Am. J. Trop. Med. Hyg., 111(3), 2024, pp. 546–553
doi:10.4269/ajtmh.23-0789

Diagnostic Prediction Model for Tuberculous Meningitis: An Individual Participant Data Meta-Analysis

Anna M. Stadelman-Behar,[1]* Nicki Tiffin,[2,3] Jayne Ellis,[4] Fiona V. Creswell,[4,5] Kenneth Ssebambulidde,[6] Edwin Nuwagira,[7] Lauren Richards,[8] Vittoria Lutje,[9] Adriana Hristea,[10] Raluca Elena Jipa,[10] José E. Vidal,[11,12,13] Renata G. S. Azevedo,[14] Sérgio Monteiro de Almeida,[15] Gislene Botão Kussen,[15] Keite Nogueira,[15] Felipe Augusto Souza Gualberto,[16] Tatiana Metcalf,[17,18] Anna Dorothee Heemskerk,[19] Tarek Dendane,[20] Abidi Khalid,[20] Amine Ali Zeggwagh,[20] Kathleen Bateman,[21] Uwe Siebert,[22,23,24] Ursula Rochau,[24] Arjan van Laarhoven,[25] Reinout van Crevel,[25] Ahmad Rizal Ganiem,[26] Sofiati Dian,[26] Joseph Jarvis,[27,28] Joseph Donovan,[29,30] Thuong Nguyen Thuy Thuong,[29,30] Guy E. Thwaites,[29,30] Nathan C. Bahr,[31] David B. Meya,[6,32,33] David R. Boulware,[33] and Tom H. Boyles[28,34]

[1]*School of Public Health, University of Minnesota, Minneapolis, Minnesota;* [2]*South African National Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa;* [3]*Wellcome CIDRI–Africa, University of Cape Town, Cape Town, South Africa;* [4]*MRC/UVRI-LSHTM Uganda Research Unit, Entebbe, Uganda;* [5]*Global Health and Infection, Brighton and Sussex Medical School, East Sussex, United Kingdom;* [6]*Infectious Diseases Institute, Makerere University, Kampala, Uganda;* [7]*Department of Medicine, Mbarara University of Science and Technology, Mbarara, Uganda;* [8]*Division of Infectious Diseases, Department of Internal Medicine, Helen Joseph Hospital, University of Witwatersrand, Johannesburg, South Africa;* [9]*Cochrane Infectious Diseases Group, Liverpool, United Kingdom;* [10]*University of Medicine and Pharmacy Carol Davila, Bucharest, Romania;* [11]*Departmento de Neurologia, Instituto de Infectologia Emílio, São Paulo, Brazil;* [12]*Divisão de Clínica de Moléstias Infecciosas e Parasitárias, Hospital das Clınicas, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil;* [13]*Laboratório de Investigação Médica, Unidade 49, Hospital das Clínicas, Universidade de São Paulo, São Paulo, Brazil;* [14]*Departmento de Infectologia, Instituto de Infectologia Emílio, São Paulo, Brazil;* [15]*Hospital de Clinicas, Universidade Federal do Paraná, Curitiba, Brazil;* [16]*Center for Reference and Training in STD/AIDS CRT DST/AIDS, São Paulo, Brazil;* [17]*Department of Clinical Tropical Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand;* [18]*Northern Pacific Fogarty Global Health Fellowship Program, National Institutes of Health, University of Washington, Seattle, Washington;* [19]*Department of Medical Microbiology and Infection Prevention, Amsterdam University Medical Centre, Amsterdam, The Netherlands;* [20]*Medical Intensive Care Unit, Ibn Sina University Hospital, Mohammed V University, Rabat, Morocco;* [21]*Neurology Division, Department of Medicine Groote Schuur Hospital, University of Cape Town, Cape Town, South Africa;* [22]*Departments of Epidemiology and Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, Massachusetts;* [23]*Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts;* [24]*UMIT TIROL—University for Health Sciences and Technology, Hall I.T., Tirol, Austria;* [25]*Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, The Netherlands;* [26]*Department of Neurology Hasan Sadikin Hospital and TB/HIV Research Center Faculty of Medicine, Universitas Padjadjaran, Bandung, Indonesia;* [27]*Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana;* [28]*Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom;* [29]*Oxford University Clinical Research Unit, Centre for Tropical Medicine, Ho Chi Minh City, Vietnam;* [30]*Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom;* [31]*Division of Infectious Diseases, Department of Medicine, University of Kansas Medical Center, Kansas City, Kansas;* [32]*Department of Medicine, Faculty of Health Sciences, Makerere University, Kampala, Uganda;* [33]*Division of Infectious Diseases and International Medicine, Department of Medicine, University of Minnesota, Minneapolis, Minnesota;* [34]*Wits Reproductive Health and HIV Institute, University of the Witwatersrand, Johannesburg, Gauteng, South Africa*

Abstract. No accurate and rapid diagnostic test exists for tuberculous meningitis (TBM), leading to delayed diagnosis. We leveraged data from multiple studies to improve the predictive performance of diagnostic models across different populations, settings, and subgroups to develop a new predictive tool for TBM diagnosis. We conducted a systematic review to analyze eligible datasets with individual-level participant data (IPD). We imputed missing data and explored three approaches: stepwise logistic regression, classification and regression tree (CART), and random forest regression. We evaluated performance using calibration plots and C-statistics via internal–external cross-validation. We included 3,761 individual participants from 14 studies and nine countries. A total of 1,240 (33%) participants had "definite" (30%) or "probable" (3%) TBM by case definition. Important predictive variables included cerebrospinal fluid (CSF) glucose, blood glucose, CSF white cell count, CSF differential, cryptococcal antigen, HIV status, and fever presence. Internal validation showed that performance varied considerably between IPD datasets with C-statistic values between 0.60 and 0.89. In external validation, CART performed the worst (C = 0.82), and logistic regression and random forest had the same accuracy (C = 0.91). We developed a mobile app for TBM clinical prediction that accounted for heterogeneity and improved diagnostic performance (https://tbmcalc.github.io/tbmcalc). Further external validation is needed.

INTRODUCTION

The most lethal and disabling form of tuberculosis (TB) is tuberculous meningitis (TBM), of which an estimated 164,000 TBM cases occur annually.[1] Tuberculous meningitis diagnostics are inadequate due to combinations of poor accuracy, high cost, and lengthy turnaround times, leading to delayed diagnosis and poor outcomes.[2] Ziehl–Neelsen acid-fast bacilli (AFB) staining of CSF has low sensitivity, and mycobacterial culture is too slow to inform treatment decisions.[3] Although recently introduced nucleic acid amplification tests (NAATs), including the Xpert MTB/Rif Ultra assay (Cepheid, Sunnyvale, CA),[4,5] can speed up diagnosis, imperfect sensitivity means that negative results cannot fully exclude TBM.[6]

One approach to improving TBM diagnosis is combining all available information in a multivariable diagnostic prediction model. At least 10 prediction models have been developed that perform well in internal validation with diagnostic sensitivity ranging from 70% to 95% but poorly in external validation.[2] The primary reason for heterogeneous model performance across different settings and populations is case mix variation, which refers to the distribution of important predictor variables such as HIV-status and TB prevalence. Case

* Address correspondence to Anna M. Stadelman-Behar, 689 23rd Ave. SE, Minneapolis, MN 55455. E-mails: stad0110@umn.edu or anna.m.behar@gmail.com

mix variation across different populations can lead to differences in the prediction models performance, even when true predictor effects are consistent.[7,8] Prior TBM diagnostic prediction models were all developed from a single population and are typically based on comparisons to one other meningitis etiology, compromising external validity and clinical utility.

Individual participant data (IPD) from multiple studies provides an opportunity for external validation of any new model. Large datasets can examine heterogeneity and improve the predictive model performance across different populations, settings, and subgroups.[9–11] Participant-level data are preferred to meta-analysis of aggregated data because multiple individual-level factors can be jointly examined and interaction terms between variables can be considered.[12]

Therefore, we performed a systematic review and used IPD from multiple studies across multiple geographical locations to develop a composite TBM diagnostic prediction model. We used logistic regression as well as machine learning techniques, classification and regression tree (CART), and random forest models.

MATERIALS AND METHODS

Systematic literature search strategy.

We undertook systematic literature review per Preferred Reporting Items for Systematic Review and Meta-Analysis of IPD (PRISMA) guidelines, as per our published protocol.[13,14] We searched using MEDLINE and EMBASE to identify all studies reporting adult TBM diagnosis.[15] Controlled and natural language terms identified key search concepts such as: "tuberculosis," "meningitis," "diagnosis," "clinical feature," and "predictor." Supplemental Appendix A presents full search strategies, conducted on September 26, 2018.

Data acquisition and synthesis.

We requested anonymized IPD from corresponding authors of eligible studies. Specific demographic and clinical variables requested are listed in Supplemental Appendix B. We selected 13 target diagnostic predictors: symptom duration, cerebrospinal fluid (CSF) white cell count, CSF white cell differential, CSF glucose, CSF protein, CSF cryptococcal antigen (CrAg), blood glucose, blood white blood cell (WBC) count, HIV status, fever, TB incidence, age, and biological sex. We excluded subjects missing >50% of target predictors. We excluded datasets with clear pattern of missingness among target predictors based on diagnosis, age, sex, or some other participant characteristic. We analyzed datasets provided by the same research group as a single dataset.

Missing data.

Blood glucose was the variable with the most missingness in every dataset, and we performed single imputation of median glucose value in each dataset. For other missing target predictors, multiple imputation by chained equations was performed. Missing data within datasets were assumed missing at random. A total of 50 imputations were used per missing variable.

STATISTICAL ANALYSES

We defined "definite" TBM as any positive CSF test of MTB/RIF, Xpert MTB/RIF Ultra, other NAAT, culture, AFB, or "definite" TBM classification per the Uniform TBM case definition.[16] We defined "probable" TBM as having no alternate diagnoses and any computed tomography scan, magnetic resonance imaging scan, or X-ray suggestive of TBM or a "probable" TBM case classification per the uniform TBM case definition.[16] Definite or probable TBM were considered a TBM case.

We used three algorithm development strategies to predict TBM cases versus non-TBM cases. First, an IPD meta-analysis using a logistic regression model with an average intercept was fitted with the target predictors.[10] We used the backward stepwise method for predictor selection using a *P*-value threshold of 0.1. We also fitted a logistic regression model with stratified intercepts for each country. Next, we developed CART and random forest models with machine learning methods with the same target predictors as well as an indicator variable for country. We internally validated models using k-fold internal-external cross-validation.[10]

We externally validated models with data from a multisite, observational cohort based in Uganda that was not used to train the models. Multiple imputation for missing data was not performed, other than for blood glucose, so participants were included if they had complete data for the predictors included in the model.

We measured performance using the calibration ratio of predicted (expected) to observed outcomes, calibration plots (slope), and C-statistic.[10,17] We summarized overall prediction model internal validity by averaging the C-statistic values and amalgamating calibration across datasets. We summarized model external validity using Brier Score, a measure of probabilistic predictions accuracy where values close to 0 indicate perfect accuracy.[18] We also calculated diagnostic accuracy for all prediction probability thresholds. All analyses were conducted using R studio version 1.3.1093. Our findings are reported in accordance with the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis statement.[19]

The University of Minnesota and Makerere University's Institutional Review Boards approved this study.

RESULTS

Search results, studies, and participants included.

After deduplication, our searches yielded 2,179 reports that underwent title and abstract screening, and 121 full texts were reviewed (Figure 1). Thirty-four studies met our eligibility criteria, and we acquired IPD from 19 studies (18 datasets) totaling 6,147 individuals. Four datasets ($N = 796$ participants) were excluded because of either clear patterns of missingness for diagnosis or target predictors were missing in >50% of their data. An additional 82 participants <5 years old and 1,589 participants missing >50% of key predictors were excluded.

The final analysis dataset included 3,671 individuals from 15 studies (Table 1).[3,20–31] Most were cohort studies ($N = 9$) or cross-sectional studies ($N = 4$), and two were screening cohort data from randomized controlled trials. No study showed high risk of bias (Supplemental Appendix C).

In total, 1,148 (31%) participants met the case definition for definite TBM and 104 (3%) for probable TBM. Of non-TBM cases, 13% had cryptococcal, 6% bacterial, and 3% viral meningitis; the remainder had no confirmed diagnosis (Supplemental Appendix D). Participant demographics are

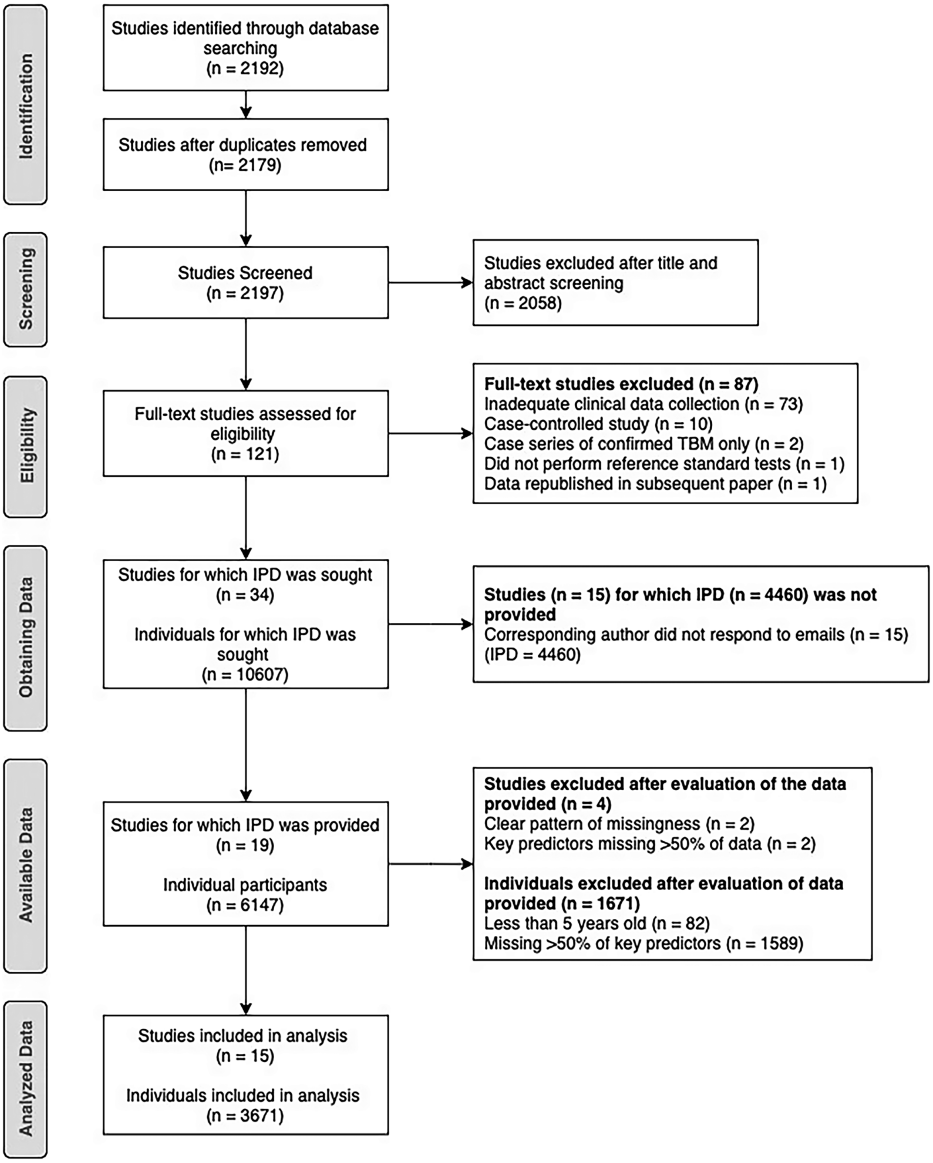

FIGURE 1. Preferred Reporting Items for Systematic Review and Meta-Analysis individual-level participant data (IPD) flow diagram of study selection process. TBM = tuberculous meningitis. This figure appears in color at www.ajtmh.org.

presented in Table 2. The final analysis dataset included 1,644 (45%) individuals who were HIV positive.

Multivariable prediction models.

The logistic regression model revealed CSF white cell count, CSF white cell differential (white cells below the detectable threshold, neutrophil predominance, or lymphocytic predominance), CSF glucose, blood glucose, CSF CrAg, and fever as significant predictors of TBM (Supplemental Appendix E). Symptom duration, blood WBC count, age, and biological sex were excluded due to complete missingness within some datasets, which could not be imputed. In a sensitivity analysis, we did not find these variables to be significantly predictive of TBM. Although not statistically significant, we retained HIV status in the model based on a 10% change in predictor values when excluded from the model as well as to account for case-mix variation. The logistic regression model with stratified intercepts for each country (with Brazil as the reference group due to largest sample size) is presented in Supplemental Appendix F. All the included target predictors were used in the development of both CART and random forest models. The resulting CART decision tree is shown in Supplemental Appendix G.

Internal–external cross-validation. Average C-statistic across datasets was 0.79 (95% CI: 0.75–0.83) for logistic regression, 0.76 (95% CI: 0.71–0.80) for CART, and 0.80 (95% CI: 0.76–0.84) for random forest (Supplemental Appendix H). The most heterogeneity in accuracy, indicated by the C-statistic, between datasets was observed with logistic regression (Supplemental Appendix I). Calibration ratio was most heterogenous in the CART model, and calibration slope was most varied in the random forest model, suggesting that the prediction diagnosis does not correspond with the observed diagnosis in both the CART and random forest models (Supplemental Appendix I). Visual inspection of the calibration plots for all three models indicates that CART is the most poorly calibrated model, with

TABLE 1
Characteristics of studies included in analysis dataset

Author/Owner	Year	Country	TB Burden	Study Design	N*	Age (range)	Men (%)	HIV (%)	Probable Cases (%)	Definite Cases (%)
Anselmo[20]	2017	Brazil	High	Cross-sectional	289	43 (6–84)	163 (56)	142 (49)	10 (3.5)	39 (13)
Gualberto[21]	2017	Brazil	High	Cohort	92	37 (8–64)	65 (71)	92 (100)	6 (6.5)	8 (8.7)
Azevedo[22]	2018	Brazil	High	Cohort	101	40 (17–73)	62 (61)	101 (100)	0 (0)	12 (12)
de Almeida[23]	2019	Brazil	High	Cross-sectional	321	40 (5–86)	188 (59)	177 (55)	13 (4.0)	13 (4.0)
Nhu[24]	2014	Vietnam	High	Cross-sectional	160	NA	NA	64 (40)	24 (15)	132 (83)
Heemskerk[3]	2018	Vietnam	High	Cohort	303	NA	NA	38 (13)	0 (0)	70 (23)
Donovan[25]	2020	Vietnam	High	Cohort	204	NA	NA	43 (21)	0 (0)	113 (55)
Jarvis[27]	2019	Botswana	Low	Cross-sectional	138	38 (5–90)	80 (58)	97 (70)	3 (2.2)	7 (5.1)
van Laarhoven and Dian[28]	2017	Indonesia	High	Cohort	761	30 (14–78)	460 (60)	146 (19)	0 (0)	339 (45)
Dendane[29]	2013	Morocco	Low	Cohort	414	32 (14–84)	221 (53)	1 (0.2)	0 (0)	246 (60)
Metcalf[30]	2018	Peru	Low	Cohort	37	40 (19–77)	27 (73)	23 (62)	11 (30)	8 (22)
Jipa[31]	2017	Romania	Low	Cohort	111	34 (18–75)	57 (51)	32 (29)	0 (0)	20 (18)
Bateman[26]	2012	South Africa	High	Cohort	93	32 (15–71)	43 (46)	49 (53)	9 (9.7)	30 (32)
Boulware[40]	2014	South Africa, Uganda	High, Low	RCT Screening	61	35 (19–75)	37 (61)	58 (95)	4 (7)	31 (51)
Rhein[41]	2019	Uganda	Low	RCT Screening	586	34 (14–75)	343 (59)	581 (99)	24 (4)	98 (17)
Total	–	–	–	–	3,671	35 (5–90)	1,746 (58)	1,644 (45)	104 (2.8)	1,148 (31)

NA = not available; RCT = randomized controlled trial; TB = tuberculosis.
* Sample size reflects the number that are included in this analysis and not the sample size from the article of origin. No study showed high risk of bias (Supplemental Appendix C).

logistic and random forest showing better (similar) calibration performance (Supplemental Appendix J). Sensitivity and specificity values at different prediction probability thresholds are displayed in Supplemental Appendices K–M. At the 0.1 prediction probability threshold for a positive test, sensitivity was higher in the CART (specificity = 0.55) and random forest (specificity = 0.32) models.

External validation.

A total of 404 participants were included in the external validation dataset (Supplemental Appendix N). Age ranged from 18 to 80 and differed with statistical significance different between TBM and non-TBM groups (Supplemental Appendix N). Most participants were HIV positive (N = 386, 96%). Thirty-two (8%) had definite TBM and 28 (7%) probable TBM (Supplemental Appendix O). The dominant meningitis etiology was cryptococcal meningitis (58%).

The mean C-statistic was the same for logistic regression and random forest (0.91) followed by CART (0.85) (Figure 2). The CART model had the calibration ratio closest to one, but the worst calibration slope value (–0.02) of the three models. Of the three models, calibration slope was closest to one for the random forest model (1.11) (Table 3). Visual inspection of the calibration plots for all three models shows that random forest is the best calibrated with the most bin midpoints falling along a 45-degree line (gray dotted line in Figure 3) of observed event percentage. The Brier scores were similar across the three models (Table 3).

At the predetermined prediction probability cutoff of 0.1, sensitivity and specificity were 0.77 and 0.89 for logistic regression, 0.70 and 0.81 for CART, and 0.87 and 0.73 for random forest (Supplemental Appendix P–R). The random forest model missed the fewest number of TBM cases with a false-negative rate (FNR) of 0.13 compared with the logistic (FNR = 0.23) and CART (FNR = 0.30) models. The logistic model had the highest proportion correctly classified with 87% of predictions correctly classifying individual participants as either TBM or non-TBM followed by CART, 79%, and random forest, 75%. Diagnostic performance of each model is summarized in Supplemental Appendices P–R.

An online/mobile application of our logistic regression model is available at https://tbmcalc.github.io/tbmcalc. Supplemental Appendix S displays logistic regression coefficients in

TABLE 2
Univariate analysis of clinical, hematological, and CSF data of individual participants with and without TBM

Variable	Non-TBM (N = 2,419)	TBM (N = 1,252)	P-Value
Age, years	35 (27–46)	32 (25–43)	<0.001
Men	1,229 (59)	517 (57)	0.260
Symptom Duration, days	7 (4–21)	11 (6–20)	<0.001
Fever	1,298 (54)	871 (70)	<0.001
HIV Positive	1,225 (51)	419 (34)	<0.001
Blood Glucose, mg/dL	103 (94–113)	104 (84–120)	0.477
CSF White Cell Count, cells/mm^3	8 (2.5–139)	140 (40–319)	<0.001
WBC Differential			
WBC <5 cells/mm^3	1,105 (46)	125 (10)	<0.001
Neutrophilic Dominance	373 (15)	332 (27)	
Lymphocytic Dominance	941 (39)	795 (64)	
CSF Protein, mg/dL	60 (31–134)	154 (86–267)	<0.001
CSF Glucose, mg/dL	53 (37–68)	23 (5.5–41)	<0.001
CSF CrAg Positive	455 (19)	13 (1.0%)	<0.001

CSF = cerebrospinal fluid; CrAg = CSF cryptococcal antigen; TBM = tuberculous meningitis; WBC = white blood cell.
Values are median (interquartile range) or N (%). P-value based on χ^2 or Kruskal–Wallis rank sum test.

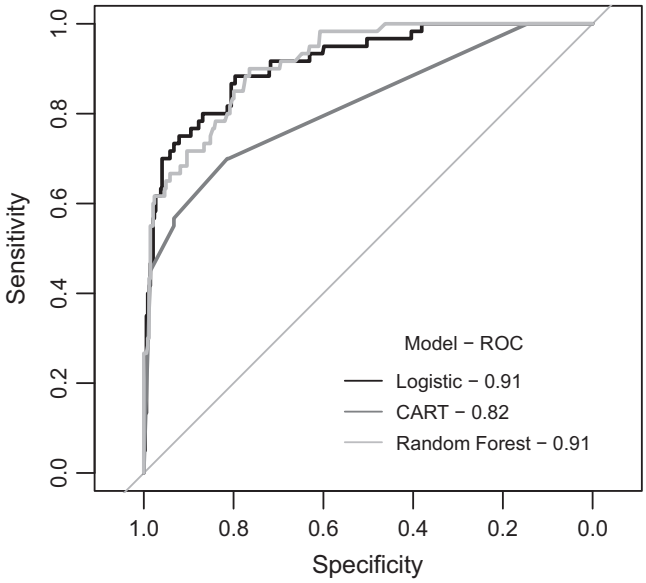

FIGURE 2. Receiver operating characteristic (ROC) curves for logistic regression, classification and regression tree (CART), and random forest models. Logistic = logistic regression. Numbers behind models indicate C-statistic (i.e., area under the receiver operator characteristic curve). This figure appears in color at www.ajtmh.org.

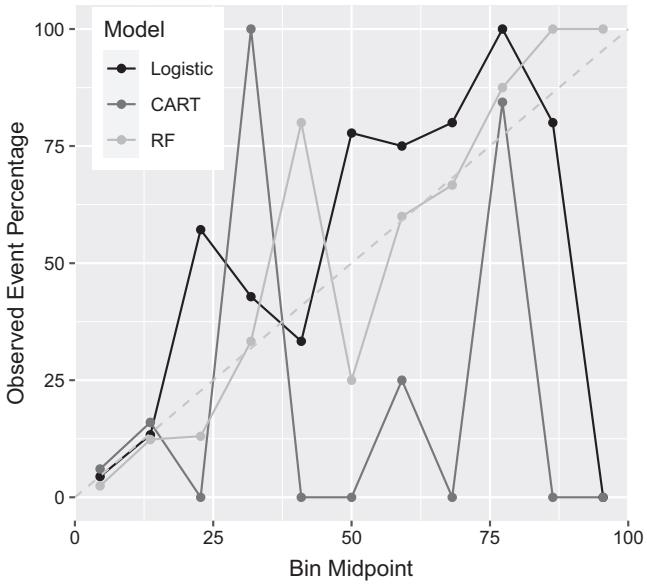

FIGURE 3. Calibration plot for Logistic, classification and regression tree analysis (CART), and random forest multivariate prediction models. Logistic = logistic regression; RF = random forest. This figure appears in color at www.ajtmh.org.

each IPD dataset, the final model, and external validation dataset.

DISCUSSION

This is the first study to develop broadly generalizable clinical multivariable prediction models for diagnosing TBM using a meta-analysis approach to synthesize data from diverse settings. The models were externally validated and showed excellent discrimination (C-statistic 0.82–0.91) and calibration (plots). Models derived using logistic regression showed similar performance to those using machine learning techniques (CART and random forest). Our models are primarily targeted to clinicians in low-resource settings with limited access to microbiologic testing and can be implemented using a smartphone application.

Previously published diagnostic models from single sites have reported higher C-statistic values than our study on internal validation but have largely shown disappointing results with external validation. One of the best known is Thwaites rule, with C-statistic of 0.99 for differentiating TBM from bacterial meningitis in Vietnam. External validation studies in Turkey,[32] India,[33] China,[34] and Colombia[35] showed good performance, but the model performed poorly in Malawi with high HIV prevalence.[36] Poor performance is largely attributed to the lack of representation of HIV-coinfected persons in the model's development dataset. HIV is a known contributor to case-mix variation thus influences a predictive model's generalizability.[2]

TABLE 3
Performance of logistic, CART, and random forest diagnostic algorithms

	Overall Calibration				
Model	Ratio (E/O)	Intercept	Slope	Brier Score	C-Statistic
Logistic Regression	0.57	26.8	0.49	0.07	0.91
CART	0.93	21.8	−0.02	0.09	0.82
Random Forest	0.70	−2.66	1.11	0.06	0.91

CART = classification and regression tree; C-statistic = area under the receiver operator characteristic curve; E/O = expected (predicted) case classification/observed case classification.

Our data clearly demonstrate that heterogeneity in clinical TBM case presentation affects performance of TBM diagnostic prediction models. Although we accounted for heterogeneity in every step of model development, internal validation revealed that performance of the three model types varied considerably in different populations and settings.

Overall, our findings indicate logistic regression performed better than machine learning approaches. This is consistent with a systematic review indicating, on average, no difference in the performance between logistic regression and machine learning approaches.[37]

In the context of suspected TBM, we suggest that the validated multivariable prediction models (based on readily attainable clinical data) should be used in conjunction with the experience of the treating clinician to guide immediate decisions about empiric TB treatment and the need for further or repeat testing. If the pretest probability is sufficiently low, it may be reasonable to exclude TBM without further testing.[38] If pretest probability is sufficiently high, empiric treatment of TBM may be initiated immediately. Using an overall TBM prevalence of 34%, our logistic regression model was able to decrease TBM probability to $<5\%$ in 694 (29%) participants without TBM and increase the probability to $>40\%$ in 969 (77%) participants with TBM by reclassification. Overall, our model can rapidly and accurately triage 45% of patients (1,663/3,671) to either no treatment of TBM (19%) and a search for alternative causes of meningitis or immediate treatment of TBM (26%) in the current study. The remaining 55% of patients would need NAAT diagnostics or findings from additional clinical investigation.

Our analysis demonstrates the significant contribution HIV infection makes to case-mix variation. The prevalence of HIV

has been identified as a significant contributor to heterogenous outcomes in TBM and modulates TBM pathogenesis.[39] All three models developed in this study performed most consistently in studies where all subjects were either all HIV positive[21,22,40,41] or all HIV negative.[29] Conversely, the models were most inconsistent in datasets with HIV prevalence ranging from 13% to 53%. Although HIV status was included in the prediction models as an independent TBM predictor, significant heterogeneity remained in performance in IPD datasets with a mix of individuals who were HIV positive and HIV negative.

A strength of this analysis is the large sample size taken from nine countries that improves generalizability in multiple settings. Our models showed good performance using laboratory and clinical evaluations that are readily available in resource-limited settings, where the burden of TBM is greatest. According to the WHO, blood glucose, HIV testing, and CrAg screenings are all considered essential diagnostics and are typically supplied in most hospitals and clinics.

Our study has several limitations. The lack of a perfect reference standard is common to all TBM diagnostic studies. We standardized our TBM case definitions across the studies including definite/probable TBM cases.[16] Our definitions are likely highly specific but may classify an unknown number of participants with TBM as non-TBM, biasing our findings toward the null. Cerebrospinal fluid volume and concentration techniques affect the sensitivity of reference standard tests.[42,43] Data on CSF volume was not available for our analysis, which may contribute to TBM cases being misclassified as non-TBM. Another limitation is missing predictor data across studies. Symptom duration, age, biological sex, and blood WBCs all predict TBM but could not be imputed due to patterns of missingness. The a priori prediction probability threshold of 0.1 is based on clinical expertise but needs further substantiating evidence. The final threshold(s) will depend on the benefit–harm balance of treating true-positive and false-negative patients, and the cost-effectiveness threshold ("willingness-to-pay") in the specific setting.[44] Finally, external validation in this analysis was limited to a population with a high prevalence of HIV and cryptococcal meningitis.

Further work should include externally validating the models using diverse cohorts with the possibility of updating them based on variables that have been significantly associated with TBM in previous studies. As previously proposed, harmonizing a minimum, essential dataset for TBM diagnostic studies would be helpful to coordinate for the future.[45] Finally, the current model should be tested in a clinical validation study coupled with further decision-analytic modeling to determine the impact on patient-relevant outcomes and cost.

Received November 10, 2023. Accepted for publication April 18, 2024.

Published online July 16, 2024.

Note: Supplemental materials appear at www.ajtmh.org.

Acknowledgments: We thank Drs. Kym Snell and Erika Helgeson for the more technical aspects of this work. We also thank authors who contributed data that was ultimately not eligible for inclusion in this work, including Drs. Ahmed Iqbal Bhigjee, Kameswar Prasad, Paul Pasco, and Omar Siddiqi. All their guidance was instrumental in completing this project.

Financial support: Fogarty International Center, NIH, USA (R01NS086312, D43TW009345). University of Minnesota Graduate School Doctoral Dissertation Fellowship. Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health (award number UL1-TR002494). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. N. Tiffin receives funding from Wellcome Trust CIDRI–Africa (203135/Z/16/Z), the NIH H3ABioNET Award (U24HG006941), and the UKRI/MRC (MC_PC_MR/T037733/1). F. V. Creswell is supported by a Wellcome PhD fellowship (210772/Z/18/Z).

Current contact information: Anna M. Stadelman-Behar, School of Public Health, University of Minnesota, Minneapolis, MN, E-mail: stad0110@umn.edu. Nicki Tiffin, South African National Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa, and Wellcome CIDRI–Africa, University of Cape Town, Cape Town, South Africa, E-mail: ntiffin@uwc.ac.za. Jayne Ellis MRC/UVRI-LSHTM Uganda Research Unit, Entebbe, Uganda, E-mail: j.ellis@doctors.org.uk. Fiona V. Creswell, MRC/UVRI-LSHTM Uganda Research Unit, Entebbe, Uganda, and Global Health and Infection, Brighton and Sussex Medical School, East Sussex, United Kingdom, E-mail: fiona.cresswell@lshtm.ac.uk. Kenneth Ssebambulidde, Infectious Diseases Institute, Makerere University, Kampala, Uganda, E-mail: kssebambulidde@gmail.com. Edwin Nuwagira, Department of Medicine, Mbarara University of Science and Technology, Mbarara, Uganda, E-mail: kanyama83@gmail.com. Lauren Richards, Division of Infectious Diseases, Department of Internal Medicine, Helen Joseph Hospital, University of Witwatersrand, Johannesburg, South Africa, E-mail: lolrichards@gmail.com. Vittoria Lutje, Cochrane Infectious Diseases Group, London, United Kingdom, E-mail: vittoria.lutje@lstmed.ac.uk. Adriana Hristea and Raluca Elena Jipa, University of Medicine and Pharmacy Carol Davila, Bucharest, Romania, E-mails: adriana_hristea@yahoo.com and ralucajipa@yahoo.com. José E. Vidal, Departmento de Neurologia, Instituto de Infectologia Emílio, São Paulo, Brazil, Divisão de Clínica de Moléstias Infecciosas e Parasitárias, Hospital das Clınicas, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, and Laboratório de Investigação Médica, Unidade 49, Hospital das Clínicas, Universidade de São Paulo, São Paulo, Brazil, E-mail: josevibe@gmail.com. Renata G. S. Azevedo, Departmento de Infectologia, Instituto de Infectologia Emílio, São Paulo, Brazil, E-mail: rguise@hotmail.com. Sérgio Monteiro de Almeida, Gislene Botão Kussen, and Keite Nogueira, Hospital de Clinicas, Universidade Federal do Paraná, Curitiba, Brazil, E-mails: sergio.ma@ufpr.br, gislene.kussen@ufpr.br, and keitenogueira@gmail.com. Felipe Augusto Souza Gualberto, Center for Reference and Training in STD/AIDS CRT DST/AIDS, São Paulo, Brazil, E-mail: felipegualberto@gmail.com. Tatiana Metcalf, Department of Clinical Tropical Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand, and Northern Pacific Fogarty Global Health Fellowship Program, National Institutes of Health, University of Washington, Seattle, Washington, E-mail: tatianametcalfg@gmail.com. Anna Dorothee Heemskerk, Department of Medical Microbiology and Infection Prevention, Amsterdam University Medical Centre, Amsterdam, The Netherlands, E-mail: a.heemskerk@vumc.nl. Tarek Dendane, Abidi Khalid, and Amine Ali Zeggwagh, Medical Intensive Care Unit, Ibn Sina University Hospital, Mohammed V University, Rabat, Morocco, E-mails: tdendane@hotmail.com, abidikhalid6@gmail.com, and aazeggwagh@gmail.com. Kathleen Bateman, Neurology Division, Department of Medicine Groote Schuur Hospital, University of Cape Town, Cape Town, South Africa, E-mail: kathleen.bateman@uct.ac.za. Uwe Siebert, Departments of Epidemiology and Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, MA, Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital Harvard Medical School, Boston, MA, and Department of Public Health, Health Services Research and Health Technology Assessment, UMIT—University for Health Sciences, Medical Informatics and Technology, Boston, MA, E-mail: uwe.siebert@umit.at. Ursula Rochau, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT—University for Health Sciences, Medical Informatics and Technology, Hall I.T., Tirol, Austria, E-mail: ursula.rochau@umit.at. Arjan van Laarhoven and Reinout van Crevel, Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, The Netherlands, E-mails: arjan.vanlaarhoven@radboudumc.nl and reinout.vancrevel@radboudumc.nl. Ahmad Rizal Ganiem and

Sofiati Dian, Department of Neurology Hasan Sadikin Hospital and TB/HIV Research Center Faculty of Medicine, Universitas Padjadjaran, Bandung, Indonesia, E-mails: rizalbdg@gmail.com and Sofiati.Dian@radboudumc.nl or sofiatidian@gmail.com. Joseph Jarvis, Botswana Harvard AIDS Institute Partnership, Gaborone, Botswana, and Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, E-mail: drjoejarvis@gmail.com. Joseph Donovan, Thuong Nguyen Thuy Thuong, and Guy E. Thwaites, Oxford University Clinical Research Unit, Centre for Tropical Medicine, Ho Chi Minh City, Vietnam, and Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom, E-mails: joseph.donovan@lshtm.ac.uk, thuongntt@oucru.org, and gthwaites@oucru.org. Nathan C. Bahr, Division of Infectious Diseases, Department of Medicine, University of Kansas Medical Center, Kansas City, KS, E-mail: nate.bahr@gmail.com. David B. Meya, Infectious Diseases Institute and Department of Medicine, Faculty of Health Sciences, Makerere University, Kampala, Uganda, and Division of Infectious Diseases and International Medicine, Department of Medicine, University of Minnesota, Kampala, Uganda, E-mail: david.meya@gmail.com. David R. Boulware, Division of Infectious Diseases and International Medicine, Department of Medicine, University of Minnesota, Minneapolis, MN, E-mail: boulw001@umn.edu. Tom H. Boyles, Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, and Wits Reproductive Health and HIV Institute, University of the Witwatersrand, Johannesburg, Johannesburg, South Africa, E-mail: drtomboyles@gmail.com.

REFERENCES

1. Dodd PJ, Osman M, Cresswell FV, Stadelman AM, Lan NH, Thuong NTT, Muzyamba M, Glaser L, Dlamini SS, Seddon JA, 2021. The global burden of tuberculous meningitis in adults: A modelling study. *PLOS Glob Public Health 1:* e0000069.
2. Wilkinson RJ, et al.; Tuberculous Meningitis International Research Consortium, 2017. Tuberculous meningitis. *Nat Rev Neurol 13:* 581–598.
3. Heemskerk AD, et al., 2018. Improving the microbiological diagnosis of tuberculous meningitis: A prospective, international, multicentre comparison of conventional and modified Ziehl–Neelsen stain, GeneXpert, and culture of cerebrospinal fluid. *J Infect 77:* 509–515.
4. Bahr NC, et al.; ASTRO-CM Trial Team, 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.
5. Cresswell FV, et al.; ASTRO-CM team, 2020. Xpert MTB/RIF Ultra for the diagnosis of HIV-associated tuberculous meningitis: A prospective validation study. *Lancet Infect Dis 20:* 308–317.
6. Boyles TH, Thwaites GE, 2015. Appropriate use of the Xpert (R) MTB/RIF assay in suspected tuberculous meningitis. *Int J Tuberc Lung Dis 19:* 276–277.
7. Riley RD, Ensor J, Snell KI, Debray TPA, Altman DG, Moons GMM, Colls GS, 2016. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: Opportunities and challenges. *BMJ 353:* i3140.
8. Debray TP, Damen JAA, Snell KI, Ensor J, Hooft L, Reitsma JB, Riley RD, Moons GMM, 2017. A guide to systematic review and meta-analysis of prediction model performance. *BMJ 356:* i6460.
9. Ahmed I, Debray TP, Moons KG, Riley RD, 2014. Developing and validating risk prediction models in an individual participant data meta-analysis. *BMC Med Res Methodol 14:* 3.
10. Debray TP, Moons KG, Ahmed I, Koffijberg H, Riley RD, 2013. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. *Stat Med 32:* 3158–3180.
11. Jolani S, Debray TP, Koffijberg H, van Buuren S, Moons KG, 2015. Imputation of systematically missing predictors in an individual participant data meta-analysis: A generalized approach using MICE. *Stat Med 34:* 1841–1863.
12. Riley RD, Lambert PC, Abo-Zaid G, 2010. Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ 340:* c221.
13. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF; PRISMA-IPD Development Group, 2015. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: The PRISMA-IPD statement. *JAMA 313:* 1657–1665.
14. Boyles T, Stadelman A, Ellis JP, Crewsswell FV, Lutje V, Wasserman S, Tiffin N, Wilkinson R, 2019. The diagnosis of tuberculous meningitis in adults and adolescents: Protocol for a systematic review and individual patient data meta-analysis to inform a multivariable prediction model. *Wellcome Open Res 4:* 19.
15. World Health Organization, 2019. *Global Tuberculosis Report 2019.* Geneva, Switzerland: WHO.
16. Marais S, Thwaites G, Schoeman JG, Török ME, Misra UK, Prasad K, Donald PR, Wilkinson RJ, Marais B, 2010. Tuberculous meningitis: A uniform case definition for use in clinical research. *Lancet Infect Dis 10:* 803–812.
17. Steyerberg EW, 2009. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.* New York, NY: Springer Science+Business Media.
18. Brier GW, 1950. Verification of forecasts expressed in terms of probability. *Mon Weather Rev 78:* 1–3.
19. Collins GS, Reitsma JB, Altman DG, Moons KG, 2015. Transparent reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). *Ann Intern Med 162:* 735–736.
20. Anselmo LMP, Feliciano C, Mauad F, Passeri do Nascimento M, Candido Pocente R, Silva JM, Bollela VR, 2017. A predictive score followed by nucleic acid amplification for adult tuberculous meningitis diagnosis in southern Brazil. *J Neurol Sci 379:* 253–258.
21. Gualberto FAS, Gonçalves MG, Fukasaw LO, Ramos Dos Santos AM, Sacchi CT, Harrison LH, Boulware DR, Vidal JE, 2017. Performance of nested RT-PCR on CSF for tuberculous meningitis diagnosis in HIV-infected patients. *Int J Tuberc Lung Dis 21:* 1139–1144.
22. Azevedo RG, Dinallo FS, de Laurentis LS, Boulware DR, Vidal JE, 2018. Xpert MTB/RIF((R)) assay for the diagnosis of HIV-related tuberculous meningitis in Sao Paulo, Brazil. *Int J Tuberc Lung Dis 22:* 706–707.
23. de Almeida SM, Borges CM, Santana LB, Golin G, Correa L, Kussen GB, Nogueira K, 2019. Validation of *Mycobacterium tuberculosis* real-time polymerase chain reaction for diagnosis of tuberculous meningitis using cerebrospinal fluid samples: A pilot study. *Clin Chem Lab Med 57:* 556–564.
24. Nhu NTQ, et al., 2014. Evaluation of GeneXpert MTB/RIF for diagnosis of tuberculous meningitis. *J Clin Microbiol 52:* 226–233.
25. Donovan J, et al., 2020. Xpert MTB/RIF Ultra versus Xpert MTB/RIF for the diagnosis of tuberculous meningitis: A prospective, randomised, diagnostic accuracy study. *Lancet Infect Dis 20:* 299–307.
26. Botha H, Ackerman C, Candy S, Carr JA, Griffith-Richards S, Bateman KJ, 2012. Reliability and diagnostic performance of CT imaging criteria in the diagnosis of tuberculous meningitis. *PLoS One 7:* e38982.
27. Mitchell HK, et al., 2019. Causes of pediatric meningitis in Botswana: Results from a 16-year national meningitis audit. *Pediatr Infect Dis J 38:* 906–911.
28. van Laarhoven A, et al., 2017. Clinical parameters, routine inflammatory markers, and LTA4H genotype as predictors of mortality among 608 patients with tuberculous meningitis in Indonesia. *J Infect Dis 215:* 1029–1039.
29. Dendane T, Madani N, Zekraoui A, Belayachi J, Abidi K, Zeggwagh AA, Abouqal R, 2013. A simple diagnostic aid for tuberculous meningitis in adults in Morocco by use of clinical and laboratory features. *Int J Infect Dis 17:* e461–e465.
30. Metcalf T, et al., 2018. Evaluation of the GeneXpert MTB/RIF in patients with presumptive tuberculous meningitis. *PLoS One 13:* e0198695.
31. Jipa R, Olaru ID, Manea E, Merisor S, Hristea A, 2017. Rapid clinical score for the diagnosis of tuberculous meningitis: A retrospective cohort study. *Ann Indian Acad Neurol 20:* 363–366.

32. Sunbul M, Atilla A, Esen S, Eroglu C, Leblebicioglu H, 2005. Thwaites' diagnostic scoring and the prediction of tuberculous meningitis. *Med Princ Pract 14:* 151–154.
33. Vibha D, Bhatia R, Prasad K, Srivastava MVP, Tripathi M, Kumar G, Singh MB, 2012. Validation of diagnostic algorithm to differentiate between tuberculous meningitis and acute bacterial meningitis. *Clin Neurol Neurosurg 114:* 639–644.
34. Zhang YL, Lin S, Shao LY, Zhang WH, Weng XH, 2014. Validation of Thwaites' diagnostic scoring system for the differential diagnosis of tuberculous meningitis and bacterial meningitis. *Jpn J Infect Dis 67:* 428–431.
35. Saavedra JS, et al., 2016. Validation of Thwaites Index for diagnosing tuberculous meningitis in a Colombian population. *J Neurol Sci 370:* 112–118.
36. Checkley AM, Njalale Y, Scarborough M, Zjilstra EE, 2008. Sensitivity and specificity of an index for the diagnosis of TB meningitis in patients in an urban teaching hospital in Malawi. *Trop Med Int Health 13:* 1042–1046.
37. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B, 2019. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol 110:* 12–22.
38. Pauker SG, Kassirer JP, 1980. The threshold approach to clinical decision making. *N Engl J Med 302:* 1109–1117.
39. Stadelman AM, et al., 2020. Treatment outcomes in adult tuberculous meningitis: A systematic review and meta-analysis. *Open Forum Infect Dis 7:* ofaa257.
40. Boulware DR, Meya DB, 2014. Antiretroviral therapy after cryptococcal meningitis. Letter. *N Engl J Med 371:* 1166–1167.
41. Rhein J, et al., 2019. Adjunctive sertraline for HIV-associated cryptococcal meningitis: A randomised, placebo-controlled, double-blind phase 3 trial. *Lancet Infect Dis 19:* 843–851.
42. Chaidir L, Annisa J, Dian S, Parwati I, Alisjahbana A, Purnama F, van der Zanden A, Ganiem AR, van Cevel R, 2018. Microbiological diagnosis of adult tuberculous meningitis in a ten-year cohort in Indonesia. *Diagn Microbiol Infect Dis 91:* 42–46.
43. Bahr NC, Tugume L, Rajasingham R, Kiggundu R, Williams DA, Morawski B, Alland D, Meya DB, Rhein J, Boulware DR, 2015. Improved diagnostic sensitivity for tuberculous meningitis with Xpert((R)) MTB/RIF of centrifuged CSF. *Int J Tuberc Lung Dis 19:* 1209–1215.
44. Siebert U, 2003. When should decision-analytic modeling be used in the economic evaluation of health care? *Eur J Health Econ 4:* 143–150.
45. Marais BJ, et al., 2017. Standardized methods for enhanced quality and comparability of tuberculous meningitis studies. *Clin Infect Dis 64:* 501–509.