Development of a clinical prediction rule for tuberculous meningitis in adults in Lima, Peru

Lely Solari¹², Alonso Soto², Patrick van der Stuyft¹³

¹ Unit of General Epidemiology and Disease Control, Institute of Tropical Medicine of Antwerp, Belgium
² Escuela de Medicina, Universidad Peruana de Ciencias Aplicadas, Lima, Peru
³ Department of Public Health, Ghent University, Belgium

Abstract

Objectives: Diagnosis of Tuberculous Meningitis (TM) is a challenge in countries with a high burden of the disease and constrained resources and Clinical Prediction Rules (CPRs) could be of assistance. We aimed at developing a CPR for diagnosis of TM in a Latin American setting with high tuberculosis incidence and a concentrated HIV epidemic.

Methods: We enrolled adult patients with clinical suspicion of TM attending two hospitals in Lima, Peru. We obtained information on potential anamnestic, clinical and laboratory predictive findings that are easy to collect and promptly available. We independently diagnosed TM according to a composite reference standard that included a series of microbiological tests. We performed bivariate analysis and constructed a logistic regression model to select the predictive findings associated with TM. With the selected predictors included in the model we developed a score-based CPR. We assessed its internal validity and diagnostic performance.

Results: Of 155 analyzed patients, 59 (38%) had TM. The CPR we derived includes three predictors: cough for 14 days or more, 10-500 cells in CSF and adenosine deaminase ≥ 6 U/L in CSF. It classifies patients into high, moderate or low score groups and has an overall area under the ROC curve of 0.87. 59% of patients were assigned to either the high or low score group, permitting prompt decision making. In patients in the high score group, it attains a positive likelihood ratio for TM of 10.6 and in patients with low scores a negative likelihood ratio of 0.10. Bootstrap analysis indicated high internal validity.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/tmi.13041

This article is protected by copyright. All rights reserved.
Conclusion: This CPR could support decision-making in patients with clinical suspicion of TM. External validation and further assessment of its clinical impact is necessary before application in other settings.

Keywords: Meningitis, score, Mycobacterium tuberculosis, adenosine deaminase

INTRODUCTION
Despite global advances in tuberculosis control, case identification is a challenge and physicians are facing diagnostic dilemmas, in particular when they manage cases with clinical suspicion of smear-negative pulmonary or extra pulmonary tuberculosis. Tuberculous meningitis (TM) and miliary tuberculosis are the most severe extra pulmonary presentations, causing considerable morbidity and mortality. Death and secondary sequelae due to TM result from delayed treatment initiation, which often occurs because patients with TM usually attend with unspecific signs and symptoms of central nervous system involvement. Still, patient management has to be decided upon with the information available within the first hours after admission. These are elements from the history, physical examination and unsophisticated laboratory tests, while awaiting the results of more specific or sophisticated tests, if available.

Furthermore, there is no single satisfactory test for TM: acid–fast bacilli (AFB) smears of cerebrospinal fluid (CSF) and cultures for mycobacteria, considered the reference standard, have low sensitivity. The added value of molecular tests is only marginal and the evaluations of other tests in CSF such as IGRAs, and dosage of C reactive protein have yielded mixed results. The utility of determining adenosine deaminase activity (ADA) in CSF, a test used in many settings, is also controversial, despite several publications showing a good diagnostic performance.

Therefore, Clinical Prediction Rules (CPRs), commonly known as “scores”, have been proposed as tools to support or rule out the diagnosis of TM. They frequently include as predictive findings the results of routine CSF analyses, such as the number and formula of leucocytes and protein and glucose levels, besides information from the clinical history and physical examination. Many of these CPRs have been designed to differentiate TM from Acute Bacterial Meningitis. Amongst these, Thwaites’ CPR, has undergone external validation with encouraging results. Cohen’s CPR, on the other hand, focuses on the differentiation of TM from cryptococcal meningitis in people living with HIV/AIDS and Hristea’s CPR on distinguishing between TM and viral meningitis. The only existing CPR that can assist clinicians in differentiating between TM and all other meningeal conditions above is intended for use in patients with a chronic meningitis syndrome of 7 or more days of duration. Of note, none of the available CPRs includes ADA as a predictive finding of TM, despite it being...
acknowledged as a useful clinical tool for diagnosing this condition.\textsuperscript{7,8} We aimed to develop a CPR to distinguish TM from the wide array of other meningeal conditions in adults presenting with clinical suspicion of TM, considering patients’ clinical signs and symptoms and simple laboratory tests in CSF, including ADA, as potential predictive findings.

METHODS

Setting and patients

Peru is a middle-income country with a high incidence of tuberculosis (119/10\textsuperscript{5}), a concentrated HIV epidemic\textsuperscript{20} and with 6.3\% of multidrug-resistant tuberculosis among new cases\textsuperscript{21}. We performed the study in 2 third-level hospitals in Lima: Hospital Nacional Hipolito Unanue and Hospital Nacional Cayetano Heredia. The sample size needed for deriving the CPR was 160 patients with clinical suspicion of TM, given an estimated TM prevalence of 60\% among suspects and 9 candidate predictors\textsuperscript{22}.

From November 2009 to February 2012, we included all patients 18 years or older admitted in the emergency and medicine wards of both hospitals with clinical suspicion of TM. To be considered for inclusion by the attending physicians they had to have at least 2 of the following: headache, irritability, vomiting, fever, seizures, neurologic deficit or altered consciousness, with no other explanatory medical condition. We excluded patients who upon admission had already a defined diagnosis and had been put on treatment, for instance, for cryptococcal or bacterial meningitis and patients on anti-tuberculous treatment.

Anamnestic data were collected through direct patient interview or by interviewing close relatives in case of altered consciousness. Lumbar puncture was performed by the attending physicians according to standard procedures.

Predictive findings of TM

We chose the potential predictive findings to be evaluated based on the case definition of TM suggested by Marais et al\textsuperscript{23}. However, results of specialized radiological examinations, Computed Tomography Scan and Magnetic Resonance Imaging were not withheld given their inconstant availability in settings similar to ours and we also excluded findings with potentially high inter-observer variability, such as cranial nerve palsies, focal deficits and “clear appearance” of the CSF as we considered that only easily identified findings could be applicable in resource-constrained settings. As “clinical-anamnestic” predictive findings, we analyzed the duration of symptoms as a continuous variable, cough for more than 2 weeks, the presence of other systemic symptoms (weight loss and night sweats) and contact with a PTB patient. As potential laboratory predictive findings we considered presence of cells in CSF, lymphocytic predominance, CSF: plasma glucose ratio and protein level. We
added ADA levels in CSF, determined through the Giusti method.

Other tests in CSF
In order to reach microbiological confirmation of the meningitides, we systematically performed AFB staining, culture on Mycobacterial Growth Indicator Tube (MGIT) and solid media (Ogawa), polymerase chain reaction (PCR) for *Mycobacterium tuberculosis* (IS6110 PCR, Qiagen Multiplex PCR), Gram staining, culture for common bacteria, India ink staining and Cryptococcal latex agglutination test on the CSF samples. The attending physicians could request other tests (PCR for viruses, specific serological tests, biopsies and cultures from other sites of the body, etc.), according to their clinical judgement. We took these results into account to establish the final diagnoses. For instance, diagnosis of bacterial and cryptococcal meningitis were based on positive cultures or presence of antigens in CSF, whereas diagnosis of viral meningitis was based on a positive PCR for virus in CSF or on high lymphocytes in the CSF plus negativity of the aforementioned tests with resolution of symptoms and pleocytosis without antibiotic or anti-tuberculous treatment.

The performance of the laboratories of both participating hospitals is monitored through the national network of public health laboratories of Peru.

Definition of TM
We took as composite reference standard for TM the presence of any of the following: a) bacteriological evidence in CSF: positive AFB in smears, positive culture or positive PCR test for *M. tuberculosis*; or b) bacteriological evidence in other specimens: positive culture for *M. tuberculosis* in other body fluids or biopsies and a negative Gram staining and cultures for bacteria, negative Cryptococcal latex agglutination test and negative culture for fungi in CSF; or c) clinical plausibility, defined as negative Gram staining and cultures for bacteria, a negative Cryptococcal latex agglutination test and negative cultures for fungi in CSF and the decision of a National Tuberculosis Program expert to initiate anti-tuberculosis therapy and good clinical response (complete resolution of the constitutional symptoms at one month after treatment initiation).

The database manager classified patients as having TM or not according to these criteria, blinded to the other results of the clinical/laboratory predictive findings.

Data analysis
We excluded from the analysis patients with laboratory results evidencing concomitant etiologies (coinfection) for the meningitis. We analyzed the association between potential predictive findings and TM using Mann-Whitney test for continuous variables and Pearson’s χ² test for categorical vari-

This article is protected by copyright. All rights reserved.
ables and selected the predictive findings that were associated with TM at a level of significance of \( p \leq 0.10 \) to include them in a multivariate logistic regression model. We performed logistic regression using backward elimination; retaining predictive findings associated with TM at \( p \leq 0.10 \). Subsequently, we converted continuous predictive findings into categorical ones. We created 10-years age groups and 7-days-of-disease groups for symptoms duration and we dichotomized laboratory variables using currently recommended cutoff points (ADA 6U/L, proteins >1 g/L, glucose CSF: serum ratio >0.5) \(^{23} \). We run the logistic regression again with the categorical variables and also calculated the Hosmer and Lemenshow goodness of fit test, the Nagelkerke R\(^2 \), odds ratios and likelihood ratios for both models.

We then transformed the best model into a score by assigning to each predictive finding a number of points according to and proportional with its odds ratio. ROC curve analysis allowed us to select the best cutoff point(s) for this score. A positive likelihood ratio of >10 and a negative likelihood ratio of <0.10 were considered convincing evidence in favor of and against TM, respectively \(^{25} \). Finally, we evaluated the performance of the score through standard diagnostic performance assessment and by means of bootstrapping in the original data set.

For reporting the results of this study, we followed the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines \(^{26} \).

**Ethical aspects**
The ethics committees of Universidad Peruana Cayetano Heredia, of both participating hospitals and of the Institute of Tropical Medicine, Antwerp approved the study. All included patients, or a direct relative in case of altered consciousness, gave informed consent for participation.

**RESULTS**
15 patients were admitted with a defined diagnosis and had already been put on treatment for a specific meningitis. 217 patients had clinical suspicion of TM without definite diagnosis and were eligible according to our criteria. 45 refused to participate and 15 died before the lumbar puncture was performed. Thus, 157 patients were included in the study but two were excluded from the analysis due to concomitant infection with 2 agents (both patients had HIV AIDS, a positive culture for *Cryptococcus neoformans* and a positive PCR for *Mycobacterium tuberculosis* in CSF). Eventually, 155 patients were analyzed.

According to our reference standard, 59 (38%) patients had TM. 18 (30.5%) of them were diagnosed by bacteriological evidence in the CSF, 7 (11.9%) had bacteriological evidence of *M. tuberculosis* in other body fluids and 34 (57.6%) patients fulfilled our definition of clinically plausible TM.
There were 96 (62%) patients with medical conditions other than TM (Non-TM). 37 (23.9%) had other bacterial (6), viral (19) or cryptococcal (12, all HIV positive) infectious meningitides. 26 (16.8%) had non-meningeal infections, including toxoplasmosis and severe sepsis with neurological impairment, while other miscellaneous conditions including subarachnoid hemorrhage and metabolic encephalopathies were diagnosed in 33 (21.3%) patients.

The mortality at one month of follow up was 25.8% for TM patients and 17.9% for non-TM patients (p=0.23). Overall, there were 55 patients with HIV AIDS, 22 (38%) in the TM and 33 (34%) in the non-TM groups.

Table 1 shows the results of the bivariate analysis for the 9 potential predictive findings in TM and Non-TM patients. There were no missing values. Eight predictive findings were associated with TM at the level of $p \leq 0.10$ and were selected for inclusion in the logistic regression models. We found no correlations between them. In multivariate analysis, only cough for more than 14 days, presence of 10-500 cells/L in CSF and ADA level in CSF remained associated with TM at $p<0.10$. The calibration of the model was good (Hosmer and Lemeshow goodness of fit test $p=0.18$) and it had a Nagelkerke R2 of 0.37. The model with continuous variables dichotomized also had an appropriate goodness of fit (Hosmer and Lemeshow $p=0.30$) and a Nagelkerke R2 of 0.40. Since both models were comparable, and in order to keep the CPR parsimonious, the later one was used for derivation of the CPR (Table 2). The number of points assigned to each predictive finding in the score was proportional to its odds ratio. With bootstrapping we found the same final model and the same odds ratios, be it with wider confidence intervals.

The area under the ROC curve for the score was 0.87 (95% C.I. 0.81-0.92). We found that no single cutoff point fulfilled our accuracy goal, and selected 2 cutoff points. A cutoff of $<1$ point had a sensitivity of 94.9% (95% C.I. 85.8-98.9) and a specificity of 52.6% (95% C.I. 42.1-63.0), implying that patients scoring less than 1 point had a very low probability of TM. A cutoff $\geq 3$ points had a sensitivity of 55.9% (95% C.I. 42.4-68.8) and a specificity of 94.7% (95% C.I. 88.1-98.2), indicating that patients with 3 or more points had a high probability of having TM. Table 3 shows the number of patients in each score category and the associated likelihood ratios. 59% of patients were assigned to either the high or low score group, while 41% remained in the intermediate (indeterminate) category. The CPR led to correct classification as non-TM in 94% of patients in the low-score group and as TM in 85% of patients in the high-score group.

We additionally performed a subgroup analysis, comparing patients with microbiological confirmation of TM to patients with other infectious meningitides. The logistic regression analysis gave very high ORs for the variables included in the score: 31.8 for cough more than 14 days, 32.0 for CSF WBC between 10-500 and 91.2 for ADA $\geq 6$U/L. As the relative values of these ORs remain roughly
unchanged, the hypothetical CPR for TM based on this model would be the same. The area under the ROC curve (0.85, 95%CI 0.76-0.96) in this sub-population was very similar to the one obtained for the complete dataset.

DISCUSSION
We constructed a CPR for the diagnosis of TM in patients with clinical suspicion of this condition using anamnestic clinical and laboratory CSF predictive findings that are easy to collect and rapidly available. It includes having cough for more than 2 weeks – the WHO definition of “respiratory symptomatic”, the presence of 10-500 cells/L in CSF and an ADA level of ≥6 U/L. It leads to categorization of suspects into low, intermediate and high TM score groups, attaining a positive likelihood ratio of 10.62 in patients with high scores, a negative likelihood ratio of 0.10 in patients with low scores and an area under the ROC curve of 0.87. Bootstrap analysis found a high internal validity. This CPR permits prompt decision-making in patients having either high or low scores, which represent 59% of the patients attending with clinical suspicion of meningitis in our setting. In patients with high scores, in whom TM is very likely, immediate anti tuberculous treatment initiation ought to be considered. In patients with low scores, in whom there is fair evidence against TM, efforts should be redirected towards confirming other etiologies for the symptoms and anti-tuberculous treatment initiation could be withheld, unless bacteriological evidence of tuberculosis becomes available subsequently. In patients with intermediate scores, TM treatment initiation could be considered based on additional criteria and on the clinician’s judgement, but not on the basis of the score. A CPR supporting decision-making in 59% of patients with clinical suspicion of TM can lead to substantial improvements in outcome over the widely variable current practice.

Only 30.5% of our TM cases were confirmed through bacteriological tests in CSF, a percentage that is within the usual range but it still constitutes a limitation of the study. If access to health services improves and patients arrive earlier in the course of the disease, this percentage would even further decrease. Diagnosis of TM through “clinical plausibility”, defined by clinical response to anti-tuberculous therapy, may be challenged by some authors, but when no suitable reference standard is available, relying on clinical arguments is an adequate approach. Bacteriological confirmation of TM could possibly have been enhanced by increasing the volume of the CSF sample used for microbiological analyses, but in routine practice, clinicians – just as we did – ask for a whole range of tests, as we did, and enough CSF volume should be available for all of these. Another constraint was the lower than expected proportion of TM in the suspect patients, which led to a relatively small absolute number of TM cases. This reduced the number of TM events, implying a somewhat lower statistical power than expected. Even so, our TM series is one of the biggest in Latin America.
The CPR we developed differs in the predictive findings included from other ones due to several reasons. Firstly, most of the existing CPRs are intended to differentiate TM from bacterial meningitis\textsuperscript{9,12}, such as Thwaites’ index\textsuperscript{9}, and so are not directly comparable to our CPR. It includes predictors like age, or blood cell count. Moreover, the spectrum of patients from which most were derived only comprises patients with microbiologically confirmed conditions. Considering a wider spectrum of possible differential diagnoses poses more constraints on the CPR’s derivation, but corresponds more closely to routine clinical practice and its challenges, which constitutes a strength of our study.

Secondly, our setting differs from that in which most others CPRs were derived, mainly Asian countries with high prevalences of bacterial meningitis among patients with suspicion of TM. Bacterial meningitis is a condition whose frequency is decreasing in Latin America\textsuperscript{29}. Along the same line, our findings may not directly apply to settings with high HIV prevalence or with low tuberculosis incidence, where the proportion of cryptococcal meningitis or other aseptic meningitides, respectively, would be higher. In addition, HIV positivity could influence the accuracy of the CPR, but a larger number of HIV positive patients than available in our series is needed to further explore this. HIV status can alter the inflammatory profile of the CSF\textsuperscript{30} and it seems to affect the performance of Thwaites’ index\textsuperscript{31}.

Thirdly, we selected as potential predictors only those ones that are inexpensive and easy to obtain information on in hospitals in low and middle-income countries and that are highly reproducible. For that reason, we excluded findings from Computed Tomography and Magnetic Resonance Imaging and from neurological examination. Notwithstanding, we acknowledge their utility in individual patient care\textsuperscript{32}. Finally, other CPRs did not investigate ADA as a potential predictive finding. Despite controversy on the contribution of ADA to the diagnosis of TM when used as an isolated test, we found that as a part of a CPR for TM ADA is an important contributor.

The potential of molecular tools for the diagnosis of TM has to be acknowledged, however, the currently deployed tests, such as GeneXpert, only provide marginal contributions and are less sensitive than liquid cultures\textsuperscript{33,31}. Notwithstanding, more sophisticated technology such as one-tube nested PCR-lateral flow strip test (OTNPCR-LFST)\textsuperscript{35} and the new GeneXpert generation, Xpert Ultra\textsuperscript{36}, may be more promising from the point of view of sensitivity in CSF. While research on the development of new diagnostic tests for TM is important, research on clinical tools such as CPRs should not be left behind. CPRs are of significant help in the absence of accurate low-cost point of care test and essential until further refined molecular tools become widely available. They have the potential of directly supporting medical decision-making based on the use of already existing, easily accessible and inexpensive tests. We derived a CPR for the diagnosis of TM that seems to be promising, but external validation of its performance and further assessment of its clinical utility and impact is necessary.

This article is protected by copyright. All rights reserved.
before wider application in other settings.

ACKNOWLEDGEMENTS
We thank Francine Matthys† for her contribution in the development of the study protocol; Daniela Salazar and Yenny Bravo for their active participation in patient recruitment; Dante Vargas, Roberto Accinelli and Juan Agapito for active participation in management of patients and laboratory procedures:

REFERENCES
(10) Youssef FG, Afifi SA, Azab AM, Wasfy MM, Abdel-Aziz KM, Parker TM et al. Differentiation of

This article is protected by copyright. All rights reserved.


Correspondence: Alonso Soto, Luis Galvani 249, Surco, Lima, Peru, Email sotosolari@yahoo.com

This article is protected by copyright. All rights reserved.
Table 1. Considered predictive findings for Tuberculous Meningitis (TM) in patients with clinical suspicion of TM, Lima, 2009-2012

<table>
<thead>
<tr>
<th>Predictive finding</th>
<th>TM (n=59)</th>
<th>Non-TM (n=96)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical-Anamnestic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptom duration (days) (median, interquartile range)</td>
<td>14 (7-26)</td>
<td>7 (3-20)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Presence of systemic symptoms (weight loss, night sweats) (%)</td>
<td>55 (93)</td>
<td>75 (78)</td>
<td>0.01</td>
</tr>
<tr>
<td>Presence of cough for more than 14 days (%)</td>
<td>20 (34)</td>
<td>12 (13)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Contact with a patient with pulmonary tuberculosis (%)</td>
<td>26 (44)</td>
<td>22 (23)</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Cerebrospinal (CSF) tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence of cells (10-500/L) in CSF (%)</td>
<td>43 (73)</td>
<td>30 (31)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Lymphocytic predominance (&gt;50% lymphocytes) (%)</td>
<td>39 (66)</td>
<td>45 (47)</td>
<td>0.20</td>
</tr>
<tr>
<td>CSF:serum glucose ratio (median, interquartile range)</td>
<td>0.39 (0.26-0.52)</td>
<td>0.52 (0.43-0.65)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adenosine deaminase level in CSF (U/L) (median, interquartile range)</td>
<td>9 (4-14)</td>
<td>2 (1-4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Protein level in CSF (mg/dl) (median, interquartile range)</td>
<td>135 (58-189)</td>
<td>47.5 (30-86)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Table 2. Best fitting multivariate model for diagnosis of Tuberculous Meningitis (TM) and number of points assigned to each predictive finding in the proposed score

<table>
<thead>
<tr>
<th>Predictive finding</th>
<th>Odds ratio (95% CI)</th>
<th>p-value</th>
<th>Points assigned in the score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of cough for more than 14 days</td>
<td>7.0 (2.2-22.5)</td>
<td>0.001</td>
<td>1</td>
</tr>
<tr>
<td>Presence of 10-500 cells/ul in CSF</td>
<td>8.2 (3.0-22)</td>
<td>&lt;0.001</td>
<td>1</td>
</tr>
<tr>
<td>Adenosine deaminase (ADA) level in CSF ≥6 U/L</td>
<td>27.9 (8.2-95)</td>
<td>&lt;0.001</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3. Prevalence of Tuberculous Meningitis (TM) in each category of the CPR for TM and associated likelihood ratios

<table>
<thead>
<tr>
<th>Score on the CPR</th>
<th>Number of patients (%)</th>
<th>TM (%)</th>
<th>Positive likelihood ratio (95% C.I.)</th>
<th>Negative likelihood ratio (95% C.I.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (&lt;1 point)</td>
<td>53 (34%)</td>
<td>3 (5.7)</td>
<td>2.0 (1.6-2.5)</td>
<td>0.10 (0.03-0.30)</td>
</tr>
<tr>
<td>Intermediate (1-2 points)</td>
<td>63 (41%)</td>
<td>23 (36.5)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>High (≥3 points)</td>
<td>39 (25%)</td>
<td>33 (88.8)</td>
<td>10.6 (4.4-25.7)</td>
<td>0.47 (0.35-0.62)</td>
</tr>
</tbody>
</table>
Figure 1. Patients with clinical suspicion of Tuberculous Meningitis (TM), final diagnosis and performance of the Clinical Prediction Rule (CPR), Lima, Peru

232 patients with suspicion of TM (see text for criteria)

- 15 already on specific antimicrobial treatment
- 45 refused to participate
- 15 died before lumbar puncture

157 patients included

- 2 with concomitant infection with 2 agents

59 patients with TM (38%)

- Low score 3 (5.1%)
- Int. score 23 (39.0%)
- High score 33 (55.9%)

96 patients with other conditions (62%)

- Low score 50 (52.1%)
- Int. score 40 (41.7%)
- High score 6 (6.3%)