

Original Article

The accuracy of tuberculous meningitis diagnostic tests using Bayesian latent class analysisHuy Ngoc Le¹, Hutch Sripilung², Virasakdi Chongsuvivatwong², Nhung Viet Nguyen¹, Tri Huu Nguyen¹¹ Vietnam National Lung Hospital, Hanoi, Vietnam² Epidemiology Unit, Prince of Songkla University, Hat Yai, Songkhla, Thailand**Abstract**

Introduction: Tuberculous meningitis (TBM) is the most dangerous form of tuberculosis with high mortality and disability rates. However, the delayed diagnostic process is often due to the absence of the gold standard tests leading to a lack of information about the sensitivity and specificity of diagnostic tests. This study aims to estimate the prevalence of TBM and determine the performance of four diagnostic procedures: the mycobacteria growth culture test, Gene Xpert assay, and analysis of protein levels and leukocyte count taken from cerebrospinal fluid.

Methodology: We used a Bayesian latent class analysis to estimate the prevalence of TBM with 95% credible interval (CI), and the specificity and sensitivity of the four diagnostic procedures. The area under the receiver operating characteristic curve (AUC) of the cerebrospinal protein levels and leukocyte count were also compared and estimated using different thresholds.

Results: A total of 1,213 patients suspected of having TBM were included. The estimated TBM prevalence was 34.8 % (95% CI: 28.8 – 41.3). The sensitivity of culture test and Gene Xpert assay was 62.7% (95% CI: 52.5 – 74.0), and 57.5% (95% CI: 51.0 – 64.0), and the specificity of Gene-Xpert was 95.9% (95% CI: 92.0 – 99.8). The AUC for leukocyte count was 76.0%, and for protein level was 73.4%.

Conclusions: This study provided better information about the performance of four routine diagnostic tests and the prevalence of TBM which can enhance disease control and improve treatment outcomes.

Key words: Tuberculous meningitis; Gene Xpert assay; Bayesian latent class analysis.

J Infect Dev Ctries 2020; 14(5):479-487. doi:10.3855/jidc.11862

(Received 21 July 2019 – Accepted 10 March 2020)

Copyright © 2020 Le *et al.* This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Tuberculosis (TB) is a leading global health concern. According to the Tuberculosis Global Report 2017 published by the World Health Organization (WHO), it is one of the ten leading causes of death worldwide. In 2016, 10.4 million TB incident cases were reported with more than half occurring in South East Asia and Western Pacific Regions [1].

Tuberculous meningitis (TBM) is the most lethal manifestation of TB due to its impact on the nervous system [2]. The mortality rate of TBM was reported to be nearly 20% in the first month of treatment and also left 50% of patients living with severe neurological sequelae [3]. As it is a heterogeneous neurological disease with complicated and non-specific symptoms, there is no single test or symptom to confirm TBM. Hence, the diagnostic process is often delayed due to a lack of information about the accuracy of diagnostic tests and clinical symptoms.

Commonly, the diagnostic procedures are based on composite reference standards, which depended on clinical symptoms and laboratory features. Among

them, the most common diagnostic compositions are the Thwaites diagnostic score [4] and the Lancet score [5]. Studies have found that clinical scoring systems have high sensitivity but low specificity when used in non- Human Immunodeficiency virus (HIV) groups [6,7].

Cerebrospinal fluid culture test has been considered as the gold standard for diagnosis and also has essential roles in antimicrobial susceptibility testing. However, the laboratory analysis is time-consuming (4-6 weeks for negative results), and the sensitivity is moderate. A study from Indonesia reported that the sensitivity of the culture test was less than 50% when compared with clinical diagnosis [8].

In 2012, the Gene-Xpert MTB/RIF assay (Cepheid, USA), a nucleic-based test, was recommended by WHO as an initial diagnostic test for diagnosing extra-pulmonary tuberculosis, especially TBM. The test has a sensitivity and specificity of 59.3% and 99.5%, respectively [9]. However, most of these studies considered the culture test as the gold standard, leading to imperfect reference bias towards low sensitivity, a

commonly encountered problem in the field of diagnostic studies. These problems have been discussed in a meta-analysis [10].

Hence, more information and knowledge about the estimation of the sensitivity and specificity of TBM diagnostic tests are crucial to have prompt diagnostic algorithms and suitable treatments, which can decrease mortality and improve patient outcomes [11]. As there has been no report on the sensitivity and specificity of laboratory tests in TBM in which the gold standard using biopsy and culture test is not ethically in all cases, especially in cases with negative test results. The main objectives of this study were to describe the characteristics and the laboratory features of TBM patients and evaluating the sensitivity and specificity of the TBM diagnostic tests, and the prevalence of TBM disease in Vietnam using the Bayesian latent class analysis approach. The main objectives of this study were to describe the characteristics and the laboratory features of TBM patients and evaluate the sensitivity and specificity of the TBM diagnostic tests. Thus, the prevalence of TBM in Northern Vietnam could be estimated with confidence. Furthermore, the results of this study would potentially give better and more precise understanding and information for the development of an appropriate diagnostic algorithm and would improve the treatment outcomes.

Methodology

Study design

We conducted a cross-sectional observational study conducted from November 2016 to April 2019.

Study setting

We collected data from patients treated at Vietnam National Lung Hospital, the largest hospital for tuberculosis and respiratory diseases in Northern Vietnam. All patients suspected of having tuberculosis meningitis (TBM) who had cerebrospinal fluid collected by lumbar puncture immediately after hospital admission and before anti-biotic usage were included. The quality of the specimens was controlled carefully and followed the standard guidelines. We tried to do the lumbar puncture as soon as possible, frequently right after the patient's hospital admission and before drug usage. All the specimens were transferred to the microbiology department immediately after the procedure and examined by experienced technicians. A total of 1ml to 3 ml of cerebrospinal fluid (CSF) was collected. In several cases with ample CSF, we centrifuged CSF to increase positive result. The specimens were analyzed by experienced technicians to

ensure that the quality of the specimens and the test result were reliable. The samples were analyzed for chemical composition, leukocyte count, Xpert test, and Mycobacteria Growth Indicator Tube (MGIT) culture test. The culture and Xpert tests were done at the Tuberculosis National Reference Laboratory. Both of the diagnostic tests were performed based on the manufacturer's instructions. As has been pointed out in previous study, the experiences of laboratory technicians were essential factors for the result of the culture test [12]. We understand the importance of this issue in diagnosing infectious diseases, especially in severe disease such as tuberculous meningitis. In our hospital routine procedure, we try to minimize the biases by following standard procedures with experienced technicians, and also providing regular training for the laboratory workers. All the specimen tests were performed at the National Reference laboratory of Vietnam, where thousands of tests are done annually. The laboratory is the reference laboratory for the whole country and is responsible for quality assurance monitoring and providing technical support to all tuberculosis (TB) laboratories in Vietnam. In 2017, this laboratory analyzed nearly 7,000 samples using Xpert tests and achieved the ISO 17043 qualification.

The tests under evaluation

According to WHO guidelines for extrapulmonary tuberculosis diagnosis and treatment from 2012, the MGIT culture and the Xpert tests were added in the Vietnamese standard protocol for diagnosis of TBM. The Xpert test was proved to be a reliable test based on its rapid and automatic characteristics, and a positive result of the MGIT culture test is considered crucial for TBM confirmation. We considered 2 more diagnostic parameters: total cell counts and the chemical analysis of the diagnostic tests. Based on the meta-analysis of Marais [5], CSF cytology and chemical parameters play significant roles in detecting the existence of TBM with approximate cut-off values. The cut-off values of these parameters were 1g/L for protein level and 500 cells/mm³ for leukocyte count. We also examined a range of thresholds to see if it resulted in better interpretations of the clinical outcomes.

The chemical parameters included glucose, protein, and chloride levels. According to a previous study [5], protein and glucose levels were also used as diagnostic parameters for TBM diagnosis.

In this study, we used the protein level obtained from CSF as a chemical diagnostic marker of TBM. According to previous research, the average protein

concentration level in cerebrospinal fluid is less than 0.45 g/Lt [5,13]. However, a cut-off value for protein level in diagnosing TBM in the previous study ranged from 0.4g/Lt to 2g/Lt. We, therefore, fit the models with 4 thresholds (0.4g/Lt, 1g/Lt, 2g/Lt, and 3g/Lt). Similarly, based on a previous study, the cut-off values for total cell counts in cerebrospinal fluid ranged from 10 cells/mm³ to 500 cells/mm³ [5]. In this study, we chose the cut-off value for the leukocyte counts as 10 cells/mm³. Moreover, we used a range of cut-off values for the cell counts: 10, 20, 500, and 1000 cells/mm³ in 4 models.

For the Xpert and culture test, all the results which reported the presence of *Mycobacterium tuberculosis* in CSF specimens were defined as positive regardless of the bacteria loads.

Statistical analysis

Statistical software

Data were analyzed using R software version 3.5.2 (Foundation for Statistical Computing, Vienna, Austria). Descriptive statistics were used to summarize the characteristics of patients using the mean and standard deviation or median and interquartile range for continuous variables, and frequencies and percentages for categorical variables.

Bayesian latent class analysis (BLCA) was used to estimate the performance of diagnostic tests and the prevalence of TBM. WinBUGS software version 1.4.3 (Medical Research Council Biostatistics Unit, United Kingdom) was used [14].

Bayesian latent class analysis

The principles of applying latent class analysis in evaluating the performance of diagnostic tests without a gold standard were firstly introduced by Hui-Walter [15] and have been developed to use with the Bayesian framework in other studies [16–18]. In the past decade, Bayesian latent class analysis (BLCA), used to assess the performance of diagnostic tests in the fields of infectious disease and cancer, has become increasingly popular. The rationale to use this method has been well-explained in studies by Limmathurotsakul *et al.* [19] and Dendukuri [17].

With this statistical method, disease prevalence can be determined without consideration of any diagnostic test as the reference standard. BLCA has been successfully used to investigate the performance of diagnostic tests for patients with typhoid fever [20] and childhood tuberculosis [21]. In 2013, a web-based application for Bayesian using R and WinBUGS programs was developed to help researchers deal with

complex data sets [18]. In 2015, the Standards for Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models (STARD-BLCM) checklist has been introduced for checking the quality of diagnostic test studies with BLCM.

Briefly, instead of considering a perfect standard test truly existed, we hypothesized that all the tests were imperfect, then we assumed that our observed data, which were the combination of test results, arose from a mixture of true TBM (+) and TBM (-) groups. This assumption helped us to calculate the sensitivity and specificity of every single test.

Model definition

In BLCA, there are 3 components: the prior distribution, observer data, and posterior distribution. While the prior distribution is often based on the experience of experts or previous knowledge, the data set we analyzed is combinations of dichotomous test results. With *k* diagnostic tests, the number of combinations is 2^{*k*}. The combinations follow a multinomial distribution, with the likelihood being the multinomial probability of each observed combination [16]. From prior distributions and observed data, the estimation for distributions of the prevalence of a disease, specificity, and sensitivity of each test can be calculated from the simulating process using the Markov chain Monte Carlo (MCMC) method.

The method of constructing prior distributions using experts' knowledge was mentioned in a previous publication [22]. In this study, we used the informative prior distribution only for the specificity of the culture test, due to its specificity being widely accepted as perfect [8,21], and we assumed that the specificity of the culture test was near perfect (99%). For other parameters, non-informative prior distributions were used. Moreover, we also analyzed with prior distributions from 2 experts at Vietnam National Lung Hospital who had more than 20 years of experience working in TB fields, including TBM. The experts' panel included the head of the General Medical Department and the head of the National Tuberculosis Reference Laboratory. We used their personal beliefs about the prevalence of disease, the sensitivity, and specificity of MGIT culture test, Xpert test, protein levels, leukocyte count at the cut-off point 1g/L for protein level and 500 cells/mm³ for leukocyte count based on the well-known diagnostic criteria, the Lancet scoring system [5]. They were asked to elicit the most probable value or "best guess" (*mode*) and their certainty about the maximum values (95th percentile). The expert elicitations then were transformed into beta

distribution, Beta (a, b). For cut-off values of protein level and leukocyte counts, which are not mentioned in the Lancet scoring system, un-informative prior distribution was used due to the lack of previous information.

Since all 4 diagnostic tests were based on different principles, we assumed that all the diagnostic tests were independent (each test result was not influenced by other tests). The models were built with method described in previous studies [17,18].

Model fitting

Data were fitted based on Gibbs sampling by MCMC methods. For the model fitting, each model had passed a fitting of 3 simultaneous chains. A burn-in process of 2000 iterations was used for every chain. The simulation process was run with 20,000 iterations. The appropriate convergence was examined using a Gellman-Rubin-Brooks plot. Model fitting was measured using the deviation information criteria. The procedures were repeated for the various thresholds of protein concentration and leukocyte count. We calculated the area under the receiver operating characteristic curve (AUC) for each threshold of protein level based on the leukocyte count and repeated the progress for the leukocyte counts.

Identifiability and sensitivity analysis

A lack of identifiability happens when the number of degrees of freedom is less than the number of estimated parameters from the model. In this study, based on previous findings [23], at least 3 diagnostic tests are needed to ensure the model is identifiable. In these studies, we used the models for 4 diagnostic tests in a single population to avoid the non-identifiable problem.

The sensitivity analysis was performed to ensure the robustness of BLCA by running parallel models with different prior distributions from our chosen prior distribution and 2 expert elicitations distribution.

To enhance the quality of the study in the Bayesian approach method, STARD-BLCM was used [24].

Missing data

We used a multiple imputation method to impute missing data based on predictive mean matching [25]. Separate datasets were run parallel and compared together to check for the reliability of the imputed data sets.

Ethical approval

Ethical approval was obtained from the Human Research Ethics Committee, Faculty of Medicine, Prince of Songkla University, Thailand (project number 61236181). The study was performed based on the principles stated in the Declaration of Helsinki.

Results

A total of 1,213 suspected TBM patients were included, 140 patients from the prospective part, and 993 patients from the retrospective part. There were 785 males (64.7%) and 428 females (35.3%). Table 1 shows a summary of the study patients and their laboratory parameters. The median age was 51.3 (Interquartile Range (IQR): 34.2 – 64.3) years, and the median length of hospital stay was 13.7 (IQR: 7.5 – 21.1) days. The white blood cell count was within normal limits. However, the leukocyte-count and protein and glucose levels obtained from the cerebrospinal fluid were all high.

A total of 269 (22.2%) patients had a positive MGIT culture test, while from the Gene Xpert assay, 275 (22.7%) patients tested positive, of which 113 had a positive MGIT culture. 156 patients (12.9%) tested positive for both MGIT culture and Gene Xpert assay.

Bayesian latent class analysis

The complete set of prior distributions based on two-expert elicitations are shown in Table 2. Since the experts worked in different fields, their beliefs were inconsistent in several parameters. However, both experts believed that the culture test had a specificity of at least 99%.

Table 1. Clinical and laboratory characteristics of the patients.

Parameter	Median (IQR)	Geometric mean (SD)	Normal range
Age (years)	51.3 (34.2 - 64.3)	40.7 (2.66)	
Length of stay (days)	13.7 (7.5 – 21.1)	11.3 (2.52)	
White blood cells (10 ⁹ /L)	9.0 (6.7 – 12.1)	9.0 (1.55)	4.00 – 11.00
CSF leukocyte counts (cell/mm ³)	70.0 (20.0 – 317.5)	84.7 (5.5)	< 10
CSF protein level (g/L)	0.72 (0.33 – 1.57)	0.69 (3.13)	< 0.45
CSF glucose level (mmol/L)	3.28 (2.42 – 4.07)	3.01 (1.85)	> 2.2

CSF: Cerebrospinal fluid; SD: standard deviation; IQR: interquartile range.

Table 2. Prior distributions used for estimation based on expert opinions.

Parameter	Prior distribution	
	Expert 1	Expert 2
Culture test		
Sensitivity	Beta (5.9, 1.0)	Beta (5.0, 7.0)
Specificity	Beta(88.0, 1.0)	Beta (88.0,1.0)
Xpert MTB/RIF test		
Sensitivity	Beta (4.6, 2.0)	Beta (5.0, 7.0)
Specificity	Beta (18.5, 1.0)	Beta (48.5, 17.0)
Cerebrospinal fluid protein level		
Sensitivity	Beta (4.6, 2.0)	Beta (47.5, 21.0)
Specificity	Beta (4.7, 2.5)	Beta (23.0, 35.0)
Cerebrospinal fluid cell count		
Sensitivity	Beta (3.5, 1.5)	Beta (23.2, 34.5)
Specificity	Beta (3.0, 1.5)	Beta (48, 13.0)
Prevalence	Beta (2.5, 1.0)	Beta (48.2, 13.0)

Table 3 summarizes the posterior interferences of test performances using the different prior distributions shown in Table 2. Although we ran the models with multiple prior distributions in parallel, the estimation of the posterior distributions differed slightly. The TBM prevalence was estimated at 34.8% (95% CI: 28.8 – 41.3). The sensitivity, specificity, negative predictive value, and positive predictive value of the four diagnostic tests are presented in Table 3.

The Latent class analysis (LCA)-derived sensitivity and specificity of protein concentration level and CSF leukocyte count through models with different

thresholds are shown in Table 4. Briefly, the AUC for the leukocyte count was higher compared with the protein level (76% versus 73%), which indicated that the leukocyte count has better diagnostic precision than protein concentration. The receiver operating characteristic curves for both diagnostic parameters are illustrated in Figure 1.

Discussion

In this study, we evaluated the sensitivity and specificity of the MGIT culture test, Gene Xpert Mycobacterium tuberculosis (MTB)/resistance to

Table 3. Posterior means of sensitivity, specificity, PPV, and NPV of Gene Xpert assay and MGIT culture test, and protein levels and leukocyte count obtained from cerebrospinal fluid.

Parameter	Model*		Expert 1		Expert 2	
	Mean	95% CI	Mean	95% CI	Mean	95% CI
Prevalence	34.8	28.8 – 41.3	32.2	26.8 – 38.2	36.8	32.2 – 41.7
MGIT culture test						
Sensitivity	62.69	52.53 – 74.00	65.50	55.66 – 75.98	62.42	54.30 – 70.89
Specificity	99.42	99.01 - 100	98.49	98.01 – 99.65	99.47	99.02 – 100.0
PPV	98.26	96.84 - 100	95.32	93.24 - 98.98	98.53	97.21 – 99.91
NPV	83.15	75.37 – 90.16	85.59	78.76 – 91.48	81.91	75.84 – 87.31
Gene Xpert MTB/RIF test						
Sensitivity	57.52	51.00 – 64.00	60.59	53.64 – 67.41	54.14	47.82 – 60.56
Specificity	95.89	92.03 – 99.81	95.11	91.74 – 98.62	92.62	90.05 – 95.00
PPV	88.00	75.72 – 99.50	85.34	74.26 – 96.12	80.95	73.78 – 87.40
NPV	80.82	75.37 – 90.16	83.49	78.10 – 87.94	77.60	72.27 - 82.27
Cerebrospinal fluid protein level (cut-off 1g/L)						
Sensitivity	60.09	54.50 – 65.73	63.15	57.02 – 69.28	62.97	57.95 – 67.95
Specificity	65.75	61.80 – 69.78	65.59	53.64 – 67.41	64.48	60.74 – 68.29
PPV	49.62	41.53 – 57.80	46.62	39.22 – 54.28	50.75	44.34 – 57.47
NPV	75.94	69.26 – 81.80	78.90	72.65 – 84.23	74.94	69.96 – 79.52
Cerebrospinal fluid leukocyte count (cut-off 500 cells/mm³)						
Sensitivity	31.06	26.00 – 36.46	32.69	27.39 – 38.39	32.67	27.91 – 37.57
Specificity	89.08	86.19 – 91.98	88.76	86.00 – 91.49	88.65	85.85 - 91.35
PPV	60.21	48.88 – 71.92	57.96	47.46 – 68.98	62.57	53.13 – 73.14
NPV	70.75	64.50 – 70.88	73.48	67.63 – 78.67	69.35	64.56 – 73.73

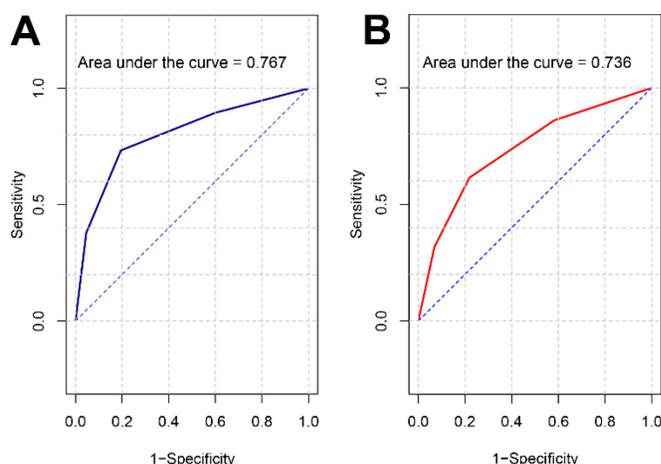
CI: Credible Interval; MGIT: Mycobacteria Growth Indicator Tube; NPV: Negative Predictive Value; PPV: Positive Predictive Value; * The model was based on a prior distribution of probability 0.99 (near perfect prediction) for MGIT culture test and uniform distribution for other tests.

rifampin (RIF) test, and the protein level and leukocyte count obtained from CSF in diagnosing TBM, and also the TBM prevalence. The study was performed using BLCA to avoid the assumption of any perfect references. This Bayesian approach was based on the application of latent class analysis for diagnostic tests introduced by previous studies [15,26]. This statistical method can assess the performance of a diagnostic test when a gold standard is unavailable or unfeasible [16,27]. By focusing on the combinations of test results rather than considering a single test as a reference, the Bayesian implementations help reduce biases in evaluating the performance of tests in the absence of a gold standard [10]. Moreover, the Bayesian framework can strengthen knowledge and information through the combination of prior knowledge of experts or from a systematic review [16,17]. Throughout the simulation process, the method can provide the most likely distribution for the performance of diagnostic tests and for disease prevalence, which is a crucial requirement for disease control and treatment.

The diagnostic process for TBM is still a big challenge since the disease is severe, with a high mortality rate. Most previous studies used the culture test or clinical examination scoring to evaluate performance of the diagnostic process. Nonetheless, both methods have limitations. The results of the culture test for Mycobacterium tuberculosis depends on various settings such as the use of antibiotics, bacterial loads, the quality of specimens, and technician experience level. On the other hand, the composite reference standards are often based on personal perspectives of local clinicians and can be biased because of non-specific symptoms [12].

Furthermore, some studies were designed with the confirmed TBM population or particular settings that could not afford complete interpretations for the general

Figure 1. Receiver operating characteristic curves of cerebrospinal fluid (CSF) leukocytes count and CSF protein concentration level for the tuberculous meningitis diagnosis. a. CSF leukocytes count, b CSF protein level.



clinical situations. Likewise, many studies lacked autopsy to confirm TBM. Hence, the performance of diagnostic tests and other clinical features was still not clearly understood [28].

While MGIT culture test was set as the gold standard in many previous studies, its accuracy is not 100%. Typically, when comparing with clinical diagnosis, the sensitivity of the test ranges from 31.8% to 48.8% [8,27,28]. Although our results confirm that the test is not 100% accurate, the estimates of the sensitivity of the test were higher than those reported in previous studies using clinical symptoms as a reference [8]. These differences can be explained by the inconsistency of examination results among clinicians based on the non-specific symptoms. Another advantage of the culture test is its ability to isolate Mycobacterium tuberculosis and distinguish non-tuberculous mycobacteria. In our study, we found that

Table 4. Sensitivity, specificity, and area under the receiver operating characteristic curve for cerebrospinal fluid leukocyte count and protein level for diagnosing tuberculous meningitis.

Protein threshold (g/L)		Leukocyte threshold (cells/mm ³)								AUC for LC at given protein threshold			
		10		20		500		1000					
		Protein	LC	Protein	LC	Protein	LC	Protein	LC				
0.4	Sensitivity	77.9	90.8	77.88	84.16	72.75	28.86	75.67	11.88	0.7673			
	Specificity	43.6	40.5	41.33	41.97	40.86	88.55	40.77	93.01				
1	Sensitivity	60.96	91.06	61.72	85.03	60.09	38.99	59.45	12.53		0.7634		
	Specificity	72.49	40.55	69.25	43.08	65.75	89.12	64.14	93.11				
2	Sensitivity	27.92	90.16	27.47	83.1	27.81	29.91	27.34	12.58			0.7645	
	Specificity	89.54	40.37	86.81	41.49	86.71	89.11	85.37	93.07				
3	Sensitivity	13.64	89.76	13.45	82.62	13.56	29.00	13.50	12.43				0.7654
	Specificity	92.50	40.32	91.69	41.35	91.65	88.58	91.04	92.85				
AUC for protein at given LC threshold		0.7306		0.7356		0.7320		0.7356					

AUC: Area under the curve; LC: leukocyte count.

1 case was caused by non-tuberculous mycobacteria. This patient was a 17-year-old male with no history of HIV and immunocompromized disease. First, his Xpert test for CSF was negative, and he was treated for TBM for 4 weeks based on his neurological symptoms before having culture test result positive with non-tuberculous mycobacteria. Although such infections are rare, several cases have been reported in the literature [29,30]. Further studies should be aware of these bacteria among the classical causal agent.

From a systematic review, the pooled sensitivity of Gene Xpert MTB/RIF assay when compared with culture test was 79.5% (95% CI: 62.0% - 90.2%), and the specificity ranged from 96% - 100% based on a 55% prevalence of clinical symptoms [31]. There have been several disagreements about the performance of the Gene Xpert assay when compared with the culture [32–34]. Interestingly, the estimate of the sensitivity of the Xpert test in our findings was higher than in earlier studies. Such disagreement could be explained by the difference in study settings, sample sizes, and statistical methods. For instance, when using the culture test as the gold standard, a positive result of the Xpert assay when the culture test was negative could be considered as a false positive [28]. The main advantages of the Xpert test are speed (results are available within 2 hours), and its ability to detect remaining DNA from unviable microbial cells. In this study, the specificity of Xpert test and other parameters were found to be lower than culture test. Nonetheless, such differences between these values could be explained by study design and statistical methodology. Bayesian statistics combine current state of knowledge with observed data to obtain posterior probabilities describing an event. The posterior inferences were based on observing data and prior knowledge. For the culture test, prior knowledge from the literature review and expert elicitation suggested a specificity around 99% to 100%. On the other hand, the prior distributions for the Xpert test were varied. However, after running the models with different prior values, this diagnostic test had nearly perfect specificity value (range from 92 – 95%), which indicated reliability of its use when diagnosing TBM compared with culture test. With its high specificity (nearly 100%) and sensitivity similar to the culture test, we recommend using the Xpert assay to aid the diagnostic process while waiting for the results of the culture test. However, a combination of the Xpert assay and clinical information with other CSF analysis parameters is preferred in place of using the test alone to rule out TBM due to the moderate negative predictive

ability [25]. Further studies with the new Xpert MTB/RIF Ultra assay should be considered [35].

We also estimated the performances of 2 standard CSF parameters, the protein concentration levels, and leukocyte count. Both had an excellent ability to distinguish TBM patients. Our findings were consistent with other previous studies in the moderate performance of these parameters [36,37]. As they are routine parameters for CSF analysis, leukocyte counts, and protein level are still useful indicators for TBM diagnosis. Similar to the previous study [4], we suggest using CSF cell counts as an indicator to diagnose TBM.

This is the first study to estimate the performance of the culture test, Xpert, CSF protein levels, and leukocyte counts in the diagnosis of the TBM without any reference standards. By using the statistical approach, we expected to broaden disease knowledge and information, which could be helpful information for clinicians dealing with TBM in high burden TB countries. The disease prevalence and the performance of routine diagnostic tests have essential implications for supporting clinician decisions and improving treatment outcomes. Through the assessment of chemicals and cytology parameters with different thresholds, our study provides a unique insight into how these parameters contribute to TBM diagnosis. Our study also aims to encourage the use of Bayesian methods in future research on TBM diseases. Other studies using this method in different populations and study settings would provide a better understanding of the disease worldwide. We hope our findings incorporated with those of other future studies on the use of Xpert-Ultra or other new diagnostic tests [35,38] will fulfill the diagnostic and treatment algorithm for TBM.

Our study had some limitations as it was conducted in a retrospective observational fashion. Due to study design, we had limited data for the CSF glucose levels and for the CSF/blood glucose ratio, which would be an important predictor for TBM diagnosis. In our study, the CSF glucose levels had a median of 3.28 (2.42 – 4.07) mmol/l or 59 (43.6 – 73.3) mg/dl, which was higher than in previous studies. However, problems about the quality of the retrospective data arose. As a national tertiary hospital, there were several cases referred from other secondary hospitals, where patients would have been treated with IV fluid or corticosteroids therapy, which could affect glucose levels. The use of CSF/Glucose ratio, instead of CSF glucose alone, should be more concise in diagnosing TBM. This problem should be addressed and considered with care in future studies.

Another limitation is the quality of the specimens. Although we followed careful guidelines and standard procedure for collecting and examining specimens, there were still limitations in quality control of the procedure and the follow-up process. Prospective studies with a plan to control the quality of specimen at the beginning and cohort studies that follow clinical course and patient outcome should be considered. A post-hoc analysis could have better validity and would enhance a more in-depth understanding of disease course and outcome. Nonetheless, this study still can provide useful information due to its large sample size and specific statistical modeling.

Future studies should be aware of several issues for the optimal use of BLCA. Firstly, it is crucial to have an appropriate prior hypothesis to increase the reliability of a Bayesian analysis. Better knowledge and experiences from experts can help improve the elicitation of prior distribution, which can help refine and solidify the final results. Although we used local experts' elicitation, the knowledge of experts can be affected by the national prevalence and epidemiology setting. A pooled experience from experts all over the world should be preferred. Secondly, in this current study, we assumed that all the diagnostic tests were independent due to their laboratory principles. In the future, other points of view, such as examining correlation assumptions and additional random effects, could be optimized to make a better fit model.

Conclusion

In summary, our findings indicated the Xpert test had sensitivity similar to the MGIT culture test and near-perfect specificity. Xpert test can be used as the primary test to combine with clinical information to achieve fast diagnosis and adequate treatment. Culture test can be used to confirm the result of Xpert test. The leukocyte count had higher AUC than that of the protein level. Among suspected patients, the prevalence of TBM patients was moderated.

Acknowledgements

This study was supported by the Scholarship Awards for Master and Ph.D. Studies Thailand's Education Hub for ASEAN Countries (TEH-AC) grant number: 059/2017. We would like to thank Professor Nguyen Van Hung, Head of Microbiology Department, Vietnam National Lung Hospital, for his valuable knowledge and advice to complete this study.

References

- World Health Organization (WHO) (2017) Global tuberculosis report 2017. WHO. Available: http://www.who.int/tb/publications/global_report/en/. Accessed 19 June 2018.
- Rock RB, Olin M, Baker CA, Molitor TW, Peterson PK (2008) Central nervous system tuberculosis: Pathogenesis and clinical aspects. *Clin Microbiol Rev* 21: 243–261.
- Heemskerck AD, Bang ND, Mai NTH, Chau TTH, Phu NH (2016) Intensified antituberculosis therapy in adults with tuberculous meningitis. *N Engl J Med* 374: 124–134.
- Thwaites G, Chau T, Stepniewska K, Phu N, Chuong L (2002) Diagnosis of adult tuberculous meningitis by use of clinical and laboratory features. *The Lancet* 360: 1287–1292.
- Marais S, Thwaites G, Schoeman JF, Török ME, Misra UK (2010) Tuberculous meningitis: a uniform case definition for use in clinical research. *Lancet Infect Dis* 10: 803–812.
- Saavedra JS, Urrego S, Toro ME, Uribe CS, García J (2016) Validation of Thwaites index for diagnosing tuberculous meningitis in a Colombian population. *J Neurol Sci* 370: 112–118.
- Vibha D, Bhatia R, Prasad K, Srivastava MVP, Tripathi M (2012) Validation of diagnostic algorithm to differentiate between tuberculous meningitis and acute bacterial meningitis. *Clin Neurol Neurosurg* 114: 639–644.
- Chaidir L, Annisa J, Dian S, Parwati I, Alisjahbana A (2018) Microbiological diagnosis of adult tuberculous meningitis in a ten-year cohort in Indonesia. *Diagn Microbiol Infect Dis* 91: 42–46.
- Nhu NTQ, Heemskerck D, Thu DDA, Chau TTH, Mai NTH, et al. (2014) Evaluation of GeneXpert MTB/RIF for diagnosis of tuberculous meningitis. *J Clin Microbiol* 52: 226–233.
- Limmathurotsakul D, Turner EL, Wuthiekanun V, Thaipadungpanit J, Suputtamongkol Y (2012) Fool's gold: Why imperfect reference tests are undermining the evaluation of novel diagnostics: A reevaluation of 5 diagnostic tests for leptospirosis. *Clin Infect Dis* 55: 322–331.
- Modi M, Garg A (2013) Tuberculosis of central nervous system. *J Postgrad Med Educ Res* 47: 202–213.
- Bahr NC, Marais S, Caws M, van Crevel R, Wilkinson RJ (2016) GeneXpert MTB/Rif to diagnose tuberculous meningitis: Perhaps the first Test but not the last. *Clin Infect Dis Off Publ Infect Dis Soc Am* 62: 1133–1135.
- Nagesh Babu G, Kumar A, Kalita J, Misra UK (2008) Proinflammatory cytokine levels in the serum and cerebrospinal fluid of tuberculous meningitis patients. *Neurosci Lett* 436: 48–51.
- Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat Comput* 10: 325–337.
- Hui SL, Walter SD (1980) Estimating the error rates of diagnostic tests. *Biometrics* 36: 167.
- Branscum AJ, Gardner IA, Johnson WO (2005) Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Prev Vet Med* 68: 145–163.
- Dendukuri N, Joseph L (2004) Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. *Biometrics* 57: 158–167. Doi: <https://doi.org/10.1111/j.0006-341X.2001.00158.x>
- Lim C, Wannapinij P, White L, Day NPJ, Cooper BS (2013) Using a web-based application to define the accuracy of diagnostic tests when the gold standard is imperfect. *PLoS ONE* 8. Available:

- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3827152/>. Accessed 20 December 2018.
19. Lim C, Paris DH, Blacksell SD, Laongnualpanich A, Kantipong P (2015) How to determine the accuracy of an alternative diagnostic test when it is actually better than the reference tests: A re-evaluation of diagnostic tests for scrub typhus using Bayesian LCMs. *PLoS ONE* 10. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4449177/>. Accessed 27 May 2018.
 20. Islam K, Sayeed Md, A, Hossen E, Khanam F, Charles RC (2016) Comparison of the performance of the TPTest, Tubex, Typhidot, and Widal immunodiagnostic assays and blood cultures in detecting patients with typhoid fever in Bangladesh, including using a Bayesian latent class modeling approach. *PLoS Negl Trop Dis* 10. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4825986/>. Accessed: 27 May 2018.
 21. Schumacher SG, van Smeden M, Dendukuri N, Joseph L, Nicol MP (2016) Diagnostic test accuracy in childhood pulmonary tuberculosis: A Bayesian latent class analysis. *Am J Epidemiol* 184: 690–700.
 22. Enøe C, Georgiadis MP, Johnson WO (2000) Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med* 45: 61–81.
 23. Toft N, Jørgensen E, Højsgaard S (2005) Diagnosing diagnostic tests: evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. *Prev Vet Med* 68: 19–33.
 24. Kostoulas P, Nielsen SS, Branscum AJ, Johnson WO, Dendukuri N (2017) STARD-BLCM: Standards for the Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models. *Prev Vet Med* 138: 37–47.
 25. Sriplung H, Yeesoonsang S, McNeil E, Bilheem S (2016) The use of a multiple imputation method to investigate the trends in Histologic types of lung cancer in Songkhla province, Thailand, 1989–2013. *BMC Cancer* 16: 389.
 26. Johnson WO (2001) Screening without a “Gold Standard”: The Hui-Walter paradigm revisited. *Am J Epidemiol* 153: 921–924.
 27. Johnson WO, Jones G, Gardner IA (2019) Gold standards are out and Bayes is in: Implementing the cure for imperfect reference tests in diagnostic accuracy studies. *Prev Vet Med* 167: 113–127.
 28. Thwaites GE, van Toorn R, Schoeman J (2013) Tuberculous meningitis: more questions, still too few answers. *Lancet Neurol* 12: 999–1010.
 29. Kwon LM, Kim ES, Lee K, Lee Y, Song JH (2018) Nontuberculous mycobacterial meningoencephalitis in a young healthy adult: A case report and literature review. *Radiol Infect Dis* 5: 85–90.
 30. Shi W, Jiao Y (2019) Nontuberculous Mycobacterium infection complicated with Haemophagocytic syndrome: a case report and literature review. *BMC Infect Dis* 19: 399.
 31. World Health Organization (WHO) (2013) Xpert MTB/RIF: WHO Policy update and Implementation manual. Available: https://www.who.int/tb/laboratory/xpert_launchupdate/en/. Accessed: 27 May 2018.
 32. Patel VB, Theron G, Lenders L, Matinyena B, Connolly C (2013) Diagnostic accuracy of quantitative PCR (Xpert MTB/RIF) for tuberculous meningitis in a high burden setting: A prospective study. *PLOS Med* 10: e1001536.
 33. Rufai SB, Singh A, Singh J, Kumar P, Sankar MM (2017) Diagnostic usefulness of Xpert MTB/RIF assay for detection of tuberculous meningitis using cerebrospinal fluid. *J Infect* 75: 125–131.
 34. Sharma K, Sharma M, Chaudhary L, Modi M, Goyal M (2018) Comparative evaluation of Xpert MTB/RIF assay with multiplex polymerase chain reaction for the diagnosis of tuberculous meningitis. *Tuberculosis* 113: 38–42.
 35. Bahr NC, Nuwagira E, Evans EE, Cresswell FV, Bystrom PV (2018) Diagnostic accuracy of Xpert MTB/RIF Ultra for tuberculous meningitis in HIV-infected adults: a prospective cohort study. *Lancet Infect Dis* 18: 68–75.
 36. Manning L, Laman M, Mare T, Hwaiwhanje I, Siba P (2014) Accuracy of cerebrospinal leucocyte count, protein and culture for the diagnosis of acute bacterial meningitis: a comparative study using Bayesian latent class analysis. *Trop Med Int Health* 19: 1520–1524.
 37. Solari L, Soto A, Agapito JC, Acurio V, Vargas D (2013) The validity of cerebrospinal fluid parameters for the diagnosis of tuberculous meningitis. *Int J Infect Dis* 17: e1111–e1115.
 38. Khonga M, Nicol MP (2018) Xpert MTB/RIF Ultra: a game changer for tuberculous meningitis? *Lancet Infect Dis* 18: 6–8.

Corresponding author

Associated Professor Hutcha Sriplung
Epidemiology Unit, Prince of Songkla University, 15,
Kanchanawit Rd., Hat Yai, Songkhla, 90110, Thailand
Tel: +66-74-451165
Fax: +66-74-429754
Email: hutchas@psu.ac.th

Conflict of interests: No conflict of interests is declared.