performance using direct patient specimens compared to M. tuberculosis isolates. 51 Caution is also warranted when using non-commercial assays without standardised procedures, reagents and techniques.
This review had several strengths, including the use of a standardised pre-specified protocol and a broad and inclusive search of the published literature. Subgroup analysis was performed according to the type of test evaluated as well as the type of specimen used. Furthermore, rigorous statistical methods were employed using bivariate random effects models and HSROC curves for estimating diagnostic performance measures.

The review was limited by the types of outcomes reported by the primary studies. Specifically, none of the studies in our review reported data on the clinical impact of detection of RMP resistance on long-term treatment outcomes such as cure rates, treatment failures and mortality. Furthermore, although rates of contaminated or indeterminate specimens were assessed, the primary pooled estimates of sensitivity and specificity do not account for these uninterpretable specimens. The true impact of implementing these assays would depend on not only the accuracy of drug resistance detection, but also on the cost-effectiveness, reliability and incremental benefit of their results on patient-important outcomes.

CONCLUSIONS

Despite generally good concordance of phage-based assay detection of RMP resistance in *M. tuberculosis* with conventional DST reference standards, significant concerns have been identified about the rate of contamination when using direct patient specimens. An antibiotic supplement to help control bacterial and fungal contamination has been developed and appears to reduce the number of uninterpretable samples without compromising sensitivity and specificity. Standardisation of methods to minimise the number of contaminated or indeterminate results is necessary before this diagnostic tool can be implemented widely. Furthermore, recent data from the company itself and other field studies have raised major concerns about the existing commercial FAST-Plaque assays. Until these concerns are resolved, there is no commercial phage assay that can be implemented. While the LRP assay appears to be the most consistently accurate test, efforts to commercialise the Bronx box LRP assay have been abandoned. Furthermore, with the 2008 World Health Organization (WHO) endorsement of highly accurate line-probe assays and their subsequent roll-out, the exact scope and role of phage-based assays, if any, need to be determined.

The evidence in this review was presented to a WHO Expert Group in September 2009 to help guide policy decisions regarding the recommendation of diagnostics for rapid detection of drug resistance in *M. tuberculosis*. Due to concerns regarding the performance of the current FASTPlaque assay on direct patient specimens, and the lack of data on the use of non-commercial phage-based assays on direct patient specimens, the Expert Group agreed that there was insufficient evidence to recommend the use of phage-based assays for rapid screening of patients suspected of having MDR-TB.

Acknowledgements

The authors are grateful to K Weyer (WHO, Geneva, Switzerland) for helpful input and feedback. The authors also thank R O’Brien (FIND, Geneva, Switzerland), H Albert (FIND, Switzerland), A Trollip (Bioteck Labs, UK), A Klein (Sequella, USA) and R McNerney (London School of Hygiene & Tropical Medicine, London, UK) for helping to identify relevant publications and information. This study was supported in part by funding from the Stop TB Department, Geneva, Switzerland, and Canadian Institutes of Health Research (CIHR MOP-89918). JM is a recipient of a Quebec Respiratory Health Training Fellowship. MP is a recipient of a CIHR New Investigator Award and a Fonds de la recherche en santé du Québec (FRSQ) establishment grant.

References

49 Foundation for Innovative New Diagnostics (FIND). FIND interrupts demonstration projects using phage-based assays for
Bacteriophage meta-analysis

**RÉSUMÉ**

OBJECTIF: Mettre à jour une méta-analyse publiée antérieurement concernant les évidences au sujet de la précision diagnostique et des caractéristiques de performance des tests basés sur les phages commerciaux ou non commerciaux pour la détection des germes tuberculeux résistants à la rifampicine.

SCHEMA ET RÉSULTATS: Revue systématique et méta-analyse de la précision des tests utilisant une régression bivariée à effets aléatoires d'une courbe SROC (summary receiver operating characteristic) hiérarchique. Les tests ont comporté les tests commerciaux FASTPlaqueTM, les tests LRP (luciferase reporter phage) et des tests faits maison d’amplification des phages. Les principaux résultats ont été la sensibilité et la spécificité pour la résistance à la rifampicine.

RÉSULTATS: Grâce à la mise à jour de recherches antérieures de la littérature, on a pu inclure dans cette méta-analyse 31 études (avec 3.085 échantillons). Les évaluations des tests commerciaux d’amplification des phages ont donné des estimations plus variables de la sensibilité (marge 81–100%) et de la spécificité (marge 73–100%) par comparaison avec les évaluations des tests faits maison d’amplification (marge de sensibilité 88–100% et marge de spécificité 84–100%). Les évaluations LRP ont fourni les estimations les plus régulières de la précision diagnostique, où sept des huit études ont signalé une sensibilité de 100% et quatre des huit études une spécificité de 100%. Les estimations de précision n’ont pas réussi à déceler que le taux de contamination et de résultats indéterminés soit un échec majeur des tests commerciaux. Celui-ci s’étalait de 3% à 36% dans les études visant à une détection directe de la résistance à la rifampicine à partir des échantillons des patients (moyenne 20%).

CONCLUSION: Les essais basés sur les phages nécessitent une amélioration ultérieure pour augmenter les résultats interprétables et réduire les déficiences techniques. Une fois ces problèmes techniques résolus, il faudra déterminer l’impact sur les résultats importants pour le patient ainsi que leur rapport coût-efficacité pour fournir les données nécessaires à une politique d’utilisation élargie.

**RESUMEN**

OBJECTIVO: Actualizar los resultados de un metanálisis publicado previamente sobre los datos científicos existentes en materia de precisión y rendimiento diagnóstico, característicos de las pruebas de tipo comercial o no comercial que utilizan bacteriófagos, en la detección de tuberculosis resistente a rifampicina.

MÉTODOS Y VARIABLES: Se llevó a cabo un examen sistemático y un metanálisis de los estudios sobre la precisión de las pruebas, aplicando un modelo de regresión bifactorial de efectos aleatorios y un modelo jerárquico de estudio de eficacia diagnóstica. Las pruebas analizadas fueron estuches comerciales FASTPlaqueTM, ensayos con el fago reporter de la luciferasa (LRP) y pruebas de amplificación de fagos, elaboradas internamente. Las principales variables de evaluación fueron la sensibilidad y la especificidad de las pruebas de resistencia a rifampicina.

RESULTADOS: Para la actualización de previas búsquedas en las publicaciones científicas, se incluyeron en este metanálisis 31 estudios (con 308 muestras). Las evaluaciones de las pruebas comerciales con bacteriófagos presentaron mayor variabilidad en el cálculo de la sensibilidad (entre 81% y 100%) y la especificidad (entre 73% y 100%) que las evaluaciones de las pruebas de amplificación elaboradas en el laboratorio (intervalo de sensibilidad 88% a 100%; intervalo de especificidad 84% a 100%). Las evaluaciones de las LRP ofrecieron valores de precisión diagnóstica más homogéneos, pues siete de ocho estudios presentaron una sensibilidad de 100% y cuatro de ocho estudios una especificidad de 100%. La evaluación de la precisión no captó el mayor defecto de las pruebas comerciales que consiste en una alta tasa de contaminación y resultados no concluyentes. Estas tasas oscilaron entre 3% y 36% en los estudios de detección de resistencia a rifampicina realizadas directamente en las muestras de los pacientes (promedio 20%).

CONCLUSIÓN: Es necesario perfeccionar las pruebas que utilizan bacteriófagos, a fin de lograr un máximo de resultados interpretables y disminuir las deficiencias técnicas. Una vez resultadas las dificultades técnicas, se debe determinar la repercusión de la aplicación de estas pruebas en variables importantes relacionadas con los pacientes y la rentabilidad de las pruebas, a fin de documentar las políticas de generalización de su uso.