

WORKSHOP REPORT

WHO Tuberculosis Diagnostics Workshop: – Product Development Guidelines –

Cleveland, Ohio, 27 July, 1997.



Sponsored by

World Health Organization

Global Tuberculosis Programme

Executive Summary

The WHO/GTB Tuberculosis Diagnostics Workshop was held on Sunday, 27 July at the Renaissance Hotel in Cleveland Ohio. Approximately 65 participants included basic and clinical scientists, a substantial number [28] of representatives from industry, representatives from U.S. and European regulatory agencies, experts from mycobacteriology laboratories, and representatives from low/middle income countries with direct experience in tuberculosis control programmes. The purpose of the workshop was to develop product performance guidelines for new diagnostics for tuberculosis needed for use in low income countries. These guidelines are intended to assist the diagnostics industry in their efforts to develop useful new products.

Product performance guidelines were prepared for four categories of products:

- 1. Tests to replace AFB microscopy in the diagnosis of smear positive pulmonary tuberculosis,**
- 2. Tests to improve the differential diagnosis of AFB smear negative tuberculosis,**
- 3. Tests to improve/simplify antibiotic susceptibility testing,**
- 4. Tests for recent and/or latent infection for use in prevalence surveys.**

The discussions leading to the formulation of the performance guidelines focused strongly on the potential use of new products in the field, in low income countries. For the tests to diagnose AFB smear positive or smear negative pulmonary tuberculosis, these guidelines emphasize the need for a high degree of specificity, improved sensitivity compared to technologies now in use in the field, speed, simplicity, and cost. Participants noted that the field performance of the current tests is far below optimal. For example, AFB microscopy may detect only 50-70% of the cases of active pulmonary tuberculosis. They also noted that "ideal" tests would be difficult to achieve but that new products were urgently needed. New tests whose performance achieved most, but not all, of the guideline recommendations would still be valuable.

Several issues remain. Further refinement of the product guidelines is needed. Additional issues relating to sample processing, levels of laboratory services in different settings, and the use of new tests in different geographical areas also need to be addressed. WHO plans to organize additional Tuberculosis Diagnostics Workshops to address these issues.

Background and venue

The aim of the Workshop was to formulate minimum useful product performance guidelines to assist industry in the development of valuable new diagnostic products. Industry asked for such assistance and WHO/GTB responded. Product performance guidelines that are established at "ideal" levels will not assist in this goal. Neither will guidelines that are set too low. The workshop sought to establish guidelines to assist in the development of new products that would be both practical to develop and useful in the field. Products whose performance meets or exceeds such guidelines should provide significant improvements in the diagnosis of tuberculosis in low income countries where most of the world's burden of this disease is located.

The WHO/GTB Tuberculosis Diagnostics Workshop was held on Sunday, 27 July at the Renaissance Hotel in Cleveland Ohio. This location was chosen to draw upon the experience of more than fifteen basic/clinical researchers with an interest in the development of new diagnostics for tuberculosis and five researchers and administrators from low/middle income countries where the prevalence of tuberculosis is high. These participants were already in Cleveland to attend the Annual Meeting of Case Western Reserve's Tuberculosis Research Unit. Other participants included 28 representatives from the diagnostics industry, five representatives from U.S. and

European regulatory agencies, experts from mycobacteriology laboratories, and representatives from low/middle income countries with direct experience in tuberculosis control programmes.

The agenda was planned to first give participants an overview of the current status of tuberculosis diagnostics in low income countries together with a synopsis of the needs for new products. Next, a brief summary of a variety of products under development was provided by industry representatives. An overview of some regulatory considerations related to new diagnostics for tuberculosis was then presented. Industry was then given an opportunity to comment on the information and topics thus far discussed. The remainder of the workshop centered on the deliberations and output from four working groups, each focused on a different diagnostic need. The working group discussions were summarized and presented for discussion to all workshop participants. Finally, with industry representatives participating as observers, workshop participants finalized workshop guidelines for the new diagnostic products most needed.

Roderick Roy

Roderick Roy, President of Biotech, opened the meeting by describing how WHO has worked, through the Global Tuberculosis Programme's Tuberculosis Diagnostics initiative, to assist industry in their efforts to develop new diagnostic products for tuberculosis. "This meeting and the manuscript included in the materials provided to participants are examples of this cooperation," said Roy. He went on to note that financial backers from the UK do not acknowledge WHO's assistance as an asset. "They view the WHO's high profile, pragmatic approach to the needs of industry as a drawback," said Roy. "It's a negative for them." They believe that WHO will want something for nothing. "My experience with the Global Tuberculosis Programme is the opposite. It's a realistic, pragmatic approach to the needs of industry," said Roy. They have clear and realistic awareness of the problems that must be tackled and understand that this can only be accomplished through cooperation. Roy said it is "difficult for industry to make decisions that involve major financial risks if we don't know where the goals posts are or what the goals are." The importance of this meeting is to establish those criteria, to look at a new generation of diagnostics to help tackle problems. "Today we have an opportunity here today to put one or two more pieces of the puzzle in place," he concluded.

Paul Nunn

Paul Nunn from WHO discussed the Global Tuberculosis Programme's policy for [GTB] support for basic and clinical research and the reasons behind these policies. He noted that about two years ago, GTB reviewed its TB strategies and redirected its support in these areas.

"It was clear that during the last twenty years, few advances in basic and clinical research were made to help with the practical control of TB in low income countries where most of the world's tuberculosis is located," Nunn said. This includes the basic and clinical research supported by GTB as well as others.

As a result of this review WHO/GTB decided to redirect its support of basic and clinical research to focus on activities that support the research process as opposed to investigator-initiated research proposals. It was decided that:

1. WHO's main role in tuberculosis research should be one of coordination among agencies and institutions with interest in TB research. As part of this coordination, WHO/GTB can provide better access to low income countries where WHO has access and to which researchers or research support agencies may not have such access.
2. WHO/GTB, working with National Tuberculosis Control Programmes, and others in the field, can present the needs, or enable those most closely involved to present their own needs, to the researchers and the research support agencies.

3. The small scale direct funding of investigator-initiated laboratory research programmes is not where WHO's comparative advantage lies. It is for this reason that WHO/GTB has ceased providing this support and directed their resources to other areas like our Tuberculosis Diagnostics Initiative and advocacy for research.

After examination of the current status of tuberculosis research in the fields of diagnostics, drugs, and vaccines, WHO/GTB concluded that the area most likely to yield useful products was diagnostics. The basic science supporting the development of new diagnostics was more advanced when compared to research needed for the development of new drugs [such as research to identify surrogate markers] or vaccines [including research to identify correlates of protection]. "We therefore concluded that the science was most advanced in the area of diagnostics and that it was in this area that useful products would be the first to appear--provided that the necessary coordination could be achieved." We therefore embarked on the Tuberculosis Diagnostic Initiative and gave John Foulds the lead responsibility "This meeting represents an important first step taken under this initiative," Nunn said.

John Foulds

The aim of the WHO/GTB Tuberculosis Diagnostic Initiative is to assist industry in their efforts to develop and test new diagnostic products. Our goal is that such new products will be available for use in the field by the end of the year 2000.

Work on the WHO/GTB Tuberculosis Diagnostic Initiative began with a series of discussions with industry to determine how WHO/GTB could best assist and enable industry in their efforts to develop new tuberculosis diagnostics. "You all have been given a copy of a manuscript prepared by Rick O'Brien and me, describing the needs for, and properties of, tuberculosis diagnostics appropriate for use in low income countries," said Foulds. This meeting is being held in response to the requests from industry to develop product performance guidelines that can be used as targets for development of these needed products.

We next plan to organize a specimen bank consisting of well characterized materials prepared from clinical specimens collected at several locations. The purpose of this specimen bank will be to provide useful materials to assist researchers and manufacturers in their efforts to evaluate new diagnostic products in the laboratory. Products that perform well with these specimens, at least in theory, will be more likely to perform better in the field. As part of the activities associated with the development of this specimen bank, we will be hosting a small meeting in early October to determine the composition of this specimen bank, how the needed specimens should be collected, stored, and distributed. "I urge those of you who have an interest in the organization and development of this bank to contact me," said Foulds.

After new diagnostic products are tested in the laboratory, the next step in product development will certainly involve field trials. WHO/GTB hopes to assist industry in this phase of product evaluation as well. We will offer assistance in the design of appropriate trials and help to establish appropriate sites in a number of international locations to help evaluate the potential value of new products. "We will also work to assist and coordinate industry efforts to obtain regulatory approvals needed to assure the widest use of appropriate new products," said Foulds.

Session I. New tuberculosis diagnostics: The perspective of low income countries

Chair: D. Kibuga

Rapporteur: J. Johnson

Opening Remarks

The presentations in this section are intended to focus on the current status of TB diagnostic methods in developing countries with an emphasis on strengths and weaknesses of current approaches.

Dr. Kibuga, director of the NTP in Nairobi, Kenya, opened the session by describing the needs for new tuberculosis diagnostics from the perspective of low income countries using his country as an example. This, he hoped, would assist the workshop participants in their discussions during the rest of the workshop.

Kibuga told participants that the Kenyan NTP is organized into 60 districts and 8 regions serving a population of 26 million with a median per capita income of 200 USD per year. Kenya is a high TB prevalence country with a prevalence of 100 cases/100,000 population. AFB smear microscopy is the major means of diagnosis. About half of these cases are AFB-smear positives and half AFB-smear negatives. Dr. Kibuga emphasized the need for new products to assist in the diagnosis of these cases and especially emphasized the need for help in the diagnosis of smear negative disease.

Bertie Squire: Tests to replace AFB microscopy in the diagnosis of smear positive pulmonary tuberculosis

Bertie Squire, University of Liverpool, discussed the functioning and effectiveness of AFB microscopy for the diagnosis of tuberculosis in the setting of a low income country's tuberculosis control programme. His presentation was based on research on done in Malawi for six months during 1995. He added that he had just returned in 1997 to find the situation no different.

Squire found that AFB microscopy may miss up to 50% of smear positive cases of tuberculosis due primarily to the requirement for multiple patient visits to the health care facility. He also found that the process is expensive costing about \$14/case treated. He noted the need for new diagnostics and found that simply using a single on-the-spot sputum sample could both increase the number of smear positive cases treated and reduce the cost to about \$2/case treated.

A four step process is used in the evaluation of patients with suspected TB at Bottom Hospital in Lilongwe, Malawi where Squire did his study. Bottom Hospital is a large referral hospital where 70% of the sputum smear positive TB patients are HIV-infected. The identification and later treatment of tuberculosis cases in Malawi begins with patients who present themselves the local hospital. TB suspects are identified based on clinical symptoms especially cough for greater than 3 weeks. These suspects are then referred to the chronic cough room where the collection of sputum specimens is coordinated. Three sputum samples are collected for evaluation: (a) a spot specimen on the first day; (b) an overnight pooled specimen which is brought to the clinic by the patient the next day; and (c) a second spot specimen which is collected when the patient returns to the clinic with the overnight specimen. This process requires two clinic visits to collect three specimens. Patients who return for their results and are found to be sputum smear positive are immediately started on anti-TB treatment. Patients who are smear negative on all three specimens are evaluated by chest X-ray.

The study by Squire and his colleagues examined the frequency and reasons for defaulting between presentation and treatment. During the six months of the study, from a total of approximately 30,000 patients who visited that hospital, 499 suspects were identified and asked to provide an on-the-spot sputum sample. Thirty-three suspects were not able to provide a sputum specimen and did not return. Of the remaining 466 suspects, only 316 provided three sputum specimens and returned to the hospital to collect the results. A total 39 cases were identified as AFB-smear-positive. Of

these, only 28 started treatment. Over 30% of the patients with AFB smear positive disease were not treated.

Squire suggested that the requirement for multiple visits to the hospital limits the number of cases of smear-positive tuberculosis that are both diagnosed and treated. He also noted that a total of 69 patients with at least one AFB-smear positive sputum were identified among the 499 suspects examined during this study. Of these 69 patients, 28 began treatment at the hospital. Another four patients were treated elsewhere. About half [37/69] of the patients who produced at least one AFB smear positive specimen either never returned to collect the results or came back but were not placed on treatment. Squire suggested that the requirements for AFB microscopy actually leads to the loss of about 50% of cases. "The major reason is the multiple visits required to complete diagnosis and to collect the drugs and begin treatment," said Squire. Squire also suggested that the poor performance of AFB microscopy has important economic and social consequences.

Squire calculated that a total of 1274 AFB smears were performed on the 499 suspects leading to the identification of a total of 39 cases with 28/39 cases being treated. The diagnosis thus costs about \$14.10/case treated. Squire also estimated that nearly 140 deaths would result from those cases that were left untreated. He went on to suggest that most patients could be identified using only one smear. Squire also indicated that the mean time from appearance at the hospital until treatment was begun was 18 days during which time the disease will be needlessly spread. He suggested that the difficulties associated with the collection of multiple specimens greatly outweigh the benefits when compared to the collection and immediate analysis of a single sputum.

In the group of 499 suspects examined in his study, Squire calculated that 53 [not 39] would have been identified as AFB-smear positive based upon the first smear. Assuming that the AFB microscopy could have been done on-the-spot, these patients could have all been started upon treatment the same day. This would mean that only 499 smears would be needed [not 1274] and that it would therefore only cost \$2.00/case treated. Squire estimated that 28 [not 140] deaths would result from the cases missed.

"In our setting there is a major priority for the development of an on-the-spot diagnostic from AFB smear positive tuberculosis," said Squire. He also noted that any new diagnostic should take into account the limitations imposed by the lack of laboratory equipment and trained personnel available in many low income countries.

Philip Hopewell: Tests to improve the differential diagnosis of AFB smear negative tuberculosis

Philip Hopewell, University of California, San Francisco, reviewed the differential diagnosis of smear-negative tuberculosis from two important points of view. First was from the epidemiologic or public health perspective. The second was from the point of view of the patient. Hopewell also reviewed a variety of procedures that are currently in use for the diagnosis of tuberculosis.

Hopewell began by discussing the somewhat arbitrary distinction between AFB smear positive and smear negative tuberculosis. He reminded the Workshop participants that the detection threshold for AFB smear microscopy is about 10^4 AFB/ml of sputum. The sputum bacillary load in some patients characterized as either smear positive or smear negative may fluctuate around this threshold. "Patients may be positive at some times and be negative at others," he said. He also noted that, although less infectious than smear-positive tuberculosis, smear negative tuberculosis remains a public health concern.

Based upon these considerations, Hopewell commented on smear status and infectiousness. He noted that, from an epidemiologic viewpoint, the infectiousness of an individual case of tuberculosis is likely to be a continuous function of the number of bacilli found in sputum such that a single patient may, over time, change from AFB smear-negative to weakly smear-positive and fall back to

smear-negative. This cycle may recur over time until, as the disease worsens, the individual becomes more firmly smear-positive. During these cycles, the number of bacilli in the sputum fluctuate above and below the theoretical value of 10^4 /ml needed give an AFB smear positive result. During this period, the infectiousness also varies, depending upon the number of organisms in the sputum. This variation in infectiousness is reflected in [rather imprecise] household contact studies that show that smear-positive, culture positive cases are more infectious than smear-negative culture positive cases which, in turn, are more infectious than smear-negative, culture negative cases. Based on molecular epidemiology, Hopewell estimated that about 15% of all cases of tuberculosis resulting from recent transmission are due to smear-negative cases.

Hopewell discussed how smear negative tuberculosis is recognized in the field. "What seems like a fairly simple and straightforward issue, the differential diagnosis of smear negative tuberculosis, is neither simple nor straightforward," he said. The first step in this process is the collection of an appropriate specimen[s]. Questions arise relating to the number of specimens, the character of the specimens, and the quality of laboratory and/or training of the personnel where the specimen is examined. Where culture can not be done, additional steps that are commonly applied in the diagnosis of smear negative tuberculosis include chest X-ray. If the patients is smear negative, many programs also empirically treat the patient with oral broad spectrum antibiotics targeted against bacterial pathogens for 2 weeks and then re-evaluate the patient. If the patient shows clinical and radiographic improvement, the patient is presumed to have had bacterial pneumonia. If there is no response to anti microbial treatment, other diagnostic possibilities, including tuberculosis, are considered.

In the setting of low income countries as elsewhere, there are a number of factors that influence the diagnosis of smear negative tuberculosis. These factors include the prevalence of tuberculosis in the population, the prevalence of HIV infection, and finally, the prevalence of other infections that may mimic tuberculosis. Additional factors include exposure and risk behaviors such as smoking which may lead to other lung diseases. At UCSF, patients referred to the TB clinic for suspected TB are screened with 3 sputum examinations for smear and culture. Smear and/or culture positive cases are started on anti-TB treatment. Smear and culture negative cases are treated with 2 weeks of broad spectrum antibiotics and undergo repeat clinical and radiographic examination. If there is no response to treatment the patient may have TB, an old arrested TB with so-called fibrotic residuals or another condition such as lung cancer.

Hopewell also referred to a study by Daley and colleagues in Tanzania¹ where patients with suspected tuberculosis and three negative smears underwent bronchoscopy with bronchoalveolar lavage [BAL]. Approximately 75% of these patients proved to be culture positive with BAL as the only source of a positive culture confirming the diagnosis of TB. TB was the most frequent diagnosis made at bronchoscopy followed by much lower numbers of cases of bacterial pneumonia and pneumocystis pneumonia. Tuberculosis was the most common cause of chronic lung disease among smear negative adults in Tanzania, a high TB prevalence country.

These examples illustrate how differential diagnosis of smear negative tuberculosis plays out in different geographic areas given the different factors that influence the array of diseases that occur. Hopewell concluded by again emphasizing the importance of appropriate diagnosis of smear negative tuberculosis, both in terms of public health to identify infectious sources more rapidly, and in terms of individual health, to identify specific diseases that can be treated. "In areas of high prevalence of tuberculosis, the most common disease that occurs in someone with the clinical signs of tuberculosis but has a negative sputum smear is still tuberculosis," Hopewell noted.

¹ Daley, CL, F Mugusi, LL Chen, DM Schmidt, PM Small, E Bearer, E Aris, IM Mtoni, JP Cegielski, G Lallingerr, I Mbaga, and JF Murray. 1996. Pulmonary complications of HIV infection in Dar es Salaam, Tanzania. Role of bronchoscopy and bronchoalveolar lavage. *Am J Respir Crit Care Med* **154**: 105-110.

Leonid Heifets: Tests to improve/simplify antibiotic susceptibility testing

Leonid Heifets, from National Jewish Medical and Research Center, discussed the current status of antibiotic susceptibility testing. He noted that there are now a number of methods available for mycobacterial drug susceptibility testing. However, the choice of the optimal method is not simple. There are four commercial systems using liquid media - BACTEC, MB/BacT, MGIT (B-D), and ESP. Results of antibiotic susceptibility testing using each of these systems has been compared to results obtained by the traditional "gold standard" agar proportion method. Overall the BACTEC 460 system performed the best. But use of the BACTEC 460 system poses additional problems related to the disposal of radioactive waste products. The agreement of the systems using liquid media and agar proportion method data is closest where testing the susceptibility of clinical strains to INH and rifampicin. Susceptibility testing results for other antibiotics, including ethambutol and streptomycin, are often less reliable.

Heifets emphasized that susceptibility testing using the agar proportion method can be performed by the direct technique (*i.e.* direct inoculation of the drug susceptibility testing plates from the clinical specimen) on smear positive sputum specimens which gives results to the clinician within 3-4 weeks. Smear-negative specimens take longer as they must first be cultured. The organisms that appear after this initial period of growth are then used for susceptibility testing (indirect method). The agar proportion method is simple and inexpensive to perform. The direct technique may fail due to contamination of the specimen or a failure to obtain adequate growth to interpret the results properly. In such cases, an indirect test, using organisms first grown in liquid medium is an important back-up.

“The key issue in obtaining rapid drug susceptibility results is not technology but the proper organization of the laboratory service,” said Heifets. It is possible to get results within 3 weeks on most of the smear-positive specimens using either the direct agar proportion or BACTEC methods. In middle income countries the best strategy may be to use the direct agar proportion method on smear-positive clinical specimens [especially those isolated from smear-positive patients] with backup testing in liquid media, when possible. For low income countries where alternative drugs may be available, determination of antibiotic susceptibility is needed.

Patrick Brennan: Tests for Infection (Prevalence Surveys) and Current Technology as Applied to Low Income Countries

Patrick Brennan from Colorado State University reviewed the current use of skin test reagents and summarized research efforts directed toward their improvement.

Brennan noted that the purified protein derivative (PPD) skin test is the only currently available method for screening populations at risk for infection with *M. tuberculosis*. It is inexpensive, familiar to health care workers and requires 2 visits. It measures cell-mediated rather than humoral immune responses and is widely recognized as the universal "gold standard" method to detect tuberculous infection. PPD skin tests can be difficult to interpret in many instances. These include individuals recently vaccinated with BCG and those living in areas where infection with cross-reactive environmental mycobacteria is frequent. Such persons may react with PPD in the absence of infection with *M. tuberculosis*. Alternatively, persons who are actually infected may not react to PPD. These include the aged, and immunosuppressed patient populations such as HIV-infected patients. PPD testing also is subject to many difficulties in administering the test and variability in reading the test. Brennan summarized these data with the observation that the PPD test may provide unreliable results in many instances. He pointed out that up to 20% of cases of true infection may be associated with classically interpreted negative PPD (< 10 mm diameter) skin tests (Grange, 1994; Bass, 1990; CDC, 1994). New, more reliable, tests are needed.

The first tuberculins were fluid medium from *M. tuberculosis* cultures initially called "lymph" and then later "tuberculin" by Koch. Early tuberculins included emulsions of killed bacteria, aqueous

extracts of pulverized bacteria, and filtrates from the medium in which bacilli had been grown. Von Pirquet introduced a heated and 10-fold concentrated glycerine-peptone broth from 6-8 week old *M. tuberculosis* cultures. When applied to scarified skin in a tuberculous child - "A papule of 5 to 20 mm in diameter appeared at the site and then gradually disappeared over a period of 8 to 10 days or longer." Ammonium sulfate precipitated preparations were introduced by Seibert in the late 1920s-1940s. In the modern era Affronti, Chaperas, Janicki, Daniel, Someya, Yamamura and others have continued the efforts to identify better tuberculins.

Several solutions have been proposed to improve or replace PPD skin testing. The first is to continue the use of skin testing but develop better antigens. New skin testing antigens are being developed at Colorado State University under an NIH contract "TB Research Materials and Vaccine Testing". Information about this program and available materials can be obtained through the World Wide Web at <http://cumbs.colostate.edu/microbiology/tb/top.htm>. Another approach has been to replace skin testing with a simple whole blood assay to measure IFN γ production after overnight stimulation with PPD or improved mycobacterial antigens.

Two approaches have been pursued to develop better skin testing materials. The first involved purification of undenatured culture filtrate or bacterial extracts. Initial work was done by Seibert to fractionate old tuberculin into smaller 2000-4000 molecular weight fractions before World War II. After WWII she fractionated PPD-S into protein fractions A, B, and C. Other collaborative work led to the characterization of US-Japan antigens 1-7. The second approach has been to fractionate PPD itself. This is difficult to do due to the large number of recognizable antigens in PPD. Someya and Yamamura, Nagai, and Kuwabera also purified small 5-10 kDa fractions which were antigenic.

"Future research should focus on a systematic approach to complement and extend current efforts to fractionate PPD," said Brennan. Initial results at CSU demonstrated that culture filtrate proteins, cytosolic, and SDS-soluble cell wall proteins (SCWP) elicited significant DTH responses in tuberculous guinea pigs. Subcellular protein fractions are separated by isoelectric focusing into 20 fractions. Each of these fractions is further separated by SDS-PAGE. Thirty fractions are separated from each of the original 20 fractions yielding a total of 500-600 fractions for each of the subcellular protein preparations. Each fraction contains one to three proteins. This procedure has been completed for the culture filtrate proteins and the cytosol preparations. All fractions will be tested using *in vitro* T-cell proliferation and IFN- γ assays against immune T cells from mice. Active fractions from the above assays will be tested in similar *in vitro* systems using human T cells from PPD-positive individuals. Fractions that are active in these *in vitro* assays will be tested by DTH skin testing in guinea pigs. Proteins of each active fraction will be further characterized. Cell wall-associated proteins have been found to be highly active and specific.

Future steps to develop better skin test materials include several ongoing and new initiatives. "We have constructed a manufacturing/pilot plant facility at Colorado State University for the production of new skin test antigens," said Brennan. He added that this facility is suitable for the production of sufficient antigens for a phase I study in Colorado and a limited phase II study to be considered by the FDA in the fall of 1997. Brennan suggested that both the WHO and industry need to become more involved in the development of these materials as they become ready for further evaluation in phase III field trials. Finally, the application of T cell epitope prediction using the Epimatrix software developed by Ann DeGroot to the whole *M. tuberculosis* genome sequence opens up exciting possibilities.

Brennan concluded with a challenge for the Workshop participants: "Important issues for this Workshop include the development of new skin testing antigens and exploring the potential uses of IFN- γ whole blood stimulation assays."

Session II: New products under development

Chair: Kathy Eisenach

Rapporteur: Mark Perkins

The purpose of this session was to inform the workshop participants of some representative new diagnostic tests that are currently under development. A number of other industry participants indicated their willingness to participate in this session but time limited the number of presentations. Presenters were chosen on a first-come basis. Presenters were asked to describe the basis of their test, its intended use, and its developmental status.

During discussion of these presentations a question arose as to the ability of a U.S.-based company to market a test outside [export] of the U.S. that had not received FDA approval. Reliable, but not official, information suggests that yes, this is possible provided that the country that will be importing the product [device] provides an authorization to import signed by a responsible official from the importing country. All products that do not have clearance or approval from FDA get the following included in their export letter: "The subject device[s] is[are] not legally in commercial distribution in the United States. Therefore, shipments of this device in the U.S. would have to be conducted in accordance with the Investigational Device Exemption [IDE] regulation (21 CFR, Part 812)."

Company/ Product	How Test Will be Used	Basis of Test	Status of Development
Roche/Amplicor MTB	Detection & speciation (replace smear/culture)	Nucleic Acid Amplification	Marketed
Bio-tec/Biophage	Detection (replace smear/culture)	Phage susceptibility	Clinical trials
CSL/Quantiferon	Detection (replace smear/culture)	Detect Gamma Interferon in blood sample in response to PPD	Marketed Australia/New Zealand
DACO/FISH	Improved microscopy	<i>in situ</i> hybridization using peptide nucleic acid [PNA] probes	Preclinical trials
Vysis/DNA Probe	Detection & speciation (replace smear)	Nucleic Acid hybridization <i>in situ</i>	Preclinical trials this year
Kreatech/EIA-TB	Detection (replace smear)	Elisa based	Preclinical trials
Immunodiagnosics Immuno-TB	Detection (replace smear)	Elisa based	Laboratory
ANDA/Gelification	Detection	1. Qualification of blood 2. Elisa	Laboratory

Bio-tec

Roderick Roy, from Bio-tec, described their BioPhage FAST Plaque Assay. This is a direct test for live *M. tuberculosis* organisms in sputum that can be modified to determine antibiotic susceptibility as well. An excess of specific bacteriophage are added processed sputum specimens containing [or not] viable *M. tuberculosis* cells.. After allowing time for adsorption and injection of the bacteriophage DNA, unabsorbed bacteriophage are inactivated using a viracidal agent. Next, a rapidly growing mycobacterial host for the specific bacteriophage [i.e. the indicator strain] is added and the mixture plated on solid media. The presence of viable *M. tuberculosis* bacilli in the sputum is reflected in the number of plaques on a lawn of growth of the indicator strain. This test can quantitatively detect the presence of *M. tuberculosis* in sputum within 24 hours. No prior growth is required. The test is compatible with routine sputum decontamination, can be

accomplished using existing technical skills, and requires little instrumentation [BL-2 hood and incubator]. The test is currently being evaluated in the field where initial results have been highly successful.

CSL

Gavin Horrigan, CSL, described their QuantiFERON-TB test to detect gamma-interferon produced by cells in blood samples after stimulation with PPD or other antigens. This test requires 24 hours and equipment including an incubator, BL-2 hood, and a centrifuge. The test is currently marketed in Australia and New Zealand. It is also now being evaluated in the U.S. in cooperation with the CDC. Field trials in Australia and New Zealand indicate that the test is 90% sensitive and 95%-98% specific. Improvements are underway to help eliminate BCG vaccination as a source of interference by using a selection of recombinant antigens.

DAKO

Henrik Stender, DAKO, reviewed the status of their PNA probe. The basis for this test is a peptide nucleic acid [PNA] probe that is designed to bind to specific nucleic acid sequences. These probes, which can have greater specificity than DNA probes, have been prepared to recognize rRNA from organisms of the *M. tuberculosis* complex. The hybridization is performed directly on sputum smears *in situ* using PNA labeled with a fluor and the presence of the organisms is detected using a fluorescence microscope. It can be accomplished in about 2 hours and is currently being evaluated in preclinical trials. The test is intended to replace AFB microscopy in the screening of suspects for AFB smear positive tuberculosis and as a replacement for culture in the differential diagnosis of smear negative disease. This test requires both a water bath for the hybridization step and fluorescent microscope to observe the results.

Vysis

Rober Koska, from Vysis [an Amoco Company] described their FISH test for tuberculosis that is currently under development. This test is a fluoromicroscopy test using fluorescent *in situ* hybridization [FISH]. The assay is a dual color test designed to screen tuberculosis suspects for active disease. It is being developed to differentiate non-tuberculosis mycobacteria from *M. tuberculosis* using as many as seven fluorochromes. The assay is rapid [approximately 2 hours], requires both a fluorescent microscope [40X] and a water bath.

KREATECH

Hendrik Houthoff described a serologic test for tuberculosis under development at KREATECH. This test is based upon an enzyme immunoassay [EIA] using KREATECH's patented antigen bar coding [ABC] methodology. It detects the presence of specific IgA in sera from patients using a combination of antigens. Preliminary evaluations of the KREATECH tuberculosis EIA kit has shown sensitivity above 80% and specificity of 80-90%. They are working toward improving the specificity with support from the European Commission. The test will be prepared for use for population screening, for patient testing, and for monitoring the progress of therapy. This test is currently ready for market in a non-FDA approved format.

Immunodiagnosics

Jay Raina, from ImmunoDiagnosics, reviewed the properties of their Immunotest TB diagnostic test kit. This test is designed to detect antibodies directed toward a combination of antigens prepared from *M. tuberculosis* and blended to optimize the signal/noise ratio. The test utilizes an indirect solid phase immunoassay technology for the semi-quantitative detection of *M. tuberculosis* antibodies. This is a rapid [less than ten minutes] test that has been shown to be more than 85% sensitive and 85% specific in screening *M. tuberculosis* suspects using either whole blood, serum, or plasma. The test is currently marketed outside the U.S. in a non-FDA approved format.

ANDA

Roland Maes, ANDA Biologicals, described their gelification test and an enzyme immunoassay based upon the detection of antibodies the A60 antigen of *M. tuberculosis* in the sera prepared from tuberculosis suspects.

The gelification test is based upon the observation that the whole blood from tuberculosis patients contains a relative excess of fibrinogen and, after mixing with an equal volume of 1.25% glutaraldehyde, will coagulate within 10 minutes. Maes was careful to point out that the rapid gelification was not, in itself, a test for tuberculosis as a number of other respiratory conditions may lead to the same result. However, he noted that the test was both simple to perform and inexpensive. He said the gelification test provided presumptive

evidence that could be confirmed using other tests. He then described the ANDA enzyme immunoassay for the detection of antibodies to the A60 antigen prepared from *M. tuberculosis*. This is a rapid test that detects IgG, IgM, and IgA antibodies. The test is simple to administer and requires only 15 minutes to complete. This test is presently on the market in a non-FDA approved format.

Research Report: Robert Wallis

Robert Wallis, Case Western Reserve University, briefly discussed some of his recent research results. He suggested that these data could lead to the development of an appropriate new diagnostic to replace AFB microscopy in screening TB suspects. Wallis has measured the amount of *M. tuberculosis* 85B [alpha] antigen in sputum directly by ELISA. He finds that the concentration of this antigen decreases within the first 7-14 days of treatment in those patients whose cultures become negative within the first month of treatment. He said that, with further work, a test based on this direct determination may be valuable as a replacement for AFB microscopy in the screening of TB suspects. Such a test may "also have the potential to identify patients who may not require a full six months treatment," he said.

Session III: Mark Goldberger

Regulatory issues for new devices

Chair: Paul Klatser

Rapporteur: Mark Perkins

Mark Goldberger, from the FDA, discussed some issues relating to the approval process for new diagnostic tests for tuberculosis. Dr. Goldberger made it plain that he was attending this workshop to trade ideas, not to establish FDA criteria nor to make official FDA proclamations. He noted that from the FDA perspective, new diagnostic tests are called “devices.” “We are here to discuss product development guidelines for the new devices needed for the diagnosis of tuberculosis in low income countries,” he said. These new products need not be ideal to be useful. Goldberger listed a number of attributes that new tests might be expected to demonstrate improvements with respect to, for example, AFB microscopy. These performance attributes include:

Test Attributes:

- 1. Sensitivity: *i.e.* more than 70%**
- 2. Specificity: *i.e.* more than 98%**
- 3. Patient population: *i.e.* suspects to be screened such as smear positive, smear negative**
- 4. Number of patients visits to the health facility before a diagnosis is received**
- 5. Source of specimen: *i.e.* sputum, blood, serum, urine**
- 6. Simplicity; *i.e.* easily administered by a technician requiring little or no training**
- 7. Equipment and reagents required: *i.e.* requiring no equipment with all reagents included with the test**
- 8. Stability: *i.e.* a shelf life of two or more years**
- 9. Robust: *i.e.* requiring no refrigeration**
- 10. Cost: *i.e.* No more expensive than current test**

Goldberger said “If anyone were to ask, as we are doing here today, what would you like in terms of a new test, I think we would all say all of them.” “Ideally,” he said, “we would like to optimize all of these. But we find that there are interrelationships and tradeoff among all of them.” He noted, for example, that molecular tests have increased sensitivity compared to AFB microscopy, and speed with respect to culture, but these tests are more expensive. “The issues is that there are clear interrelationships and tradeoffs, among a number of these attributes” he said. New tests can be better, faster, or even easier but they will likely be more expensive.

Goldberger likened the process of development of new diagnostics tests to that of buying a new car. Everyone would like a car that gets excellent gas mileage, has outstanding performance characteristics, is completely safe, and is easily affordable. “Everyone would like a car like that but they realize that they will most likely have to compromise, because such a car doesn’t exist” he said. But, if we are willing to get, say, two or three out of four, we should be able to get a satisfactory new car. Goldberger then suggested that we can similarly achieve many of the attributes of a new diagnostic test for tuberculosis but we most likely will not be able to get everything we want.

This meeting brings together both the potential users of new diagnostic tests and the manufacturer and provides a unique opportunity. “The critical thing from this meeting is the opportunity for the potential users to describe their needs to the manufacturers.” These needs include which products are most needed and a description of how these products, when available, would likely be used in the field. Next, we have here today an opportunity for the manufacturers, in turn, to describe the

products [and their performance characteristics] that might become available within the next few years based upon the available technology. Obviously, there will eventually have to be some prioritization of needs and performance with respect to the ability of technology to respond to these needs. It is likely that more than one test will be developed and that different tests may fulfill different needs. For example, one test may be more appropriate for use in settings where only limited laboratory and/or human resources are available whereas another test may be designed for situations where laboratories are better equipped and technicians are better trained. The specific use for each product would then depend upon the particular properties of the test.

When we talk about what is currently possible, we need to ask how much data is needed to know to make this determination. It is one thing to have information from the research microbiology laboratory, another to have information from clinical trials. These questions relating to performance and intended use of new diagnostic products will need to be addressed for each new product. It is the information from actual field trials that will have the most weight. This is true both for the approval process and for the decision as to the appropriate use for any new product. Goldberger went on to describe how the behavior of any new test would be evaluated after a field trial.

The evaluation of new tests for the diagnosis of tuberculosis include considerations of sensitivity, specificity, and positive predictive value. To assist in the determination of these key performance parameters, data from field trials are usually summarized in a two-by-two table. To illustrate this, Goldberger constructed such a table for a new diagnostic ["D"] where D⁺ refers to a positive result and D⁻ refers to a negative result for this test. These results are then compared to results on the same specimens achieved by the "gold standard" test. For the purposes of this discussion we can consider the diagnostic test "D" to refer to AFB microscopy and the "gold standard" as culture.

	Smear ⁺	Smear ⁻	Total
Culture ⁺	a	b	a+b
Culture ⁻	c	d	c+d
Total	a+c	b+d	

$$\text{Sensitivity} = a/a+c$$

$$\text{Specificity} = d/b+d$$

$$\text{PPV} = a/a+b$$

$$\text{NPV} = d/c+d$$

The numbers [a, b, c, and d] come directly from the results of a trial. It is critical to remember that the power to demonstrate the sensitivity depends entirely upon the absolute number of positive specimens [a+b] in the sample population. Thus, for a trial involving 500 patients conducted in an area with a prevalence of 10%, predictions of sensitivity will be based on a total of 50 cases [n=50]. The size of the confidence interval is related to the square root of the number of cases. Thus for a trials where the sensitivity turned out to be say, 80%, the 95% confidence interval would be:

Samples	Range
25	63%-95%
50	70%-91%
100	72%-88%

“At some point it becomes worthwhile to look at absolute numbers,” Goldberger said. Using as an example a test that is 98% specific and 70% sensitive that is applied to 1000 tuberculosis suspects where the prevalence may range from 2% to 20%, Goldberger noted that different strategies would likely be used to evaluate a test, depending upon the conditions under which the test would be employed. Thus, measures to improve the tests in the example where Sm = ZN smear and TB = culture.

Prevalence				
	2%	5%	10%	20%
Sm⁺, TB⁺ [True positives]	14	35	70	140
Sm⁺, TB⁻ [false positives]	20	19	18	16
Sm⁻, TB⁺ [False negative]	6	15	30	60
Sm⁻, TB⁻ [True negatives]	960	931	882	784

In low prevalence areas, of course, specificity is VERY important, as there are a large number of true negative patients being smeared. In the above example at 2% prevalence, there are, in fact, equal numbers of false positive cases [20] and positives [14 plus 6=20]. In a high-prevalence area, you might be willing to compromise more on specificity due to the lower numbers of non-infected patients being screened and the relatively higher numbers of cases “missed” [*i.e.* false negatives].

Interaction of prevalence and specificity on PPV [At a sensitivity of 95%]

Prevalence					
Specificity	1%	3%	5%	10%	20%
95%	16	37	50	68	83
96%	19	42	56	73	86
98%	32	59	71	84	92
99%	49	75	83	91	96
99.5%	66	85	91	95	98

Goldberger discussed how sensitivity, specificity, and prevalence interact in the calculation of positive predictive value for a test. But Goldberger warned that when we use these numbers, we

need to remember how comfortable we actually are that a sensitivity or specificity is EXACTLY as reported. In fact, each of these numbers will be reported as a range to include the standard deviation or 95% confidence limits.

It is not enough to have the technology, we need data from sites where the test might ultimately be used.

It is important to remember that your power to show sensitivity depends entirely on the number of positives. So even if you do a 500 patient trial, if there are only 50 patients with the disease, you are basing sensitivity predictions on the basis of $n=50$

Finally, Goldberger summarized some of the important attributes of a well planned field trial. He said such a trial may be best if it is:

- 1. Multicenter**
- 2. Sufficient number of TB+ and neg. patients**
- 3. Limit the number of duplicate samples/pt**
- 4. Standardize confirmatory testing**
- 5. Integrated clinical characteristics**
- 6. Include culture-, clinical TB (as many as 15-20% in the us)**

Industry concerns: a workshop discussion period

Salmon Siddiqi, from Becton-Dickinson, reviewed some of the industry considerations related to the development of new diagnostics for tuberculosis. He indicated that industry recognizes the need to develop diagnostics appropriate for use in low and middle income countries where most of the world's burden of disease is located. He added that there remains a number of areas where WHO can assist in this process. "Industry needs guidelines," he said. Siddiqi then made a list of areas where questions may exist.

Guidelines for industry

1. Type and format of test
2. Direct diagnosis?
3. Patient population?
 - a. Smear positive
 - b. Smear negative
4. Intended use?
 - a. Replacement of smear
 - b. Follow-up
 - c. Differential diagnosis
 - d. Replace culture
5. Antibiotic susceptibility
2. Expected sensitivity and specificity
3. Stability and shelf life
 - a. Attention to robustness
4. Packaging
5. Pricing

Siddiqi said that a number of these either topics were either on the agenda for this Workshop or had been addressed in the draft manuscript by Foulds and O'Brien that is included in our Workshop folder.

Next, Siddiqi asked what, exactly, would WHO be able to do to help with the development, testing, deployment, and use of new diagnostics for tuberculosis. Some of these overlap with the items listed above.

1. Development
2. Field trials
3. Approval process
4. Introduction & use of device/reagent - how to do it?
5. Scope of test use in different geographic areas
6. Extent of worldwide market
7. Pricing issues
8. Regulatory issues
9. Other specifications

Siddiqi next asked some questions about the performance characteristics of a new diagnostic whose intended use would be to replace AFB microscopy. Can a new test be useful if it is less than ideal? If so, which performance characteristics are most important?

1. Improve the smear in terms of sensitivity or simplicity

2. Improve the smear in terms of ability to distinguish *M. tuberculosis* and non-tuberculous mycobacteria
3. Could a serologic test replace smear?
4. Must a test to replace AFB microscopy in the screening of pulmonary tuberculosis suspects also replace AFB microscopy's function in follow-up of treatment of smear positive cases?
5. Must define any test aimed at replacing smear in terms of its sensitivity in smear-negative and smear-positive samples. A replacement must be highly sensitive in smear-positive patients.

Siddiqi expressed the hope that at least some of these questions could be answered during this Workshop.

Henrik Stender, DAKO, noted that the aim of this Workshop is to establish product performance priorities. He noted that minimum requirements for products are what is needed. "These must be defined," he said.

Sample preparation has not been discussed, but should be. Improving sample processing would improve the sensitivity of all tests. For example cytocentrifugation.

Siddiqi stressed that HOW a test is introduced is important. Foulds discussed the role of the WHO in this process. He suggested that the product performance guidelines that we are seeking to define during this Workshop could provide the WHO with appropriate standards by which new tests could be evaluated.

Nunn described three areas in which WHO may be able to assist the diagnostics industry in the development, testing, and approval needed for new diagnostics. These are:

1. Science
 - encourage the development and lab-based trials of new tests
 - field trial organization
 - decision and advice about specific tests: need specimen bank
2. Regulation
 - the WHO has some experience with drug combinations and could help with approach to regulatory bodies
3. Interaction with industry
 - changes in the coming years may help define the role of WHO in industry advice and regulatory help
 - comment from industry that there is a real need for specimen banks to help develop these assays.

Working Group I: Tests to replace AFB microscopy in the screening of tuberculosis suspects

Rapporteurs: Larry Geiter and Don Cave

The group emphasized that a positive result from a test to detect smear positive cases should be adequate to initiate a regimen of short-course chemotherapy.

During discussion the following points were raised:

Specificity

Overall, the discussion relating to the specificity [98-99%] of AFB microscopy occupied center stage. This high specificity is based on the observation that few patients in areas with high rates of disease have illnesses other than tuberculosis that result in AFB-positive sputum. It will be difficult to replace AFB microscopy in the diagnosis of tuberculosis for that reason.

High specificity [95% to 99%] is needed to avoid both the costs and potential adverse effects associated with unnecessary treatment. In many countries resources are already severely constrained. But the current test [rated 98% to 99% specific] may well miss as many as 33% to 50% of AFB smear positive cases due to difficulties associated with both the test itself and how it is employed [see Squire's presentation]. These missed cases can be expected to have a significant impact on future transmissions. Tony Catanzaro, UC San Diego, suggested that we need a simple model to compare and contrast the impact of cases missed on increased transmission and to estimate the number of false positive cases inappropriately treated.

Sensitivity

Henrik Stender, DAKO A/S, Denmark, noted that the diagnostic sensitivity of a test is a combination of the analytical sensitivity of the test, the number of bacilli in the specimen, and the recommendations for use. In the case of AFB microscopy, the analytical sensitivity is 100% [for smear-positive specimens that result in at least one bacillus/slide]. However in combination with the recommendations for use, the actual diagnostic sensitivity may be as low as 50%.

Tiered tests

It was pointed out that a system of two tests could be practical and acceptable, i.e., one test with a very high sensitivity and a second to improve on the specificity.

Speed

There was broad consensus that a one-day turn around for results was very important. However, on discussion it was also agreed that the three key features, 95% to 99% specificity, sensitivity equal to or greater than three smears, and same day results, were all important but that a test with two of these three features could also be both a valuable new tools to assist in the diagnosis of tuberculosis and acceptable for use in many locations.

Follow-up

While not a requirement, those participants most closely involved with tuberculosis treatment and control programmes emphasized that it would also be useful if the same test used for diagnosis could also be used to measure patient response to therapy or therapy outcome. Without this property, no test can truly "replace" AFB microscopy.

Summary: Product Performance Guidelines

A test to replace AFB microscopy

- 1. Sensitivity equivalent to three sputum smears [50%-70%].**
- 2. Test result can lead to decision to treat.**
- 3. Specificity = 95%-99% as compared to culture.**
- 4. For use in areas of high prevalence [at least 10%].**
- 5. Source of clinical specimen = blood [finger stick] urine sputum.**
- 6. Time = Approximately 1-4 hours for processing and 10-20 minutes technician time [same day results].**
- 7. Quantitative or at least semi-quantitative results.**
- 8. Equipment equivalent to light microscope and biosafety hood.**

Working Group II: Tests to facilitate the differential diagnosis of AFB smear negative tuberculosis

Rapporteur Larry Geiter

Scenario for use of new products for the differential diagnosis of smear negative tuberculosis:

Where, in the scheme of the differential diagnosis of AFB smear negative tuberculosis are new products most likely to be used? Hopewell described some possible scenarios. Patients who present themselves at a health care facility with a chronic cough [three weeks] will be evaluated using a number of procedures including AFB microscopy, chest X-ray [CXR] and/or a short course of treatment with a broad spectrum antibiotic. Those with abnormal CXR and/or no response to the antibiotic need to be evaluated further for possible pulmonary tuberculosis. Normally, this evaluation would include sputum culture. It is at this point that a new diagnostic for the diagnosis of AFB smear negative disease would most likely be used. A more useful test would replace AFB-microscopy in the earliest stages of patient evaluation. "Ideally, this new test would be one that could replace AFB microscopy in early this scheme rather than culture at the end," said Hopewell.

AFB smear-negative cases of pulmonary tuberculosis account for about 30% of pulmonary cases and 50% of all cases. Individuals who have already been screened for TB have a lower prior probability of a TB diagnosis than a group of un-screened individuals. For field use, a specificity of 95% is acceptable.

Due to the urgent need, tradeoffs among sensitivity, specificity, speed, ease of use, and the robustness of the test may need to be accepted. Thus, a very robust test appropriate for field use in low income countries may be useful even with performance characteristics than outlined here.

During discussion the following points came out:

Specificity

The desired range for specificity of 95% to 99% expressed during the Working Group discussions for specificity were not derived in an objective manner but were "guestimates". The group stated that specificity was very important but that higher levels may not be practical in field application. They also noted that 95% specificity was much better than what is now achieved.

Specimen

For children, a test that works on specimens other than sputum , [e.g., blood, urine, BAL, aspirates] would be highly valued. Overall lower performance [specificity and positive predictive value] may be acceptable for children since they generally tolerate treatment very well [*i.e.* fewer adverse effects than adults].

Prevalence

A prevalence of 10% was assumed although it was noted that smear-negative patients with sign and symptoms of tuberculosis have a lower prior probability of a TB diagnosis than unscreened patients.

Summary: Product Performance Guidelines

A test to improve the differential diagnosis of AFB smear-negative cases of pulmonary tuberculosis

- 1. Specificity = 95%-99% as compared to culture.**
- 2. Sensitivity equivalent to 80% culture.**
- 3. Prevalence assumed to be 10%.**
- 4. Any test should not be affected by HIV status.**
- 5. Positive and easy to use tests for bronchitis and/or bacterial pneumonia may be very useful in the differential diagnosis of smear negative patients.**
- 6. The test should be simple and require little or no specimen processing,**
- 7. Same day results**
- 8. The cost should be comparable to costs currently incurred for diagnosis of smear negative including a chest radiograph and culture.**
- 9. Test for extra pulmonary disease would be highly valued.**

Working Group III - Antibiotic susceptibility tests

Rapporteurs: Lucy DesJardin and Rick O'Brien

What is the role of routine drug susceptibility testing?

How important is drug susceptibility testing in the context of national tuberculosis control programs in low income countries? Is susceptibility testing important for individual patient care or surveillance. Participants agreed that it would be important for surveillance purposes especially since pockets of infection with drug resistance are suspected to exist in developing countries. IUATLD maintains that testing should be used for surveillance purposes only. The cost, level of technology required for testing, and lack of alternative drugs make routine testing a practice of little value and questionable ethics.

Dr. Tyagi thought that drug testing for patient care was needed in India but the cost was prohibitive. Most felt that because of the increase in drug resistance, an attempt should be made to test on an individual patient care basis. Heifets noted that this is especially true for newly diagnosed patients. All agreed that it was important to make routine susceptibility testing possible in low income countries. This need become urgent in areas with high or increasing rates of initial drug resistance.

What is the need for drug susceptibility testing?

The need for new diagnostic products to evaluate antibiotic susceptibility is not as great as, for example, the need to improve the diagnosis of AFB smear positive disease.

For surveillance studies, the standard methods are quite acceptable in terms of cost, speed, and complexity.

For patient care, most patients with isoniazid-resistant organisms can be treated successfully with standard short-course regimens that include four drugs. However, this is not the case for patients with rifampicin-resistant disease, especially if concomitant resistance to isoniazid is present. With the growing problem of rifampicin-resistance in some countries, an inexpensive test to identify rifampicin resistance would be useful. Participants noted that most rifampicin-resistant strains have a mutation in a small segment of the genome. This permits a rapid determination of resistance by molecular methods.

Test design: Direct or culture-based

This question generated the most debate within the group. Dr. Heifets strongly supported agar-based plating of cultures for direct drug susceptibility testing. He said this procedure is "affordable, accurate, and does not require advanced equipment." He also noted that "in the U.S., the cost of managing cases with undetected drug-resistance [leading to multi-drug resistance] was much higher than any expense for susceptibility testing." Some other participants did not support agar-based plating for susceptibility testing. Dr. Tyagi pointed out that India has been unsuccessful in implementing culture based systems. "The culture of organisms for long periods of time under sterile conditions is challenging," she said. In some countries, the requirements of advanced equipment, such as centrifuges, incubators, and continuous power makes culture based methods unreachable. A core reference laboratory was also considered impractical where infrastructure including reliable transportation was lacking.

A susceptibility test that did not require isolation and growth of a pure culture would be of much greater value. For example, Bio-tec's BioPhage system does not require propagation of pure culture but does require reliable electric power, incubators, and biosafety hoods. Heifets pointed out that "any test without cultivation would likely require no less technical skill than standard [microbiology] laboratory procedures and may well require more sophisticated laboratory equipment." He ridiculed

the notion of performing any laboratory test [no matter how simple] with a potentially infectious specimen, without a properly equipped laboratory.

Which drug susceptibility tests are most important

A test for RIF and INH resistance would be of greater clinical desirability, however, a test devised for RIF alone that did not require culture of the organisms, would be worthwhile. Since the mutations which lead to RIF resistance are localized to the *rpoB* gene it might be technically possible to devise such a test.

The reason for testing for RIF alone would be that RIF resistance is the only indication for a clinician to deviate from the standard 4 drug regimen. Even though it was felt that INH resistance occurred at a greater frequency, resistance to INH would not change the treatment a particular patient was given. The caveat to this strategy was that in some areas only two drugs (INH and RIF) are routinely given for treatment of tuberculosis. If the patient was resistant to INH then this would impact the type of drugs prescribed. Also, if a patient was multiply drug resistant then the clinician might elect to withhold treatment entirely. This would be done to prevent development of rifampicin resistance and because effective chemotherapy for that particular patient might not be possible with the drugs available.

Time from sample collection to susceptibility profile

It was assumed that the diagnosis of *M. tuberculosis* infection would be made prior to and separate from susceptibility testing, so patients would know that they were infected with *M. TB* and would be returning to the clinic for treatment. In that context, it was generally agreed that results within a 7 to 14 day time span would be acceptable. There was little benefit to having results within one day.

Summary: Product Performance Guidelines

Antibiotic susceptibility tests

- 1. Results available within 7 to 14 days.**
- 2. Robust-- in reagent stability and in efficiency of readout from samples (i.e. able to make a determination of some sort on 90-95% of samples submitted for testing).**
- 3. Requires only minimal laboratory equipment. Preferably culture of isolated organisms not required.**
- 4. Importance RIF>>INH>>EMB**
- 5. A test which determined RIF resistance alone would be worthwhile.**
- 6. For use on smear positive samples.**
- 7. Greater than 98% agreement compared to agar-based methods.**
- 8. Able to detect 1% resistant organisms in patient specimen.**
- 9. Requires 10 to 15 minutes of hands-on time by a technician without special skills or background in laboratory techniques in microbiology.**
- 10. Low cost (less than US\$ 10/test).**
- 11. Recommended for surveillance and patient treatment depending upon local availability of alternate treatment regimens**

Working Group IV: Test to enable detection of recent/latent infections

Rapporteurs: Chris Whalen/Rick O'Brien

The possible uses of a better test for infection are several: 1) improved tuberculosis control by understanding better tuberculosis transmission and proposing interventions to decrease transmission; 2) assessment of prevalence of infection as measure of tuberculosis transmission in a population; 3) detection of 'high risk persons who might be considered for preventive therapy; and 4) research studies (NOS). For these purposes, the PPD-tuberculin skin test is generally satisfactory, except for failing to distinguish recent from remote infection, information which might assist in decisions about preventive therapy for some persons.

In addition to providing for a way of distinguishing recent from remote infection, desirable characteristics of a new test include providing for distinguishing among sensitization by BCG, MOTT infection and infection from *M. tuberculosis*, quantifying the risk of disease given infection, and requiring only a single patient encounter for testing. Obviously, issues of cost are important. Other considerations include the target population for intended use, e.g., developing vs. industrialized countries, evaluation of individuals in high-risk groups (including contacts) vs. prevalence surveys of infection, and patient-specific factors which may relate to test reactivity (e.g., MOTT exposure, genetic factors).

Possible test formats include antibody detection systems which can require little equipment or training for administration, with the possibility of improvement in test specificity through the use of several purified antigens. Tests based on the determination of cytokine production may have greater specificity than the PPD skin test but may give positive results earlier in the course of infection. In addition, these tests will require specialized equipment, making their application in low-income countries impractical. Finally, skin testing may be improved by the administration of a 'cocktail' of antigens selected on the basis of epitopes identified by EpiMatrix or from information coming from the mycobacterial genome projects.

During the discussion, it was stated that the requirement for a blood specimen would be a drawback, although the use of a finger prick specimen may be more easily implemented. Most of these tests would have limited application in persons with HIV infection

Product Types

Test Type	Advantages	Disadvantages
Antibody test	Ease of application Single visit Combined antigens	
Interferon gamma whole blood	Sensitive - positive before PPD Specificity could be enhanced	Fresh whole blood New equipment Limited applicability in developing country
Recombinant mixture of antigens	Improved specificity Standardized recombinant technology	Not yet developed
Single Antigen preparations	High specific for <i>M. TB</i>	
Urine antigen test	Specific Ease of collection Single visit	Not yet developed

New skin test antigens can be used to develop new antigens for skin testing. These antigens/epitopes can be identified using technology of EpiMatrix and the use of a database of mycobacteria containing genetic sequences/proteins.

Use	Comment
TB Control	The test would be useful if preventive therapy is used in the TB control program
Case Management	If recent infection would be distinguished from remote infection, it would identify individuals at high risk for developing disease
Epidemiologic assessment	Useful to assess impact of control program (ARI) Measure burden of disease
Research	

Product Development Guidelines for two types of screening tests for TB; skin test and serologic/urine test

	Performance Specifications Proposed	Guidelines Recommended for <u>New Skin Test</u>	Guidelines Recommended for <u>Serologic or urine test</u>
Overall	Identify infected persons	Identify persons infected with M. TB	Identify persons infected with M. TB
Specimen	None	None	Blood - antibody Urine - antigen
Quantifiable	No	Yes, size of reaction	Yes
Technician time	<5 minutes	<5 minutes	<5 minutes to collect
Laboratory time	None	None	10-15 minutes
Lab equipment	Syringe/needle	Syringe/needle	Syringe/needle/vacutainer Lancet/filter paper Urine container ELISA kits Equipment for measuring antibody/antigen
Shelf life	6-24 months	6-24 months	6-24 months
Refrigeration	Not required	Not required	Not required
Safety	None	TB and HIV protection needed	TB and HIV protection needed with phlebotomy
Patient visits	Two	Two	One
Cost	\$0.50/test	\$0.50/test; cost of test must include multiple visits	\$0.50/test

Summary Discussion

Discussion relating to a test to replace AFB microscopy for the screening of pulmonary tuberculosis suspects focused on the specificity of both the current test [AFB microscopy] and new tests currently under development. When asked for information on this point, manufacturers indicated that a specificity of 98-99% would be difficult to achieve using a single test that is both rapid and simple. Tests based upon amplification can achieve this degree of specificity but require both a well equipped laboratory and highly trained personnel. They are also expensive. A test that uses a specific bacteriophage [e.g. Biophage] can, at least in theory, be highly specific. But this test requires both equipment [BL-2 hood and incubator] and at least 24 hours. A high degree of specificity could also be achieved by combining tests. For example, use of a single AFB smear [as recommended by Squire] in combination with a simple antigen [or antibody] detection procedure, could provide results that are rapid, highly specific, and technically simple. But the added cost of an additional test may not be practical for many low income countries. These are examples of the trade-off described by Goldberger. New tests can achieve some, but not all, of the properties that are desired. If such trade-offs are acceptable, these tests could be developed using existing technologies.

Dr. Kibuga said that new tests to improve the differential diagnosis of AFB smear negative tuberculosis are urgently needed. The working group that focused on these tests agreed. These cases normally represent 30% of all cases. In some instances be this percentage may be much higher. The working group agreed with Hopewell who calculated that AFB smear negative pulmonary tuberculosis is less infectious [approximately 15%] when compared to AFB smear positive cases.

Tests to improve antibiotic susceptibility testing and tests to detect recent infection [skin tests] were given lower priorities. These tests would be valuable, but in the context of low and middle income countries, only under defined circumstances.