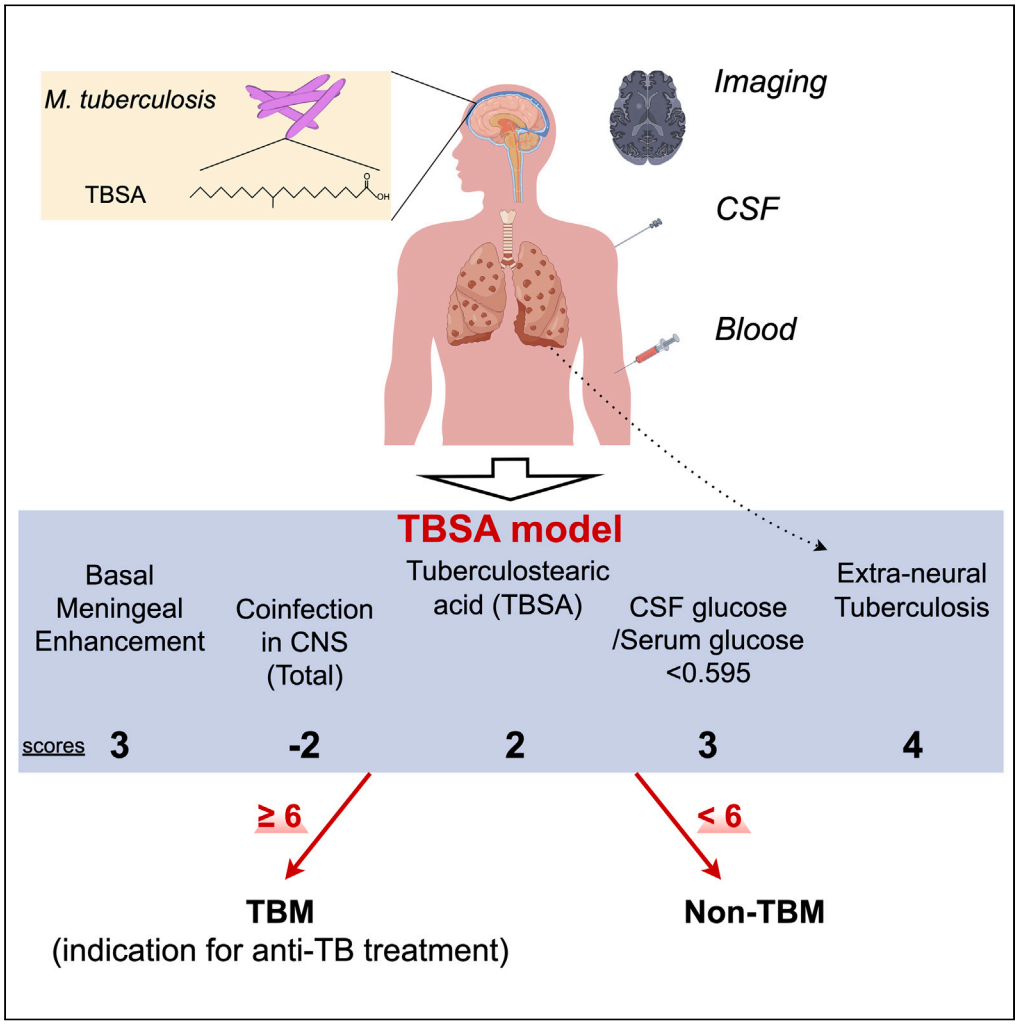
iScience

CellPress
OPEN ACCESS

Article

Tuberculostearic acid incorporated predictive model contributes to the clinical diagnosis of tuberculous meningitis



Tsz Hei Fong, Wangpan Shi, Guohui Ruan, ..., Chung Lam Ng, Yafang Hu, Haishan Jiang

jianghs@smu.edu.cn

Highlights

TBSA-combined scoring system shows high accuracy in TBM rule-in diagnosis

TBSA model aids early diagnosis of TBM for timeliness anti-TB treatment initiation

Model including TBSA confers solid reliability comparing traditional scoring system

Five-parameters model was validated by SVM modeling and 10-fold cross-validation

Fong et al., iScience *26*, 107858
October 20, 2023
https://doi.org/10.1016/j.isci.2023.107858

iScience

CellPress
OPEN ACCESS

Article

Tuberculostearic acid incorporated predictive model contributes to the clinical diagnosis of tuberculous meningitis

Tsz Hei Fong,[1,5,7] Wangpan Shi,[2,6,7] Guohui Ruan,[3,7] Siyi Li,[2,7] Guanghui Liu,[1] Leyun Yang,[4] Kaibin Wu,[3] Jingxian Fan,[1] Chung Lam Ng,[2] Yafang Hu,[1] and Haishan Jiang[1,8,*]

SUMMARY

The conventional confirmation tests of tuberculous meningitis (TBM) are usually low in sensitivity, leading to high TBM mortality. Hence, sensitive methods for indicating the presence of bacilli are required. Tuberculostearic acid (TBSA), a constituent from Mycobacterium *tuberculosis* had been evaluated as a promising marker, but fails to demonstrate consistent results for definite TBM. This study retrospectively reviewed medical records of 113 TBM suspects, constructing a TBSA-combined scoring system based on multiple factors, which show sensitivity and specificity of 0.8148 and 0.8814, respectively, and the area under the receiver operating characteristic curve of 0.9010. Multivariate analyses revealed four co-predictive factors strongly associated with TBSA: extra-neural tuberculosis, basal meningeal enhancement, CSF glucose/Serum glucose <0.595, and coinfection in CNS (Total). The subsequent machine learning-based validation showed correspondent importance to factors in the TBSA model. This study demonstrates a simple scoring system to facilitate TBM prediction, yield reliable diagnoses and allow timely treatment initiation.

INTRODUCTION

Tuberculous meningitis (TBM) is one of the most life-threatening central nervous system (CNS) infectious diseases caused by extrapulmonary Mycobacterium *tuberculosis* (Mtb) infection.[1] In spite of the fact that TBM accounts for about 1–1.8%.[2–4] of tuberculosis (TB) in HIV-negative people, its mortality rate reaches 30–40%.[4,5] Clinicians frequently fail to order anti-tubercular (anti-TB) treatment because of unspecific symptoms and difficulties in identifying the Mtb at an early stage.[1,6] In addition, the gold-standard tests (acid-fast bacilli microscopy, Mtb cultures, and nucleic acid amplification tests (NAATs)) used to detect Mtb hold limitations in confirming this fatal disease in a timely manner.[7–9] In the absence of early clinical intervention, patients often end up in a coma, severely disabled, or even die. Nevertheless, TBM can be treated successfully with prompt diagnosis and treatment, resulting in a 95% recovery rate.[1,6]

Hence, a scoring system that could indicate the probability of TBM was proposed by an international TBM workshop in 2010,[10] herein referred as the Lancet scoring system, aiming at facilitating clinical research for better patient care by categorizing patients as Definite, Probable, Possible, and Not TBM. However, in a recent study, Sulaiman et al.[11] revealed unsatisfactory performance of the Lancet system in distinguishing true TBM and non-TBM patients. It was found that concurrently using another scoring system, such as Thwaites' scoring system, with the Lancet could greatly enhance differential diagnostic efficiency,[12] but an effective rule-in system remains developed.

In the past, while some indirect molecular indices such as Mtb-specific antibodies, and interferon-gamma release assays (IGRAs) level in cerebrospinal fluid (CSF)[13,14] have been reported to be valuable for TBM diagnosis, yet a near-direct biomarker, tuberculostearic acid (TBSA), was overlooked in the past two decades,[15] which showed 83.3%–100% sensitivity and 98.9%–100% specificity, and could be rapidly detected by gas-chromatography/mass spectrometry (GC/MS) in 1–2 days from last century's studies.[16–19] It was known that TBSA, a mycobacterial origin lipid marker, is currently thought as a good candidate to indicate Mtb burden by lipidomics within a day in non-CNS infection,[20] but unexpectedly, not in TBM.

We previously found that positive TBSA results shared a similar distribution of definite TBM patients classified by the Lancet scoring system,[15] suggesting a promising role of TBSA to improve TBM diagnosis. In the present study, we sought to integrate the benefits of TBSA with

[1]Department of Neurology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China
[2]The First Clinical Medical School, Southern Medical University, Guangzhou 510515, China
[3]School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
[4]Department of Biostatistics, Bioinformatics, and Biomathematics, Georgetown University, Washington, DC 20057-1484, USA
[5]Present address: School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen 518055, China
[6]Present address: Department of Pathology, University of California San Diego Health System, San Diego 92103, CA, USA
[7]These authors contributed equally
[8]Lead contact
*Correspondence: jianghs@smu.edu.cn
https://doi.org/10.1016/j.isci.2023.107858

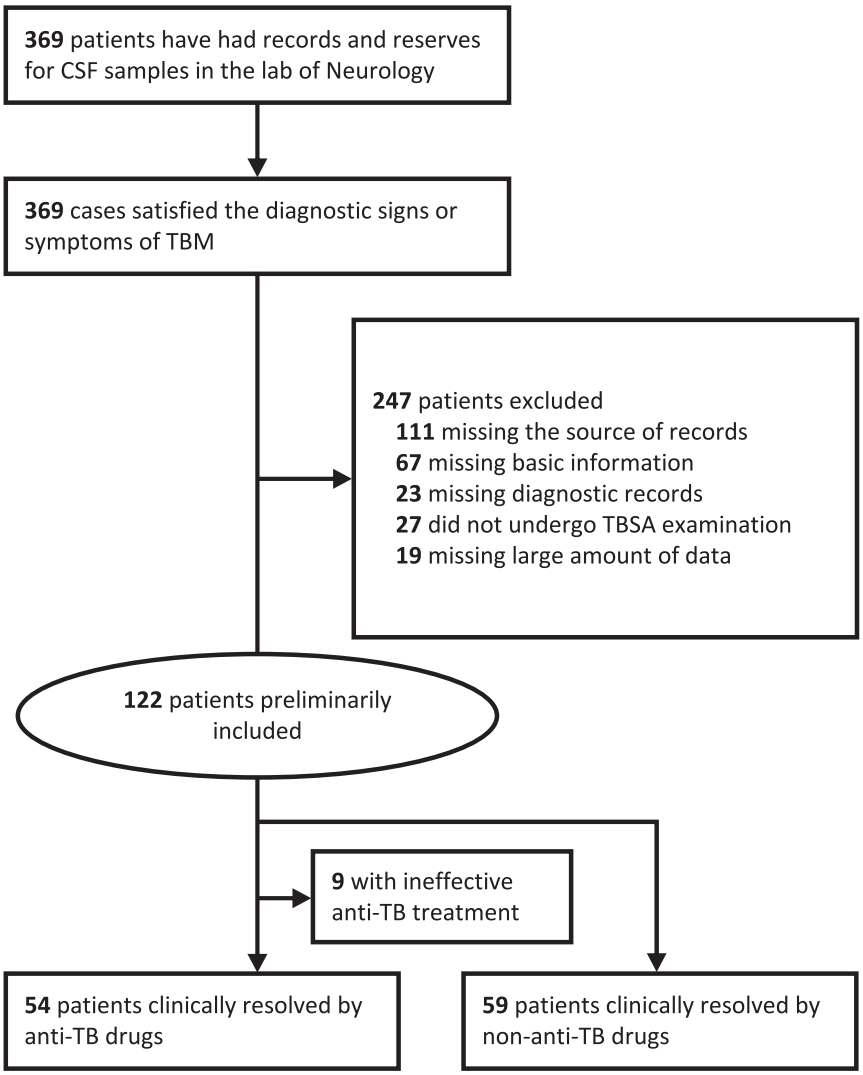


Figure 1. Study flowchart and grouping procedure of suspected TBM patients

CSF = cerebrospinal fluid, TBM = tuberculous meningitis, TBSA = tuberculostearic acid, TB = tubercular.

multiple diagnostic criteria to build up a diagnostic system, aiming to overcome the insufficient sensitivity of gold-standard tests. As we know that, constructing a scoring system based on gold-standard tests limited its diagnostic accuracy in the actual clinical situation where TBM confirmation tests showed low sensitivity to certain patients. In order to provide a better TBM rule-in strategy, an alternative for patients grouping is based on their prescribed regimens and treatment outcomes in the real clinical setting. Here, a multivariate logistic model, which was advantageous in constructing diagnostic scoring systems in various clinical settings,[21,22] would be applied to design a highly sensitive TBM diagnostic model. Concurrently, the support vector machine (SVM), which has been implemented in diagnoses and prognosis predictions in non-CNS tuberculous diseases,[23–25] was utilized as a maximum spacing classifier to accurately sort two independent cohorts based on inputted features, and to further validate the classification model performance. Consequently, we developed a TBSA-combined scoring system with high sensitivity to compensate for the experiences-based clinical judgements and sensitivity-limited gold-standard diagnostic bias in early TBM, in order to provide a prompt therapeutic strategy.

RESULTS

Clinical characteristics of patients

In our study, clinical records of 369 patients (ages ≥ 11) were preliminarily reviewed from the suspected TBM patients between January 1, 2009 and January 31, 2019. A total of 256 patients were excluded, including 247 with incomplete data and nine with poor response to TBM treatment. In the end, 113 patients who had positive treatment outcomes were included in the analysis and they were divided into the "Clinical resolution by anti-TBM treatment (CRA, n = 54)" group and the "Clinical resolution by non-TBM treatment (CRnA, n = 59)" (Figure 1). The

median age of patients in the CRA and CRnA group was 39 (Interquartile range 32–51) and 43 (27–58) years, respectively. The parameters utilized for evaluating patients' manifestation are shown in Table 1 and laboratory findings are in Table S1.

The presence of Mycobacterium *tuberculosis* was confirmed by microbiological and nucleic acid tests in 23 (20.4%) patients' cerebrospinal fluid (hereafter, definite TBM patients), among which 19 (17.6%), 3 (5.2%), and 3 (3.8%) positive results were found in Ziehl-Neelsen staining (patients received the test, n = 108), Mtb culture (n = 58), and NAAT (n = 79), respectively (Table S1). Tuberculostearic acid was detected in CSF of 39 (34.5%, n = 113) patients by GC/MS, whereas six definite TBM patients were found negative in TBSA. It should be noted that, two patients showing positive responses to non-*anti*-TBM treatments but detected positive in CSF Mtb were grouped into CRnA due to the severely lagging Mtb discovery after successful inpatient management. The resolution of symptoms may be attributed to their responses to corticosteroids.

Tuberculostearic acid scoring system

The classification model was developed based on the existing collected data containing a total of 70 parameters (Tables 1 and S1), showing that 33 CRA parameters were significantly different from the CRnA group ($p < 0.1$, Tables 1 and S1), which were therefore incorporated into the logistic regression. Univariate logistic analyses revealed 27 significant variables ($p < 0.1$) for the subsequent multivariate analyses which further found ten significant covariates ($p < 0.05$) associated with TBSA ($p < 0.001$, Table S2). The forward (conditional) multivariate logistic regression of these ten variables, together with TBSA, stepwise identified five significant predictors contributing to the classification between CRA and CRnA group, including TBSA ($p = 0.005$), extra-neural tuberculosis ($p = 0.0069$), basal meningeal enhancement (BME, $p = 0.0009$), CSF glucose/Serum glucose <0.595 ($p = 0.0012$), and coinfection in CNS (Total) ($p = 0.044$), which is referred as the TBSA model (Figure 2). The model included four risk factors (OR > 1.0) and one protective factor (OR < 1.0) regarding to the causation of TBM, and the score of each factor was separately assigned with the corresponding odds ratio (Figure 2).

Support vector machine model for validation

In addition, to verify the capacity of the TBSA model, a feature-based SVM model was trained using the 70 parameters (Tables 1 and S1), then visualized by PCA through the reduction of dimensionality to display effective TBM classification. The down-dimensioned model reflected partial information from the original SVM model, nonetheless, patients were shown to be distinctly separated into two groups (Figures 3 and S1). As shown in Figure 3, 84.6% of TBSA-positive patients (circled in the Figure) were sorted to the right side (i.e., therapeutically grouped TBM), implicating the distribution of TBSA along to the resolution of anti-TBM treatment (Figure 3).

Our SVM model renders 0.8704 of recall (values converged on 2-dimension: 0.7407), 0.9153 of specificity (0.8644), 0.9038 of precision (0.8333), 0.8852 of NPV (0.7846), and 0.8867 of F1 score (0.7843) for classification (Table S3), which was verified by 10-fold cross-validation before an optimal model was formed. Each divided part of the features subset was tested and shown in terms of area under curve (AUC), sensitivity, and specificity (Figure S2), in which eight (AUC >0.8000) out of ten testing sets of data verified efficient training with solid consistency, indicating high robustness of the selected features for the classification model.

The features subset was generated by RFE-CV modeling with 52 features (Figure 4), representing the optimal number of features and the best model for CRA and CRnA classification. The sign of weight indicates the correlation between features and the presence of TBM (+: positive correlation, -: negative correlation). Our results revealed that those co-influencing factors established in the TBSA model above also contributed with greater weights in the validated SVM model (Figure 4), showing consistency between models and implicating the validity of the TBSA model.

Comparing diagnostic methodologies

We next examine the performances of the models established in this study and that of those previously published models. The predictive scores of 113 patients summated through the TBSA model were plotted as the ROC curve with an AUC 0.9010 (Figure 5). The TBSA model with a cut-off value 5.5 demonstrates a high sensitivity (0.8148, 95% CI: 0.6916–0.8962) and specificity (0.8814, 0.7748–0.9413), indicating that patients scoring at 6 or above present TBM (Figure 5). Moreover, the classification capabilities of the SVM model, Lancet scoring system, and Thwaites' scoring system were also evaluated through sensitivity, specificity, and AUC (Figure 5; and Table S3). Among all four models, the SVM model exhibits the best sensitivity (0.8704, 95% CI: 0.7558–0.9358), specificity (0.9153, 0.8165–0.9633), and AUC (0.9369, 0.8897–0.9842), while the Lancet and Thwaites' scoring system performed unsatisfactorily in classifying CRA group when compared to both SVM and TBSA model. The Lancet scoring system showed good specificity (0.8475, 95% CI: 0.7348–0.9176), which equates to using TBSA only, but low in sensitivity (0.6667, 0.5336–0.7776), as when solely using TBSA (0.5556, 0.4238–0.6800) for diagnosis (Table S3). The Thwaites' scoring system failed to differentiate CRA patients from CRnA in this study ($p = 0.3285$) due to multi-type coinfected patients included in both groups. Thus, the TBSA model exhibited comparable performance to the SVM model with remarkably fewer parameters.

Tuberculostearic acid model demonstrating sensitive tuberculous meningitis diagnosis

We further compared the distribution of patients who were categorized into the group requiring anti-TB treatments under four different grouping approaches: i) clinical outcome with the resolution of TBM symptoms (CRA), ii) positive results from gold-standard Mtb direct examinations (Definite TBM), iii) the "Probable" and "Possible" groups of patients under Lancet scoring system (Lancet scoring), and iv) patients with scores higher than the optimum cut-off value evaluated by the TBSA model.

Table 1. Clinical characteristics of patients clinically resolved by anti-TBM or non-*anti*-TBM treatments

Parameter	Clinical resolution by anti-TBM treatment (n = 54)	Clinical resolution by non-*anti*-TBM treatment (n = 59)	*p* Value
Age, median (IQR), y	39 (32–51)	43 (27–58)	0.9894
Age <18	1 (2)	5 (8)	0.1484
Female/Male (% of male)	13/41 (76)	22/37 (63)	0.1292
Smoke	16 (30)	14 (24)	0.4780
Clinical manifestation			
Duration of illness, median (95% CI), d	14 (0–29)	10 (0–24)	0.0386
Symptom duration of more than 5 days[a]	52 (96)	49 (83)	0.0224
Systemic symptoms suggestive of tuberculosis[a]	19 (35)	6 (10)	0.0014
History of recent[a]	4 (7)	0 (0)	0.0182
Focal neurological deficit (excluding cranial nerve palsies)[a]	5 (9)	6 (10)	0.8705
Cranial nerve palsy[a]	5 (9)	7 (12)	0.6534
Altered consciousness[a]	16 (30)	22 (37)	0.3894
Extra-neural tuberculosis	19 (35)	1 (2)	<0.0001
Coinfection in CNS (Total)	9 (17)	18 (31)	0.0848
Bacteria coinfection (CNS)	2 (4)	1 (2)	0.5070
Virus coinfection (CNS)	6 (11)	14 (24)	0.0792
Fungi coinfection (CNS)	2 (4)	3 (5)	0.7214
Coinfection outside CNS (Total)	10 (19)	21 (36)	0.0422
Bacteria coinfection (non-CNS)	5 (9)	13 (22)	0.0638
Virus coinfection (non-CNS)	4 (7)	4 (7)	0.8966
Fungi coinfection (non-CNS)	1 (2)	1 (2)	0.9496
Mycoplasma coinfection (non-CNS)	0 (0)	5 (8)	0.0287
Imaging results			
Hydrocephalus[a]	13 (24)	9 (15)	0.1219
Basal meningeal enhancement[a]	30 (56)	13 (22)	0.0002
Tuberculoma[a]	4 (7)	0 (0)	0.0333
Infarct[a]	15 (28)	12 (20)	0.3543
Pre-contrast basal hyperdensity[a,b]	0 (0)	0 (0)	NA
TB elsewhere			
Chest radiograph suggestive of active tuberculosis[a]	8 (15)	0 (0)	0.0022
CT/MRI/ultrasound evidence for tuberculosis outside the CNS[a]	13 (24)	1 (2)	0.0003
Positive Ziehl-Neelsen staining or CSF Mtb culture for extra-neural specimen[a,b]	0 (0)	0 (0)	NA
Positive NAAT for extra-neural tuberculosis[a,b]	0 (0)	0 (0)	NA
Lancet categories without TBM confirmation[c]			
Probable TBM	25 (46)	1 (2)	NA
Possible TBM	27 (50)	45 (76)	NA
Non TBM	2 (4)	13 (22)	NA
Treatments[c]			
Anti-tubercular treatment	54 (100)	0 (0)	<0.0001
Corticosteroids	15 (28)	22 (37)	0.2844

(*Continued on next page*)

Table 1. *Continued*

Parameter	Clinical resolution by anti-TBM treatment (n = 54)	Clinical resolution by non-*anti*-TBM treatment (n = 59)	*p* Value
Other supportive/bacterial/viral/fungal treatment	15 (28)	59 (100)	<0.0001
Intrathecal injection	1 (2)	0 (0)	0.2938

Values are median (IQR) or N (%).

IQR = interquartile range, CI = confidence intervals, NA = not applicable.

[a]Clinical parameters suggested by the Lancet scoring system.

[b]Parameters excluded due to no collected data.

[c]Treatments and Lancet categories not involved as the parameters in the prediction model.

As shown in Figure 6, CRA group showed the actual number of patients who undertook anti-TB medication, in which only 16 patients were simultaneously thought to need anti-TB treatments by the rest of the grouping methods. The Definite TBM group could determine the true TBM patients, but two (8.7%) of them were ruled out from receiving anti-TB therapies due to lagged-detected Mtb evidence and also not detected as TBSA positive along with lower scores (<6) in the TBSA model. The Lancet scoring identified 46 (46.9%) patients as TBM suspects needing anti-TB treatment apart from 52 (53.1%) CRA patients, showing a severely high false-positive rate. Fourty-four patients (86.3% of the TBSA model) identified based on the TBSA model overlapped with patients in CRA, including 42 overlapped with patients from the Lancet scoring. Additionally, the TBSA model demonstrates comparable results with the SVM model (positive prediction n = 52) which identified 47 (90.4%) patients overlapping with the CRA group.

Overall, through step-by-step feature selection and validation by the SVM model, we proposed a scoring system incorporating four clinical parameters along with positive-TBSA results, which demonstrated high sensitivity and accuracy to predict TBM, and therefore could aid the early diagnosis of TBM by overcoming the insufficient timeliness and effectiveness of direct Mtb examinations and reducing its strong dependence on clinicians' experience.

DISCUSSION

Tuberculous meningitis is a severe CNS infectious disease with particular diagnostic challenges due to the insensitivity of mycobacterial detection in CSF.[7,26] Uncertain or delayed diagnosis protracts the initiation of anti-TB therapy, which results in detrimental consequences and high mortality of TBM.[1,27] Therefore, TBM treatment should be initiated even without microbiological proof,[7,28] but it is necessary to identify those patients who will positively respond to anti-TB drugs in a timely manner before treatment commences. We developed a sensitive evaluation system to retrospectively define uncertain TBM cases with good responses to anti-TB treatment using sensitive near-direct molecular evidence, TBSA. Moreover, this scoring system could be used to speculate TBM patient cohorts by inferring Mtb existence for further clinical therapies.

In the mycobacterial plasma membrane (PM), palmitic (16:0), oleic (18:1), and tuberculostearic (19:0) acid constitute the major fatty acid components, in which TBSA is capable to regulate lateral membrane partitioning of intracellular membrane domain[29] and highly related to Mtb existence in host.[20] In this study, we considered the advantages of TBSA for sensitive (83.3%–100%) and specific (98.9%–100%) TBM diagnosis,[17–19,30] and avoided solely usage of TBSA results that may cause false-positive diagnosis.[31] Hence, co-diagnostic factors were further screened from multiple parameters. Despite several diagnostic factors were also proved to indicate TBM, such as adenosine deaminase,[32] procalcitonin,[33] Pandy's test,[34] serum sodium,[35] and cerebral infarction,[36] and so forth, these methods could not directly confirm Mtb existence in CSF and all of them were excluded in TBSA model by showing no strong correlation with TBSA. Therefore, TBSA and its strong related factors included together could greatly suggest the presence of Mtb in CSF.

The co-diagnostic factors we found in the TBSA model were all known as common clinical references for TBM evaluation, in which BME, extraneural tuberculosis, and CSF to serum glucose ratio are applied in the Lancet scoring system.[10] The cut-off value of CSF glucose/Serum glucose was redefined as less than 0.595 rather than 0.5 suggested in the Lancet scoring system, which represents the feature of patients with good anti-TBM treatment outcomes. Moreover, other coinfections in CNS were usually considered to eliminate TBM suspicion, but TBSA is suggested to be concerned in all suspected meningitis cases even if strong co-infected evidence was provided, especially in endemic areas where TB co-infection with other pathogens is observed. Consequently, TBSA model incorporated benefits from diverse approaches for TBM diagnosis.

Currently, emerging direct strategies for Mtb confirmation seemingly showed huge potential for future TBM diagnosis. NAAT and GeneXpert MTB/RIT Ultra demonstrate high sensitivity and specificity for TBM confirmation,[8] but several Chinese studies reported sensitivity of Xpert MTB/RIT Ultra or Xpert Ultra was not higher than 47% in diagnosing TBM,[37–39] which is different from the previous sensitive results.[40] Metagenomic next-generation sequencing (mNGS) and lipoarabinomannan (LAM) showed excellent specificity, but they hold the limitations of expensiveness and relatively low sensitivity (22–33%), respectively.[9,41] Nevertheless, the construction strategy of TBSA model provides an integrated option to compensate for single diagnostic defects.

Taking advantage of the powerful Lancet scoring system that is helpful in diagnosing TBM patients with high sensitivity,[12] most of the patients could be classified into a wide-range category, but the results of grouping, which extensively requires labor and expertise investigation, cannot impact the expert clinical judgment and subsequent treatment for suspects.[42] Surprisingly, 57% of Definite patients (n = 21) in CRA group were classified as Possible (n = 11) or Not (n = 1) TBM, which means clinicians might not be able to make appropriate treatment

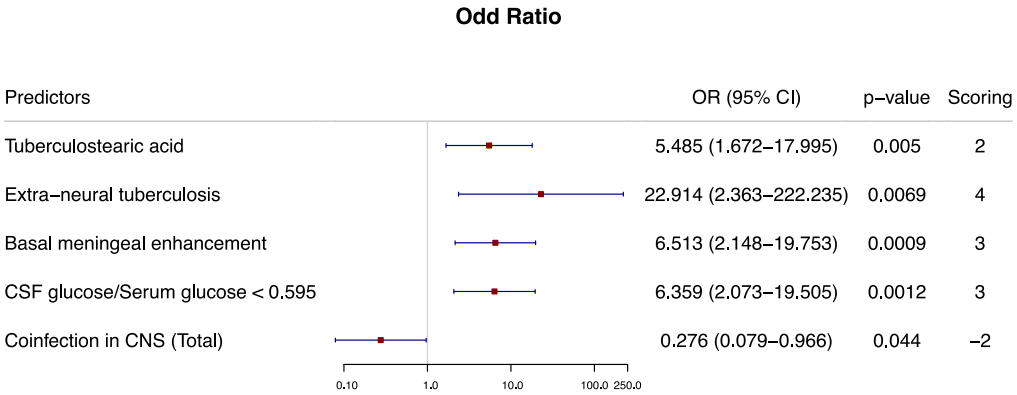


Figure 2. Multivariate logistic regression analysis of finally included variables for TBSA model

Odd ratio plot for factors associated with TBM shows tuberculostearic acid, extra-neural tuberculosis, basal meningeal enhancement, and CSF glucose/serum glucose <0.595 as the risk factors, as well as coinfection in CNS (Total) as the protective factor for TBSA model. The scores were assigned according to the OR value. OR = odds ratio, CI = confidence intervals, CSF = cerebrospinal fluid, CNS = central nervous system.

decisions for these patients if they were found negative in gold-standard results. However, positive TBSA demonstrated a similar distributive pattern of Definite TBM patients in the Lancet scoring system, which may indicate the highly suspected patients in the Possible or Not group.[15] TBSA model could find out 72% of Possible or Not TBM patients in CRA, in contrast to 59% using TBSA alone. Furthermore, the TBSA model was able to identify most (92.3%) of the Probable patients, showing this simpler and more practicable model with an improved capacity for TBM diagnosis. Owing to the TBSA model demonstrating a good rule-in sensitivity, it could be considered to apply together or after a differential diagnosis, for example the Thwaites' scoring system to discriminate TBM and bacterial meningitis.[43]

In order to verify the straightforward TBSA model, a more comprehensive and complicated SVM model was trained for TBM prediction, which showed higher precision and accuracy. The feature selection of SVM model was not based on the correlation with TBSA and showed its ranking of weights associated with the presence of TBM, in which co-diagnostic factors in the TBSA model all contribute great importance. When a risk factor was found as a parameter's value smaller than a cut-off value in logistic analyses, it would be shown as a negatively correlated feature in SVM model, such as CSF glucose, CSF glucose/Serum glucose, Serum Sodium, and so forth. As a result, the SVM model

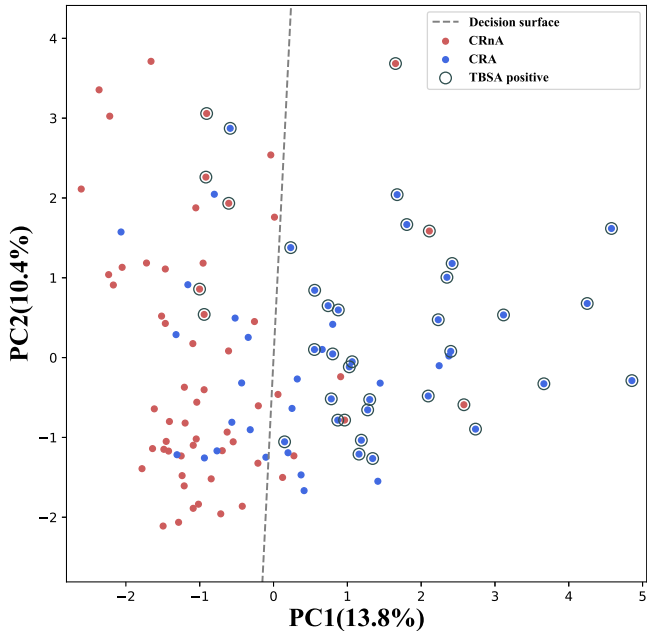


Figure 3. Principal components analysis for the 2-D demonstration of SVM classification model

The principal components analysis exhibits the down-dimensioned SVM classification in 2-D for clear visualization. The 2-D plane shows an obvious difference between CRA and CRnA, reflecting 24.3% information of high dimension SVM classification model; circled dots indicate the patients with positive result of TBSA detection. 2-D = 2-dimension, CRA = Clinical resolution by anti-TBM treatment, CRnA = Clinical resolution by non-*anti*-TBM treatment, SVM = support vector machine, PC = principal component, TBM = tuberculous meningitis, TBSA = tuberculostearic acid.

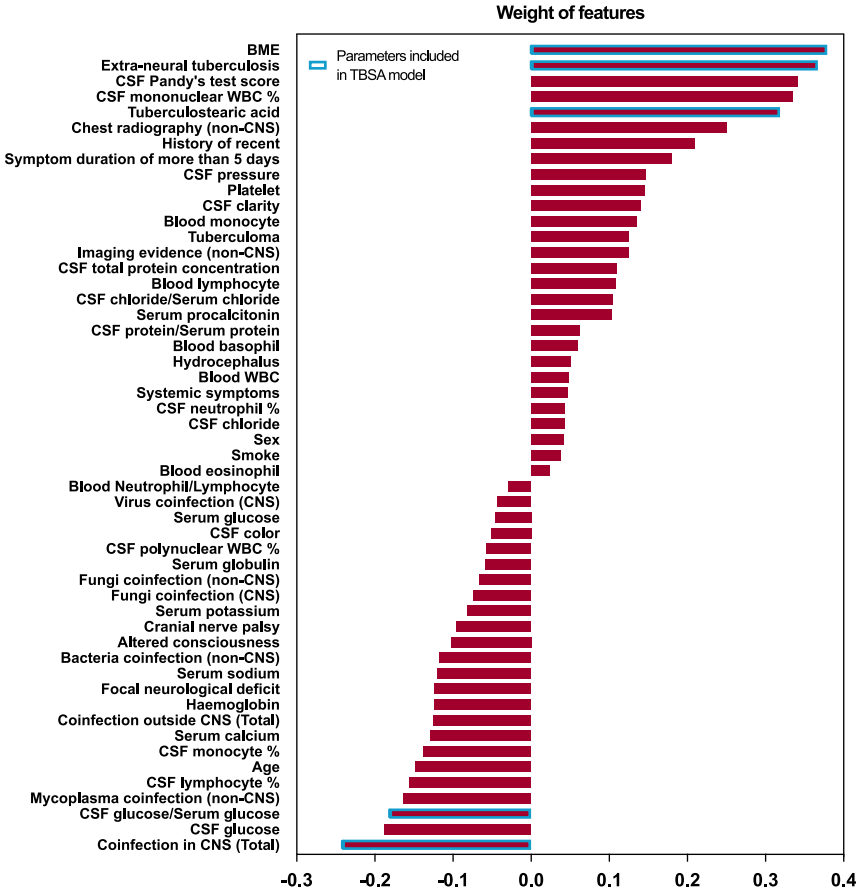


Figure 4. Significant features incorporated for SVM modeling with respective weighting

RFE-CV extracted 52 features demonstrated with weighting which contributes to the grouping of patients. The absolute value of the weight indicates the importance where the signs indicate the correlation to the presence of TBM (+: positive correlation, -: negative correlation). Five features incorporated in the TBSA model are displayed using the blue frame and share high weighting in the SVM model. TBSA = tuberculostearic acid, BME = basal meningeal enhancement, CSF = cerebrospinal fluid, CNS = central nervous system, WBC = white blood cell, RFE-CV = recursive feature elimination-cross validation.

validates features consistency between models but 52 features are required to achieve its optimum performance, justifying the TBSA model as a good fit for easy and rapid clinical application.

The grouping strategies in this study for models' construction are critical to imitate the realistic situation for TBM diagnosis. In the goal of defining TBM, concurrently considering the misdiagnosed patients excluded by insensitive gold-standard tests, we only preliminarily included patients with good treatment outcomes in order to establish distinctive groups with definite responses to anti-TBM or non-*anti*-TBM treatments. Due to the unaffordable cost to patients, none of the nine excluded patients with ineffective responses underwent multi-drug-resistant detection, so the drug resistance parameter was not considered in this study. Besides, in the view of the biosynthesis rate of TBSA from oleic acid by an enzyme Cfa that supports further research in revealing specific indications of different periods of Mtb infection, dormancy, and reactivation,[29] the grouping in this study implicates several potential considerations for further Mtb metabolic studies, especially the parameters highly correlated with TBSA levels, such as CSF chloride and CSF to serum chloride balance (Table S2).

In an effort to facilitate prompt diagnosis and early treatment for suspected TBM patients, a developed TBSA-combined scoring system with four co-predictors can infer the presence of Mtb, which was validated by an SVM model and accurately indicates the latent TBM patients from a suspected cohort when compared to contemporary commonly used scoring systems.

Limitations of the study

This study has several limitations. First, since the number of TBM patients available in non-specialist TB hospitals is limited, it was not practicable for independent cohort verification, thus 10-fold CV was used to verify the reliability of a predictive model, providing only algorithmic evidence for clinical reference. Second, due to the broad limitations of nowadays TBM confirmation tests,[9] we classified patients based on their regimen and outcome of treatments to reflect the real clinical situation, but this grouping strategy relied on professional clinical judgements and may exist a minor extent of bias compared to using gold-standard for patients grouping. In order to generate

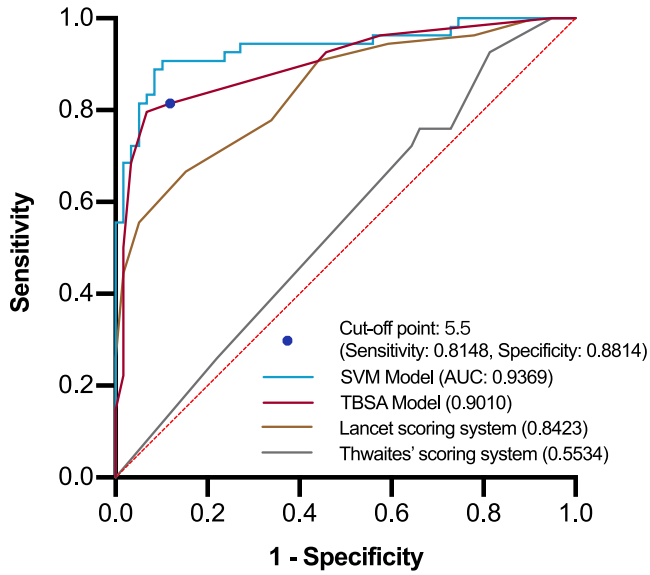


Figure 5. Receiver operating characteristic curve for different TBM diagnostic models

The ROC curves exhibit the diagnostic performance of the SVM model, TBSA model, Lancet and Thwaites' scoring system; the cut-off point shows the optimal threshold score (5.5) of TBSA model for classification.

more indisputable results, more confirmation tests will be applied in future studies. Third, the detection procedures for TBSA in CSF require up-to-date advanced protocols[44] to provide more sensitive results. Last, our model cannot predict patients with drug resistance or HIV/AIDS. Our scoring system is applicable to patients with uncertain TBM diagnoses, particularly in China, and should be prospectively studied in real-world validation with more diverse clinical settings which include patients with drug resistance or HIV co-infection, as well as other populations.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Study design, patients, and ethical approval
  - Inclusion and exclusion criteria
  - Therapeutic information
- METHOD DETAILS
  - Grouping strategies
  - Application of Clinical Scoring System
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Construction of a scoring system with TBSA
  - SVM classification model for validation
  - Statistical analysis
- ADDITIONAL RESOURCES
  - Trial registration

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107858.

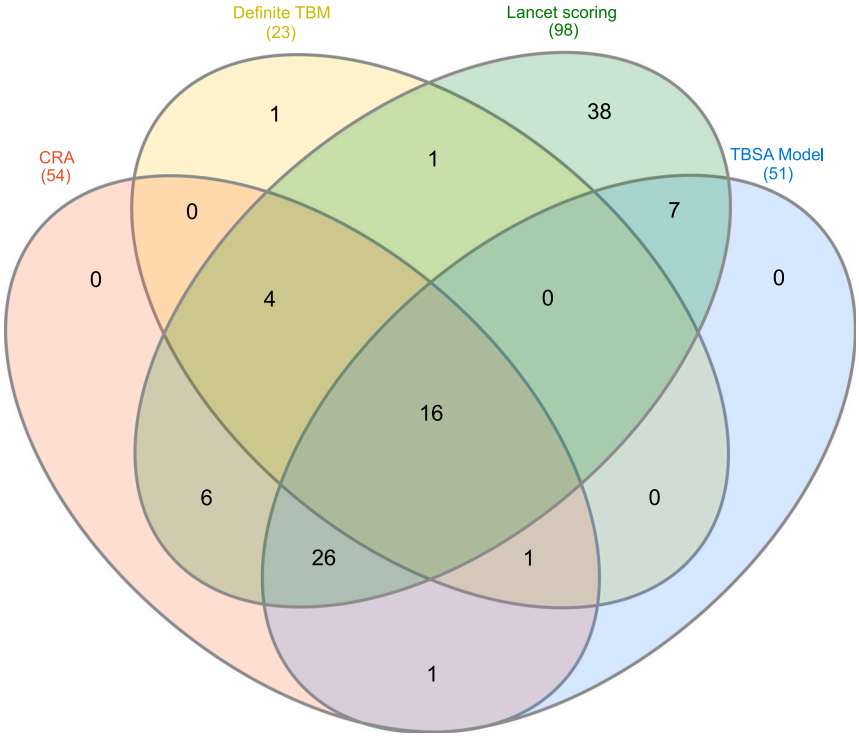


Figure 6. Venn diagram for the patients diagnosed as requiring anti-TB treatments by different grouping methods

The Venn diagram shows the patients diagnosed as requiring anti-TB treatments under four respective methods of (1) conventional TBM confirmation (Definite TBM): acid-fast bacilli, Mtb culture or NAAT, (2) Lancet scoring system (Lancet scoring): patients grouped in "Probable" or "Possible" TBM were regarded as requiring anti-TB treatments while not considering the results of conventional TBM confirmation, and (3) TBSA model, against the CRA group. CRA = Clinical resolution by anti-TBM treatment, TBM = tuberculous meningitis, TBSA = tuberculostearic acid.

ACKNOWLEDGMENTS

We thank Shengli An (Ph.D.) from Southern Medical University, Department of Biostatistics, Boyi Yang (Ph.D.) and Gongbo Chen (Ph.D.) from Sun Yat-Sen University, School of Public Health, and Xuewu Liu (M.D., Ph.D.) from Nanfang Hospital of Southern Medical University, Department of Neurology, for their valuable comments and input on this article. We thank Home for Researchers editorial team (www.home-for-researchers.com) for language editing service, and Figdraw (www.figdraw.com) for providing cartoon elements and platform in creating Graphical Abstract. This study was supported by the China International Medical Foundation [Grant No. Z-2016-20-1801] and Science and Technology Projects in Guangzhou, China [Grant No. 202102080451].

AUTHOR CONTRIBUTIONS

H.J., T.H.F., and W.S. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. T.H.F., W.S., G.R., and S.L. contributed equally to this work. T.H.F., S.L., W.S., and H.J. conceived and designed the study. T.H.F., S.L., G.L., and H.J. identified the appropriate data for the research. T.H.F., S.L., W.S., J.F., and C.L.N. conducted the data acquisition. T.H.F., G.R., K.W., and L.Y. performed data analyses and interpretation of the results. T.H.F., S.L., W.S., Y.H., and H.J. drafted and revised the article. H.J. supervised the study. All the authors revised and approved the final article before submission.

DECLARATION OF INTERESTS

The authors declare that they have no conflict of interest.

Received: March 21, 2023
Revised: August 5, 2023
Accepted: September 6, 2023
Published: September 9, 2023

REFERENCES

1. Schoeman, J.F., and Donald, P.R. (2013). Tuberculous meningitis. Handb. Clin. Neurol. *112*, 1135–1138. https://doi.org/10.1016/b978-0-444-52910-7.00033-7.
2. Thwaites, G., Fisher, M., Hemingway, C., Scott, G., Solomon, T., and Innes, J.; British Infection Society (2009). British Infection Society guidelines for the diagnosis and treatment of tuberculosis of the central nervous system in adults and children. J. Infect. *59*, 167–187. https://doi.org/10.1016/j.jinf.2009.06.011.
3. Ducomble, T., Tolksdorf, K., Karagiannis, I., Hauer, B., Brodhun, B., Haas, W., and Fiebig, L. (2013). The burden of extrapulmonary and meningitis tuberculosis: an investigation of national surveillance data, Germany, 2002 to 2009. Euro Surveill. *18*, 20436.
4. Dodd, P.J., Osman, M., Cresswell, F.V., Stadelman, A.M., Lan, N.H., Thuong, N.T.T., Muzyamba, M., Glaser, L., Dlamini, S.S., and Seddon, J.A. (2021). The global burden of tuberculous meningitis in adults: A modelling study. PLOS Glob. Public Health *1*, e0000069. https://doi.org/10.1371/journal.pgph.0000069.
5. Navarro-Flores, A., Fernandez-Chinguel, J.E., Pacheco-Barrios, N., Soriano-Moreno, D.R., and Pacheco-Barrios, K. (2022). Global morbidity and mortality of central nervous system tuberculosis: a systematic review and meta-analysis. J. Neurol. *269*, 3482–3494. https://doi.org/10.1007/s00415-022-11052-8.
6. Thwaites, G.E. (2013). Advances in the diagnosis and treatment of tuberculous meningitis. Curr. Opin. Neurol. *26*, 295–300. https://doi.org/10.1097/WCO.0b013e3283602814.
7. Wilkinson, R.J., Rohlwink, U., Misra, U.K., van Crevel, R., Mai, N.T.H., Dooley, K.E., Caws, M., Figaji, A., Savic, R., Solomons, R., et al. (2017). Tuberculous meningitis. Nat. Rev. Neurol. *13*, 581–598. https://doi.org/10.1038/nrneurol.2017.120.
8. Pormohammad, A., Nasiri, M.J., McHugh, T.D., Riahi, S.M., and Bahr, N.C. (2019). A Systematic Review and Meta-analysis of the Diagnostic Accuracy of Nucleic Acid Amplification Tests for Tuberculous Meningitis. J. Clin. Microbiol. *57*, e01113-18. https://doi.org/10.1128/jcm.01113-18.
9. Huynh, J., Donovan, J., Phu, N.H., Nghia, H.D.T., Thuong, N.T.T., and Thwaites, G.E. (2022). Tuberculous meningitis: progress and remaining questions. Lancet Neurol. *21*, 450–464. https://doi.org/10.1016/s1474-4422(21)00435-x.
10. Marais, S., Thwaites, G., Schoeman, J.F., Török, M.E., Misra, U.K., Prasad, K., Donald, P.R., Wilkinson, R.J., and Marais, B.J. (2010). Tuberculous meningitis: a uniform case definition for use in clinical research. Lancet Infect. Dis. *10*, 803–812. https://doi.org/10.1016/s1473-3099(10)70138-9.
11. Sulaiman, T., Medi, S., Erdem, H., Senbayrak, S., Ozturk-Engin, D., Inan, A., Civljak, R., Nechifor, M., Akbulut, A., Crisan, A., et al. (2020). The diagnostic utility of the "Thwaites' system" and "lancet consensus scoring system" in tuberculous vs. non-tuberculous subacute and chronic meningitis: multicenter analysis of 395 adult patients. BMC Infect. Dis. *20*, 788. https://doi.org/10.1186/s12879-020-05502-9.
12. Imam, Y.Z., Ahmedullah, H., Chandra, P., Almaslamani, M., Alkhal, A., and Deleu, D. (2020). Accuracy of clinical scoring systems for the diagnosis of tuberculosis meningitis in a case mix of meningitides a retrospective cohort study. J. Neurol. Sci. *416*, 116979. https://doi.org/10.1016/j.jns.2020.116979.
13. Philip, N., William, T., and John, D.V. (2015). Diagnosis of tuberculous meningitis: challenges and promises. Malays. J. Pathol. *37*, 1–9.
14. Török, M.E. (2015). Tuberculous meningitis: advances in diagnosis and treatment. Br. Med. Bull. *113*, 117–131. https://doi.org/10.1093/bmb/ldv003.
15. Fong, T.H., Shi, W., Li, S., Liu, G., Ng, C.L., and Jiang, H. (2021). Tuberculostearic acid, a potential parameter for scoring system construction for tuberculous meningitis diagnosis. Int. J. Mycobacteriol. *10*, 428–432. https://doi.org/10.4103/ijmy.ijmy_162_21.
16. Mårdh, P.A., Larsson, L., Høiby, N., Engbaek, H.C., and Odham, G. (1983). Tuberculostearic acid as a diagnostic marker in tuberculous meningitis. Lancet *1*, 367. https://doi.org/10.1016/s0140-6736(83)91676-8.
17. French, G.L., Teoh, R., Chan, C.Y., Humphries, M.J., Cheung, S.W., and O'Mahony, G. (1987). Diagnosis of tuberculous meningitis by detection of tuberculostearic acid in cerebrospinal fluid. Lancet *2*, 117–119. https://doi.org/10.1016/s0140-6736(87)92328-2.
18. Elias, J., De Coning, J.P., Vorster, S.A., and Joubert, H.F. (1989). The rapid and sensitive diagnosis of tuberculous meningitis by the detection of tuberculostearic acid in cerebrospinal fluid using gas chromatography-mass spectrometry with selective ion monitoring. Clin. Biochem. *22*, 463–467. https://doi.org/10.1016/s0009-9120(89)80099-2.
19. Brooks, J.B., Syriopoulou, V., Butler, W.R., Saroglow, G., Karydis, K., and Almenoff, P.L. (1998). Development of a quantitative chemical ionization gas chromatography-mass spectrometry method to detect tuberculostearic acid in body fluids. J. Chromatogr. B Biomed. Sci. Appl. *712*, 1–10. https://doi.org/10.1016/s0378-4347(98)00158-3.
20. Brandenburg, J., Heyckendorf, J., Marwitz, F., Zehethofer, N., Linnemann, L., Gisch, N., Karaköse, H., Reimann, M., Kranzer, K., Kalsdorf, B., et al. (2022). Tuberculostearic Acid-Containing Phosphatidylinositols as Markers of Bacterial Burden in Tuberculosis. ACS Infect. Dis. *8*, 1303–1315. https://doi.org/10.1021/acsinfecdis.2c00075.
21. Lee, S.A., Kim, S.W., Chang, H.H., Jung, H., Kim, Y., Hwang, S., Kim, S., Park, H.K., and Lee, J.M. (2018). A New Scoring System for the Differential Diagnosis between Tuberculous Meningitis and Viral Meningitis. J. Kor. Med. Sci. *33*, e201. https://doi.org/10.3346/jkms.2018.33.e201.
22. Solari, L., Soto, A., and Van der Stuyft, P. (2018). Development of a clinical prediction rule for tuberculous meningitis in adults in Lima, Peru. Trop. Med. Int. Health *23*, 367–374. https://doi.org/10.1111/tmi.13041.
23. Agranoff, D., Fernandez-Reyes, D., Papadopoulos, M.C., Rojas, S.A., Herbster, M., Loosemore, A., Tarelli, E., Sheldon, J., Schwenk, A., Pollok, R., et al. (2006). Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum. Lancet *368*, 1012–1021. https://doi.org/10.1016/s0140-6736(06)69342-2.
24. Gao, X.W., and Qian, Y. (2018). Prediction of Multidrug-Resistant TB from CT Pulmonary Images Based on Deep Learning Techniques. Mol. Pharm. *15*, 4326–4335. https://doi.org/10.1021/acs.molpharmaceut.7b00875.
25. Li, C., Hou, L., Sharma, B.Y., Li, H., Chen, C., Li, Y., Zhao, X., Huang, H., Cai, Z., and Chen, H. (2018). Developing a new intelligent system for the diagnosis of tuberculous pleural effusion. Comput. Methods Progr. Biomed. *153*, 211–225. https://doi.org/10.1016/j.cmpb.2017.10.022.
26. Slane, V.H., and Unakal, C.G. (2023). Tuberculous Meningitis. In StatPearls (StatPearls Publishing Copyright © 2023, StatPearls Publishing LLC.).
27. Schaller, M.A., Wicke, F., Foerch, C., and Weidauer, S. (2019). Central Nervous System Tuberculosis : Etiology, Clinical Manifestations and Neuroradiological Features. Clin. Neuroradiol. *29*, 3–18. https://doi.org/10.1007/s00062-018-0726-9.
28. Luo, M., Wang, W., Zeng, Q., Luo, Y., Yang, H., and Yang, X. (2018). Tuberculous meningitis diagnosis and treatment in adults: A series of 189 suspected cases. Exp. Ther. Med. *16*, 2770–2776. https://doi.org/10.3892/etm.2018.6496.
29. Prithviraj, M., Kado, T., Mayfield, J.A., Young, D.C., Huang, A.D., Motooka, D., Nakamura, S., Siegrist, M.S., Moody, D.B., and Morita, Y.S. (2023). Tuberculostearic Acid Controls Mycobacterial Membrane Compartmentalization. mBio *14*, e0339622. https://doi.org/10.1128/mbio.03396-22.
30. Daikos, G.L., Brooks, J.B., Michos, A., Roma, E., and Syriopoulou, V. (2004). Detection of tuberculostearic acid in serum and other biological fluids from patients with tuberculosis by electron capture-gas chromatography and chemical ionisation-mass spectrometry. Int. J. Tubercul. Lung Dis. *8*, 1027–1031.
31. Heap, B.J. (1991). False-positive tuberculostearic acid tests in cysticercosis. Lancet *337*, 498–499. https://doi.org/10.1016/0140-6736(91)93440-k.
32. Ye, Q., and Yan, W. (2023). Adenosine deaminase from the cerebrospinal fluid for the diagnosis of tuberculous meningitis: A meta-analysis. Trop. Med. Int. Health *28*, 175–185. https://doi.org/10.1111/tmi.13849.
33. Kim, J., Kim, S.E., Park, B.S., Shin, K.J., Ha, S.Y., Park, J., Kim, S.E., and Park, K.M. (2016). Procalcitonin as a Diagnostic and Prognostic Factor for Tuberculosis Meningitis. J. Clin. Neurol. *12*, 332–339. https://doi.org/10.3988/jcn.2016.12.3.332.
34. Zhang, W., Guo, X.X., Wang, X.F., and Wang, M.S. (2016). Cerebrospinal fluid IgG against TB-SA for diagnosis of tuberculous meningitis. Eur. Rev. Med. Pharmacol. Sci. *20*, 462–466.
35. Na, S., Kim, T., Song, I.U., Chung, S.W., Kim, S.H., Oh, Y.S., Oh, J., and Kim, W. (2021). The association between serum sodium level and tuberculous meningitis compared with viral and bacterial meningitis. Sci. Rep. *11*, 10906. https://doi.org/10.1038/s41598-021-90358-5.
36. Wasay, M., Farooq, S., Khowaja, Z.A., Bawa, Z.A., Ali, S.M., Awan, S., Beg, M.A., and Mehndiratta, M.M. (2014). Cerebral infarction and tuberculoma in central nervous system tuberculosis: frequency and prognostic implications. J. Neurol. Neurosurg. Psychiatry *85*, 1260–1264. https://doi.org/10.1136/jnnp-2013-307178.

37. Shao, L., Qiu, C., Zheng, L., Yang, Y., Yang, X., Liang, Q., Zhang, Y., Che, N., Pang, Y., and Duan, H. (2020). Comparison of diagnostic accuracy of the GeneXpert Ultra and cell-free nucleic acid assay for tuberculous meningitis: A multicentre prospective study. Int. J. Infect. Dis. *98*, 441–446. https://doi.org/10.1016/j.ijid.2020.06.076.

38. Sun, L., Yao, L., Fu, G., Lin, L., Zhu, E., and Huang, J. (2021). A comparison of the accuracy of the CapitalBio Mycobacterium real-time polymerase chain reaction and the Xpert MTB/RIF assay for the diagnosis of tuberculous meningitis. Int. J. Infect. Dis. *104*, 92–96. https://doi.org/10.1016/j.ijid.2020.12.044.

39. Huang, M., Wang, G., Sun, Q., Jiang, G., Li, W., Ding, Z., Jia, H., Gao, M., Huang, H., and Li, Q. (2021). Diagnostic accuracy of Xpert MTB/RIF Ultra for tuberculous meningitis in a clinical practice setting of China. Diagn. Microbiol. Infect. Dis. *100*, 115306. https://doi.org/10.1016/j.diagmicrobio.2020.115306.

40. Donovan, J., Cresswell, F.V., Thuong, N.T.T., Boulware, D.R., Thwaites, G.E., and Bahr, N.C.; Tuberculous Meningitis International Research Consortium (2020). Xpert MTB/RIF Ultra for the Diagnosis of Tuberculous Meningitis: A Small Step Forward. Clin. Infect. Dis. *71*, 2002–2005. https://doi.org/10.1093/cid/ciaa473.

41. Wang, S., Chen, Y., Wang, D., Wu, Y., Zhao, D., Zhang, J., Xie, H., Gong, Y., Sun, R., Nie, X., et al. (2019). The Feasibility of Metagenomic Next-Generation Sequencing to Identify Pathogens Causing Tuberculous Meningitis in Cerebrospinal Fluid. Front. Microbiol. *10*, 1993. https://doi.org/10.3389/fmicb.2019.01993.

42. Kurien, R., Sudarsanam, T.D., Samantha, S., and Thomas, K. (2013). Tuberculous meningitis: a comparison of scoring systems for diagnosis. Oman Med. J. *28*, 163–166. https://doi.org/10.5001/omj.2013.47.

43. Thwaites, G.E., Chau, T.T.H., Stepniewska, K., Phu, N.H., Chuong, L.V., Sinh, D.X., White, N.J., Parry, C.M., and Farrar, J.J. (2002). Diagnosis of adult tuberculous meningitis by use of clinical and laboratory features. Lancet *360*, 1287–1292. https://doi.org/10.1016/s0140-6736(02)11318-3.

44. Arslan, F.N., Kolk, A., and Janssen, H.G. (2019). Methods for one- and two-dimensional gas chromatography with flame ionization detection for identification of Mycobacterium tuberculosis in sputum. J. Chromatogr., B: Anal. Technol. Biomed. Life Sci. *1124*, 204–217. https://doi.org/10.1016/j.jchromb.2019.06.012.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
IBM→ SPSS→ Statistics Version 25.0	IBM, Chicago, IL, USA	https://www.ibm.com/products/spss-statistics
GraphPad Prism 8 for macOS Version 8.2.1 (279)	GraphPad Software	https://www.graphpad.com/
RStudio Version 1.1.456	RStudio Inc	N/A
Python Version 3.7.6	Python Software Foundation	https://www.python.org/
Scikit-learn toolkit Version 0.21.2	Python Software Foundation	https://pypi.org/project/scikit-learn/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Haishan Jiang (jianghs@smu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The complete original data reported in this study cannot be deposited in a public repository because these data are confidential medical records. The dataset used and analysed during the current study is available upon reasonable request, contact Mr. Tsz Hei Fong (andrewfongth@163.com).
- This paper does not report original code.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study design, patients, and ethical approval

The patient cohort in this retrospective study was selected from the suspected TBM patients, who have had records and reserves for cerebrospinal fluid samples, over ten years from January 1, 2009 to January 31, 2019 at the First affiliated hospital of Southern Medical University (i.e., Nanfang Hospital), Guangzhou, China. The criteria for TBM clinical suspicion, microbiology diagnostic workup, and empirical therapies were established to handle patients presented with stiff neck, headache, fever, or vomiting in addition to one of the following indicative signs: 1) more than 5 days of disease course, 2) neurologic symptoms, and 3) cranial nerve palsies. Patients' data was extracted from the electronic medical record system, including clinical manifestation, results of microbiological examinations (Ziehl-Neelsen staining, Mtb culture, or NAAT), TBSA detection, brain radiological analysis, CSF analysis, and blood laboratory findings. All included participants' data were from Chinese patients (n = 113; males = 78, 69.0%; children = 6, 5.3%; mean age = 41.3 years). This study included no patients with HIV co-infection. This study was approved by the Medical Ethics Committee of Nanfang Hospital to collect and use patients' data for scientific research with informed consent (approval letter no.NFEC-2018-087).

Inclusion and exclusion criteria

In the current study, a total of 369 suspected TBM patients were preliminarily reviewed, all of whom met the clinical entry criteria suggested by the Lancet consensus of TBM diagnosis (headache, irritability, vomiting, fever, neck stiffness, convulsions, focal neurological deficits, altered consciousness, or lethargy).[10] Patients with missing records, basic information, diagnostic details, TBSA results, or a large amount of laboratory examination data (more than four parameters for each patient) were then excluded.

Therapeutic information

According to the therapeutic criteria of Nanfang Hospital, patients with mycobacterial confirmed or highly clinically suspected TB firstly received four first-line drugs (rifampicin, isoniazid, pyrazinamide, and ethambutol). In cases the first-line anti-TB treatment failed, linezolid, fluoroquinolone, or levofloxacin were prescribed subsequently. Patients with moderate or severe TBM accompanied with inflammatory reactions were prescribed with corticosteroids. Suspected TBM patients found to have other infections or confirmed TBM patients with

co-infections were both prescribed correspondent medications, including antiviral drugs, or antibiotics, etc. For retrospective assessment of TBM diagnostic accuracy, we compared therapeutic histories and outcomes with diagnostic results for each patient. Additionally, only the results of diagnostic tests and clinical evaluations prior to treatment were collected and used in the study.

METHOD DETAILS

Grouping strategies

Based on the type of therapy received, the finalized patient cohort was divided into two groups. Patients resulting in resolution of TBM symptoms (including night sweats, weight loss, persistent cough, lethargy, altered consciousness, headache, and fever) treated with anti-TB medications and treated without anti-TB drugs were, respectively, categorized into the "Clinical resolution by anti-TBM treatment (CRA)" group and the "Clinical resolution by non-TBM treatment (CRnA)" group. Importantly, CRnA patients had either been directly or indirectly confirmed to be free of Mtb infection based on clinical, laboratory, or radiologic evidence, as well as positive responses to non-TBM therapies. The initial and subsequent responses of both CRA and CRnA patients to therapy were evaluated throughout the inpatient course. Patients with ineffective medical treatment were then removed from the study due to interruption of treatment, treatment outcome not conforming to the course of the relative disease, comorbidity or death with a critical illness. Therefore, the patients showing definite treatment responses in CRA and CRnA groups support the categorization of TBM and non-TBM patients in this study, respectively.

Application of Clinical Scoring System

The Lancet scoring system for categorizing patients with suspected TBM was recommended for use in clinical research, dividing them into groups of "Definite", "Probable", "Possible", and "Not" TBM. Additionally, Thwaites' scoring system was utilized to differentiate TBM from bacterial meningitis by five clinical variables, including age (years) $\geq$ 36, blood white-cell count (10^3/mL) $\geq$ 15000, history of illness (days) $\geq$ 6, CSF total white-cell count (10^3/mL) $\geq$ 900, and CSF % neutrophils $\geq$ 75. Our classification methods were compared with those from the Lancet and Thwaites for further evaluation.

QUANTIFICATION AND STATISTICAL ANALYSIS

Construction of a scoring system with TBSA

To establish a scoring system, we used a multistep procedure to identify the best-fitted predictors for the logistic model. First,we extracted variables that significantly differed between the CRA and CRnA groups using different statistical tests for various types of variables: the student T-tests for normally distributed variables, the Mann-Whitney tests for non-normally distributed variables and categorical variables, and the Chi-square tests for dichotomous variables (all continuous variables underwent a normality test by examining the Q-Q/P-P plot). Only variables survived at $P < 0.1$ would be considered in the following logistic regression analyses. Before entering next step, all primarily included continuous variables were converted into dichotomized forms using optimum cut-off value determined by Youden's index, separately for each variable. Then univariate logistic analyses were applied to identify significant variables ($P < 0.1$) as covariates to be included in the subsequent multivariate analyses. These covariates were inputted with TBSA in the multivariate logistic model to select the variables ($P < 0.05$) that significantly associated with TBSA ($P < 0.001$). Based on the variables selected from the previous step, including TBSA, a stepwise forward (conditional) logistic regression was performed to finally identify the key predictors for CRA and CRnA classification.

We scored each identified variable based on its odds ratio (OR) presented with 95% confidence interval (CI). The patients' final scores were calculated by summing the scores of all variables. Optimum cut-off point for the prediction algorithm was identified using the receiver operating characteristic (ROC) curve. Sensitivity (i.e. recall), specificity, positive predictive value (PPV, or precision), negative predictive value (NPV), and F_1 score of our model were compared with those of the existing scoring system.

SVM classification model for validation

Patients' data were input into a linear kernel SVM and divided into two classes by identifying a hyperplane defined by $\boldsymbol{w}^T\boldsymbol{x} + b = \mathbf{0}$, where $\boldsymbol{w}$ is the weight vector, $\boldsymbol{x}$ is the feature vector and b is the bias term, in order to construct a classification model for sorting patients into "CRA" or "CRnA" based on features weighting. Missing values in the data were filled with the mean of the corresponding variable of remaining patients.

Feature selection using Recursive Feature Elimination (RFE) was performed to assign weights (coefficients of a linear SVM) and rank the importance of features in a current feature set (the initial set included all available features). The least important feature (or features) was/were pruned recursively per loop of modelling until all features were ranked. The subsets were then derived and comprised of different number of features from each loop of RFE. The k-fold (k = 10) cross-validation (CV) based on RFE was used to validate the performance of each subset model with different features, selecting the best feature set corresponding to the highest mean value of metric (F_1 score) of 10 folds. The feature number in the best-validated set represents the optimal number of features in the SVM model. L2 regularization was applied in the above RFE-CV modelling to avoid overfitting. The final SVM model with a specific feature set was accomplished by training the whole dataset to show the maximal classification interval.

For visualization of the classification performance of the SVM model, we reduced the dimensionality of the model using principal component analysis (PCA). Down-dimensioned data reflected approximately 20-30% information of the original parameters and exhibited the classification capability of the SVM model.

Statistical analysis

All statistical analyses in this study were performed with IBM→ SPSS→ Statistics (Version 25), GraphPad Prism 8 for macOS (Version 8.2.1 (279)), and RStudio (Version 1.1.456). SVM and PCA were implemented in Python (Version 3.7.6) using Scikit-learn (Version 0.21.2) toolkit for machine learning.

ADDITIONAL RESOURCES

Trial registration

This study was retrospectively registered with the Chinese Clinical Trial Registry (Registration number: ChiCTR1900028161, https://www.chictr.org.cn/showprojEN.html?proj=46907) on 14 December 2019.