Predicting the Outcome of Therapy for Pulmonary Tuberculosis

ROBERT S. WALLIS, MARK D. PERKINS, MANIJEH PHILLIPS, MOSES JOLOBA, ALICE NAMELE, JOHN L. JOHNSON, CHRISTOPHER C. WHALEN, LUCILEIA TEIXEIRA, BARBARA DEMCHUK, REYNALDO DIETZE, ROY D. MUGERWA, KATHLEEN EISENACH, and JERROLD J. ELLNER

Case Western Reserve University, Cleveland, Ohio; Universidade Federal do Espírito Santo, Vitória, Brazil; Makerere University, Kampala, Uganda; Duke University Medical Center, Durham, North Carolina; and University of Arkansas Medical Center, Little Rock, Arkansas

Patients vary considerably in their response to treatment of pulmonary tuberculosis. Although several studies have indicated that adverse outcomes are more likely in those patients with delayed sputum sterilization, few tools are available to identify those patients prospectively. In this study, multivariate models were developed to predict the response to therapy in a prospectively recruited cohort of 42 HIV-uninfected subjects with drug-sensitive tuberculosis. The cohort included 2 subjects whose initial response was followed by drug-sensitive relapse. The total duration of culture positivity was best predicted by a model that included sputum M. tuberculosis antigen 85 concentration on Day 14 of therapy, days-to-positive in BACTEC also performed adequately (R = 0.58). Both models predicted delayed clearance of bacilli in both relapses (> 85th percentile of all subjects) using information collected during the first month of therapy. Stratification of patients according to anticipated response to therapy may allow TB treatment to be individualized, potentially offering superior outcomes and greater efficiency in resource utilization, and aiding in the conduct of clinical trials.

There is substantial variability in the response to therapy for pulmonary tuberculosis, even in those patients with fully drug sensitive isolates. In some patients, bacilli are killed rapidly and cleared quickly from sputum. In others, viable organisms persist for many weeks or months, despite multidrug treatment. In yet others, bacilli are cleared, only to reappear after therapy is stopped. These observations form the basis for the definitions of treatment failure and relapse, respectively. The causes of this phenomenon are not well understood, but they may involve mycobacterial and host biologic factors as well as host behavioral factors.

The search for tools to monitor tuberculosis therapy and predict outcome is made particularly complex by the observation that mycobacterial killing is not a single uniform process. Most actively replicating bacilli are killed rapidly during the first 1 to 2 wk of therapy. This phase of treatment can be measured by quantitative sputum culture (early bactericidal activity or EBA) (1–5). There is, however, no known relationship between EBA and the outcome of treatment. Prolonged treatment is required to eradicate persisting organisms with reduced or otherwise altered metabolic activity. Nonreplicating bacilli show reduced susceptibility to the bactericidal activities of antimycobacterial drugs (6, 7). This later, sterilizing phase of therapy appears to be distinct from the first, based in part on the differential activities of antimycobacterial drugs during the two phases.

Two studies indicate that, in contrast to EBA, the time to sterilization is an important determinant of outcome. The first, by Aberg and Nunn (8), determined that relapses occurred more frequently in those patients whose sputum cultures remained positive on agar by the third month of therapy. This forms the basis for the WHO recommendation that treatment be prolonged for those patients with delayed sputum sterilization, though this end point is infrequently reached during modern chemotherapy, even in patients who later relapse. In the second, Mitchison (9) reviewed the time to sputum sterilization of subjects enrolled in published comparative tuberculosis chemotherapy trials. He found that regimens with superior sterilizing activity at 2 mo had lower relapse rates, and he suggested that this parameter might be used as an early indicator of the relative efficacy of various regimens, if not to predict outcome in individual patients. These measures have not been widely used clinically, however, because they are primarily retrospective measures. The time to sputum sterilization is usually not known until treatment is essentially complete, particularly when agar cultures are used.

Two recent studies have indicated that earlier measures of the response to tuberculosis therapy might be useful to stratify patients according to risk of adverse outcome. In 1997, Epstein and colleagues (10) reported that failure (sustained culture positivity despite therapy) can be readily identified after 4 to 6 wk of therapy by measuring days-to-positive in mycobacterial growth indicator tube (MGIT) cultures. In 1998, a report from this laboratory indicated that relapse may be predicted by measuring M ycobacterium tuberculosis antigen 85 in sputum after 2 wk of therapy (11). In the present report, the interactions of these and other parameters were examined. The objective was to determine whether a multivariate model could predict the overall risk of failure and relapse in the treatment of drug-sensitive tuberculosis.

METHODS

Data were analyzed from a prospective study conducted in Uganda and Brazil, and reported previously (11). Briefly, patients with initial episodes of pulmonary tuberculosis who had not received prior therapy were prospectively recruited at tuberculosis control clinics in Kampala, Uganda, and Vitória, Brazil. All patients gave informed written consent for HIV testing and study participation. The study protocol was approved by the institutional review boards of Case Western Reserve University (Cleveland), Makerere University (Kampala), Universidade Federal do Espírito Santo (Vitória), Duke University (Durham) and the University of Arkansas (Little Rock). Tuberculosis was presumptively diagnosed by a positive acid-fast smear of sputum and a compatible chest radiograph and was subsequently confirmed by culture. Serology for HIV-1 was performed on all sub-
jects; seropositives were excluded. Subjects were also subsequently excluded from this analysis if their initial isolates were resistant to isoniazid, rifampin, pyrazinamide or ethambutol, or the duration of follow-up was less than 180 d.

Baseline information was collected as to age, body mass, and radiographic extent of disease, using criteria established by the National Tuberculosis and Respiratory Disease Association (12). Subjects were classified as minimal disease (coded with a value of 1) if lesions were noncavitary, of slight to moderate density, and involving a small part of one or both lungs. The total extent was required to be less than the volume of one lung above the second chondrosternal junction and the spine of the fourth or body of the fifth thoracic vertebra. Subjects were classified as having moderately advanced disease (coded as 2), if they had more than minimal disease but had a total extent of slight or moderately dense lesions limited to the total volume of one lung, and that of dense lesions limited to one-third the volume of one lung. Cavitary lesions were required to be < 4 cm in diameter. Subjects were classified as having far advanced disease (coded as 3) if lesions were more extensive than moderately advanced.

Subjects were treated with daily isoniazid, rifampin, ethambutol, and pyrazinamide at standard doses for 2 mo, followed by daily isoniazid and rifampin for 4 mo. Patients were evaluated on Days 0, 2, 4, 7, 14, 30, monthly until therapy was completed, and then bimonthly. At each evaluation, history and physical examination were performed, and multiple sputum specimens were obtained, which were processed and analyzed as described below. Chest radiography was repeated at regular intervals. Patients were hospitalized for the initial 2 wk of treatment, after which therapy was self-administered. Compliance was assessed at each clinic visit by review of dispensing records, pill-dispensing and clinic-attendance data, and by urinary isoniazid metabolite testing (Mycobacterium tuberculosis). The probability of rare events was determined using the Poisson method (16).

RESULTS

The clinical trial forming the basis of this analysis included 42 evaluable subjects. The frequency distributions of the day of collection of the last sputum specimen to indicate rapid growth of Mycobacterium tuberculosis (BACTEC medium, or any growth on solid medium, are represented in Figure 1. Rapid growth in BACTEC was defined as that detected within 20 d of inoculation; this threshold was selected based on the report of Epstein and colleagues (10) that this parameter was an early indicator of treatment failure. The distributions of the agar and BACTEC parameters differed, in that once therapy had started, cultures were more likely to be rapidly positive in BACTEC than positive on solid medium (p = 0.049 using Fisher’s exact test). This indicates that agar culture is less sensitive than BACTEC for the detection of viable mycobacteria once therapy has been initiated.

Two relapses were identified. Both initially responded, becoming culture negative on agar after Days 14 and 30, and in BACTEC on Days 30 and 90, respectively. Both subjects had recurrence of clinical disease and repeated isolation of drug-sensitive Mycobacterium tuberculosis on agar after completion of therapy. Nonadherence during the last month of therapy was documented in only one case (by pill-dispensing and clinic-attendance records). In both cases, at least one culture in BACTEC was rapidly positive during the last month of therapy, again indicating that this method may be more sensitive for early detection of recrudescent disease. For the purpose of data analysis, the day of last positive culture (both BACTEC and agar) for these two cases was recorded as 180 (the day of completion of tuberculosis therapy).

Multiple linear regression analysis was performed to identify early predictors of the duration of culture positivity, using the best subsets method. The analysis included the Day 7 to 30 values of all four microbiologic measures, as well as all baseline clinical parameters (including study site). As indicated in Table 1, antigen 85 on Day 14 and BACTEC DTP on Day 30 were strong independent predictors of the duration of culture positivity. The synergy between these two parameters appeared to contribute substantially to the accuracy of the model. Radiographic extent of disease entered the model but was of marginal statistical significance. The overall significance value of the regression model was p = 0.005, R = 0.63. The relationship between observed and predicted values is shown in the left panel of Figure 2. The predicted values of the relapsed cases were 74 and 76 d, representing the 85th and 88th percentiles, respectively, of the predicted values of all subjects.

Regression analysis can occasionally be unduly influenced by variables that are not normally distributed, or by a small number of outlying observations. Several tests were performed to examine these concerns. A log transformation, extent was no longer a significant predictor of outcome, but Day 14 antigen 85 and Day 30 DTP remained highly significant. To determine the influence of each single observation on the regression coefficients, the Cook’s distance value for each
data point was calculated (17). This parameter measures the extent to which regression coefficients change when a single observation is deleted from the data set. Cook and Weisberg (18) suggest that points with values > 1 should be evaluated in more detail. The greatest Cook’s distance identified in this data set was 0.24. These findings suggest that the model is statistically sound.

Most cases of tuberculosis occur in regions of the world where laboratory facilities are limited. The question therefore arises whether low-cost alternative models can be developed that do not require ELISA plate readers or BACTEC. Although Day 14 antigen 85 values correlated highly with those on Day 7, they did not correlate with any of the other quantitative microbiologic parameters. As a consequence, it was not possible to develop satisfactory models that did not include antigen 85. In contrast, BACTEC DTP, quantitative AFB smear, and CFU were very highly collinear (p ranging from 10^-4 to 10^-2). Therefore, there were many alternatives to BACTEC. The model with greatest potential application in low-income regions is shown in Table 2. This model does not require mycobacterial culture but instead substitutes quantitative acid-fast microscopy. The correlation coefficient (R) of the model was 0.58. The relationship between observed and predicted values is indicated in the right panel of Figure 2. Additionally, no models could be developed to predict the last day of positive culture on solid medium; this likely reflected the increased skewness and kurtosis of that data.

Two of the three parameters included in the models described above were collected prior to the initiation of treatment or during its in-hospital phase. These parameters were therefore unaffected by patient adherence. The third parameter was obtained on Day 30, 2 wk after discharge from hospital. To objectively assess patient adherence during the outpatient phase of therapy (which was not directly observed), urine was collected at each clinic visit, and was tested for a metabolite of isoniazid. All Day 30 specimens were collected; of these, all but one (98%) were positive. Of the 210 specimens to be collected during Months 2 to 6, 198 (94%) were actually collected, and of these, all but 16 (92%) were positive. This trend toward reduced adherence in Months 2 to 6 did not reach statistical significance (p = 0.16 by Poisson analysis). There was no correlation between the proportion of positive tests and the day of last rapidly positive culture (p = 0.9).

**DISCUSSION**

The eradication of tuberculosis has proven to be an elusive goal. Strategies for eradication based on treatment of active cases require highly effective regimens for success. However, high cure rates often are not achieved in the clinical setting. For example, the WHO recently reported cure rates of 78% in countries that used directly observed short course therapy (DOTS), and 45% in those that did not (19). True rates may be lower, as the WHO classification system does not require culture evidence of cure and does not reclassify “cures” that subsequently relapse. Despite its potential advantages, implementation of DOTS has been limited by its cost, which may be more than three times that of self-administered therapy (20, 21). Other strategies may not be widely implemented either, unless they demonstrate reduced costs as well as improved outcomes.

The predictive models developed in this study may help to address this need. This report indicates that parameters measured during the first month of treatment may be used to predict the microbiology of specimens obtained many months later. Larger trials will be required to determine the extent to which the predicted parameter—the day of last rapidly positive culture—correlates with true risk of treatment failure or relapse, as these clinical outcomes occurred in only two patients in this study. If this approach were validated, prospective trials would be indicated in which the intensity or duration of treatment was increased for those at high risk. Because the models can predict rapid as well as delayed bacillary clearance, trials in which the duration of therapy was reduced for selected subjects might also be indicated. Studies of such “ultra-short” tuberculosis therapy indicate that many patients can be cured with as little as 3 or 4 mo of treatment (22). Stratification according to risk may thus potentially offer increased efficiency in resource utilization as well as improved outcomes.

The most serious potential limitation to these models here is the extent to which they can be implemented where the need is greatest. The analysis indicates that the requirement for culture can be replaced by the use of quantitative acid-fast smear, but the need for ELISA remains. ELISA for antigen 85

**TABLE 1**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-3.985</td>
<td>25.691</td>
<td></td>
</tr>
<tr>
<td>Antigen 85, Day 14</td>
<td>0.459</td>
<td>0.146</td>
<td>0.004</td>
</tr>
<tr>
<td>BACTEC DTP, Day 30</td>
<td>-0.564</td>
<td>0.217</td>
<td>0.016</td>
</tr>
<tr>
<td>Extent of disease</td>
<td>17.879</td>
<td>8.501</td>
<td>0.046</td>
</tr>
</tbody>
</table>

* The overall significance of the regression model is 0.005, R = 0.63.
was performed in this study using enzyme signal amplification. This method increases the sensitivity of the assay, but also increases its complexity. In retrospect, such sensitivity may not be required for analysis of these specimens, as the critical threshold for antigen 85 appears to be approximately 50 pg/ml. This level may be detected readily by conventional ELISA. It may also be detectable by methods based on agglutination or through the use of antibody-impregnated paper matrices. Utilization of such a test would depend on the extent to which it reduced the costs of therapy in those patients at low risk for relapse. Like most ELISA's, the cost of the antigen 85 assay is between $5 and $10 per specimen. One possible strategy for its use is in the selection of patients for directly observed therapy. Zwarenstein and colleagues reported that for patients who lived in South Africa, self-administered therapy was cost effective even if only a small proportion were shifted to self-administered therapy.

Predictive models such as these may also assist in the evaluation of new drugs in early clinical trials. Conventional trials, in which one new drug is added to an existing regimen may fail to demonstrate an additive effect unless large numbers of subjects are studied. An alternative strategy for phase II evaluation of promising but unproven compounds has been suggested in which brief treatment with an experimental regimen is followed by full treatment with a standard regimen. The overall significance of the relapse risk associated with culture correlates with outcome in patients receiving treatment for pulmonar tuberculosis. Predicted values are not shown for those subjects with incomplete data.

In summary, a simple model was developed to predict the duration of culture positivity in pulmonary tuberculosis. Given the relationship between this parameter and clinical outcome, the model may be of value in the evaluation of new tuberculosis therapeutics, as well as in the care of individual patients.

**References**


---

**TABLE 2**

**LINEAR REGRESSION MODELS TO PREDICT THE DAY OF LAST RAPIDLY POSITIVE SPUTUM CULTURE, BASED ON MEASURES OBTAINED DURING THE FIRST MONTH OF THERAPY EXCLUDING CULTURE**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-41.909</td>
<td>27.048</td>
<td>0.008</td>
</tr>
<tr>
<td>Antigen 85, Day 14</td>
<td>0.400</td>
<td>0.140</td>
<td>0.001</td>
</tr>
<tr>
<td>qAFB, Day 30</td>
<td>23.415</td>
<td>10.036</td>
<td>0.027</td>
</tr>
<tr>
<td>(qAFB Day 30)²</td>
<td>-2.907</td>
<td>1.551</td>
<td>0.071</td>
</tr>
<tr>
<td>Extent of disease</td>
<td>13.004</td>
<td>7.455</td>
<td>0.091</td>
</tr>
</tbody>
</table>

* The results of quantitative AFB microscopy (qAFB) were best described using a second-order equation, with box x and x² components. Extent of disease was forced into this model to allow comparison with that in Table 1. The overall significance of the regression model is 0.002, R² = 0.58.


