Personalised risk-prediction tools for cryptococcal meningitis mortality to guide treatment stratification in sub-Saharan Africa: a prognostic modelling study based on pooled analysis of two randomised controlled trials

Thomas H A Samuels, Sile F Molloy*, David S Lawrence, Angela Loyse, Cecilia Kanyama, Robert S Heyderman, Wai Shing Lai, Sayoki Mfinanga, Sokoine Lesikari, Duncan Chanda, Charles Kouanfack, Elvis Temfack, Olivier Lortholary, Mina C Hosseinipour, Adrienne K Chan, David B Meya, David R Boulware, Henry C Mwandumba, Graeme Meintjes, Conrad Muzoora, Mosepele Mosepele, Chiratidzo E Ndhlovu, Nabila Youssouf, Thomas S Harrison†, Joseph N Jarvis†, Rishi K Gupta†*

Summary

Background Cryptococcal meningitis is a major driver of global HIV-related mortality, and validated approaches to stratify mortality risk could help to target effective treatment strategies. We aimed to develop and validate models to predict risk of all-cause mortality in people with HIV-associated cryptococcal meningitis in sub-Saharan African countries.

Methods For this prediction modelling study, we pooled individual-level data from the ACTA (ISRCTN45035509) and AMBITION-cm (ISRCTN72509687) randomised controlled trials. Data in ACTA were collected between Feb 12, 2013, and Jan 10, 2017, and data in AMBITION-cm were collected between Jan 31, 2018, and June 11, 2021. Adults aged 18 years or older with a first episode of HIV-associated cryptococcal meningitis were recruited to both trials. Exclusion criteria included pregnancy or lactation; receipt of high-dose anti-fungal treatment doses before screening; and contraindications to trial medication. Participants were recruited from nine hospitals across Cameroon, Malawi, Tanzania, and Zambia in ACTA and eight hospitals across Botswana, Malawi, South Africa, Uganda, and Zimbabwe in AMBITION-cm. We developed two primary multivariable logistic-regression models for the primary outcome of 2-week mortality: a basic model for use in a resource-limited setting that contained only candidate predictors that are routinely, programmatically obtained at hospital admission and a research model for which all predefined candidate predictors were considered for inclusion. We used internal–external cross-validation to evaluate model performance across countries within the development cohort (ie, data from all countries except Malawi participants in AMBITION-cm), before validation of discrimination, calibration, and net benefit in held-out data from Malawi.

Findings We included 674 eligible participants from ACTA and 814 from AMBITION-cm in the pooled analysis (total sample size 1488). 1263 participants were included in model development, with 225 from the Malawi site in AMBITION-cm held out for validation. 222 (17·6%) of 1263 participants in the development set and 21 (9·3%) of 225 participants in the validation set met the primary model outcome of 2-week mortality. We retained five predictors in the basic model and seven in the research model. Predictors in both models were Glasgow Coma Scale score, Eastern Cooperative Oncology Group performance status, haemoglobin, blood neutrophil count, and treatment. Additional predictors in the research model were cerebrospinal fluid opening pressure and log10 cerebrospinal fluid quantitative cryptococcal culture. Discrimination was relatively consistent between study sites for both models (pooled C statistic 0·75 [95% CI 0·68–0·82] for the basic model and 0·78 [0·75–0·82] for the research model), but calibration was more heterogeneous (pooled calibration slope 0·87 [95% CI 0·57 to 1·17] and 0·83 [0·69 to 0·97], pooled calibration in the large 0·00 [–0·54 to 0·55] and –0·02 [–0·46 to 0·42], for the basic and research models, respectively). In held-out validation, discrimination of both models was slightly higher than estimates from internal–external cross-validation (C statistic 0·78 [95% CI 0·70–0·87] in the basic model and 0·85 [0·79–0·92] in the research model). Calibration assessment suggested overestimation of risk, particularly in the high-risk range: calibration slope 1·04 (95% CI 0·54 to 1·55), calibration in the large –0·55 (–1·02 to –0·07). When comparing single, high-dose liposomal amphotericin B plus 14 days of flucytosine plus fluconazole with 1 week of amphotericin B plus flucytosine in AMBITION-cm, hazard ratios were 0·50 (95% CI 0·26–0·97) in the low-risk stratum and 0·96 (0·67–1·37) in the high-risk stratum for the basic model, and 0·61 (0·31–1·18) in the low-risk stratum and 1·03 (0·72–1·47) in the high-risk stratum for the research model.

Interpretation Both models accurately predicted 2-week mortality in people with HIV and have the potential to be incorporated into future treatment-stratification approaches in low-income and middle-income countries.

Funding National Institute for Health Research.

Lancet Glob Health 2025; 13: e920–30

*Joint first authors

†Joint senior authors

Division of Infection and Immunity (T H A Samuels MSc, Prof R S Heyderman FMedSci, Prof S Mfinanga PhD) **and UCL Respiratory Research Department, Division of Medicine** (R K Gupta PhD), **University College London, London, UK; Centre for Global Health, Institute for Infection and Immunity, City St George's University of London, London, UK** (S F Molloy PhD, A Loyse MD(Res), Prof T S Harrison MD); **Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK** (D S Lawrence PhD, N Youssouf PhD, Prof J N Jarvis PhD); **Botswana Harvard Health Partnership, Gaborone, Botswana** (D S Lawrence, Prof M Mosepele MD, N Youssouf, Prof J N Jarvis); **School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa** (D S Lawrence); **University of North Carolina Project–Malawi, Lilongwe, Malawi** (C Kanyama MD, Prof M C Hosseinipour MD); **Department of Medicine, Kamuzu Central Hospital, Lilongwe, Malawi** (C Kanyama, Prof M C Hosseinipour); **Malawi–Liverpool–Wellcome Trust Clinical Research Programme, Blantyre, Malawi** (Prof R S Heyderman, Prof H C Mwandumba PhD); **National Institute for Health and Care Research Biomedical Research Centre Clinical and**

Research Informatics Unit, University College London Hospitals NHS Foundation Trust, London, UK (W S Lai BSc); National Institute for Medical Research, Muhimbili Medical Research Centre, Dar es Salaam, Tanzania (Prof S Mfinanga, S Lesikari MD); Institute for Medical Research and Training, Lusaka, Zambia (D Chanda FRCP); Department of International Medicine, University Teaching Hospital, Lusaka, Zambia (D Chanda); Department of Public Health, Faculty of Medicine and Pharmaceutical Sciences, University of Dschang, Dschang, Cameroon (C Kouanfack PhD); Research Center on Emerging and Re-emerging diseases, Yaoundé, Cameroon (C Kouanfack); Site Agence Nationale de Recherche sur le Sida Cameroun, Hôpital Central Yaoundé, Yaoundé, Cameroon (C Kouanfack); Molecular Mycology Unit, Institut Pasteur, Paris, France (E Temfack MD); Department of Internal Medicine, Douala General Hospital, Douala, Cameroon (E Temfack); Department of Infectious Disease and Tropical Medicine, University of Paris, Paris, France (Prof O Lortholary PhD); Department of Mycology, National Reference Center for Invasive Mycoses and Antifungals, Institut Pasteur, Paris, France (Prof O Lortholary); Institut Imagine, Necker Enfant Malades University Hospital, Assistance Publique–Hôpitaux de Paris, Paris, France (Prof O Lortholary); Department of Medicine, University of North Carolina, Chapel Hill, NC, USA (Prof M C Hosseinipour); Division of Infectious Diseases, Department of Medicine, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada (A K Chan MD); Dignitas International, Zomba Central Hospital, Zomba, Malawi (A K Chan); Infectious Diseases Institute, College of Health Sciences (D B Meya PhD) and Department of Medicine, School of Medicine (D B Meya, Prof D R Boulware MD), Makerere University, Kampala, Uganda; Department of Medicine, University of Minnesota, Minneapolis, MN, USA (D B Meya, Prof D R Boulware); Department

Copyright © 2025 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Introduction

Cryptococcal meningitis is a major driver of global HIV-related mortality. There are approximately 112 000 deaths per year worldwide, with more than 75% occurring in Africa, accounting for 19% of all AIDS-related deaths.[1] Despite rapid roll-out of antiretroviral therapy, incidence of cryptococcal meningitis remains high and 10-week mortality ranges from 24% to more than 50%, depending on the setting and treatment used.[2–7] Stratifying people by disease severity at hospital admission could allow for targeted therapy; people who do not have severe disease could avoid extended hospitalisation and the most intensive and toxic treatments, whereas those who are severely unwell could be identified early for treatment escalation. Validated approaches to prognostication could be used in interventional trials to direct treatment stratification.

Previous studies have identified factors associated with increased risk of HIV-associated cryptococcal meningitis mortality, including increased age,[8–10] clinical measurements at presentation (eg, low bodyweight or BMI,[8,11,12] reduced conscious level,[8,9,12–15] or raised cerebrospinal fluid [CSF] opening pressure),[8,9] and laboratory measurements (eg, low CSF white cell count,[12–14,16] increased CSF fungal burden,[8,10,12,13,15] high peripheral white blood cell count,[8] low CD4 cell count,[9] low haemoglobin,[8] or high serum C-reactive protein).[17] However, the majority of studies published to date are small, with few conducted in Africa.[8,11] There are currently no practical tools combining these factors for clinical use to identify individuals at highest risk of death. Prognostic models have been developed for a range of acute infectious diseases, notably COVID-19,[18,19] for which extensive external validation has been done for a model to predict in-hospital mortality. For cryptococcal meningitis, existing prognostic models have been developed in studies from China via conventional statistical methods[9,20] and for people without HIV via machine learning approaches.[21] However, there are currently no validated models to guide clinical decision making for people living with advanced HIV in Africa, where the disease burden is highest.

We aimed to develop and validate models to predict risk of all-cause mortality in people with HIV-associated cryptococcal meningitis in African countries.

Methods

Study design and data sources

For this risk prediction modelling study, we pooled individual-level data from the two largest phase 3 randomised controlled trials to date in people with HIV-associated cryptococcal meningitis: ACTA (ISRCTN45035509) and AMBITION-cm (ISRCTN72509687).[4,5] Data in ACTA were collected between Feb 12, 2013 and Jan 10, 2017, and data in

Research in context

Evidence before this study

We searched PubMed for studies published between database inception and Jan 12, 2024, using the terms "cryptococcal meningitis", "HIV", "human immunodeficiency virus", "immunocompromised", "predict*", and "model*", with no language restrictions. Three previous studies, all conducted in China, developed prognostic models for cryptococcal meningitis mortality. Of these, two used statistical methods and the third used machine learning but only focused on people without HIV. We identified no studies conducted in Africa, specifically targeting people living with HIV, or using both statistical and machine learning approaches. Well developed and validated tools to predict risk of cryptococcal meningitis mortality and guide treatment stratification are thus lacking for resource-limited settings in Africa.

Added value of this study

To our knowledge, ours is the largest modelling study to date that includes development and validation of prediction models for HIV-associated cryptococcal meningitis mortality. We combined high-quality data from the two largest randomised controlled clinical trials conducted to date for cryptococcal meningitis treatment, with a total sample size of 1488 participants, of whom 236 (15·9%) met the 2-week mortality outcome and 469 (31·5%) met the 10-week mortality outcome. We developed two models, a basic model and a research model, to enable use in both resource-limited and research settings, where additional prognostic markers (eg, measurements of cerebrospinal fluid [CSF] opening pressure and CSF fungal burden) might also be available. Both models predicted risk of mortality with consistent discrimination and calibration across eight sub-Saharan African countries. Comparison between the logistic regression research model and a machine learning approach revealed no added value of the machine learning method. In exploratory analyses, treatment effects varied by predicted 2-week mortality risk, thus providing proof of concept for future treatment-stratification approaches.

Implications of all the available evidence

Our models accurately predicted risk of mortality among people with HIV-associated cryptococcal meningitis and showed consistent performance across two trials conducted in Africa. Predictions from the models could be used to direct treatment-stratification approaches in future clinical trials, with people who are at lowest predicted risk of mortality receiving therapy with reduced intensity and toxicity.

AMBITION-cm were collected between Jan 31, 2018 and June 11, 2021.

Briefly, adults aged 18 years or older with a first episode of HIV-associated cryptococcal meningitis were recruited to both trials. Exclusion criteria included pregnancy or lactation; receipt of high-dose anti-fungal treatment doses before screening; and contraindications to trial medication. Full details of the inclusion and exclusion criteria have been published previously.[4,5] In ACTA, participants were randomly assigned (1:1:1) to receive either fluconazole plus flucytosine for 14 days (ie, oral combination regimen), amphotericin B (amphotericin B refers to conventional, non-liposomal amphotericin B deoxycholate throughout, unless otherwise specified) plus either flucytosine or fluconazole for 7 days followed by 7 days of fluconazole, or amphotericin B plus either flucytosine or fluconazole for 14 days. Participants were recruited from nine hospitals across Cameroon, Malawi, Tanzania, and Zambia.[4] In AMBITION-cm, participants were randomly assigned (1:1) to receive either single, high-dose (10 mg/kg) liposomal amphotericin B plus 14 days of flucytosine plus fluconazole or amphotericin B (1 mg/kg per day) plus flucytosine for 7 days, followed by fluconazole for 7 days. Participants were recruited from eight hospitals across Botswana, Malawi, South Africa, Uganda, and Zimbabwe.[5] We chose a priori to include all data in model development apart from the Malawi site in AMBITION-cm, which was excluded from development and used for validation as Malawi was the only country where participants were recruited to both trials.

Data quality was assessed by visualising missingness and reproducing original trial report findings. Uncertainties were clarified with original trial investigators, where required. Sex was self-reported with the options of male or female.

Both studies were approved by the London School of Hygiene & Tropical Medicine Research Ethics Committee and local ethics and regulatory authorities in each country (appendix p 6). All participants provided written informed consent. If a participant had abnormal mental status, written informed consent was obtained from the next of kin; if a participant recovered the capacity to provide consent, written informed consent was obtained from that participant.

Model development and validation

We developed two primary multivariable logistic-regression models for the primary outcome of 2-week mortality, as we considered early mortality more likely to be directly associated with cryptococcal meningitis severity compared with later mortality. We developed a basic model for use in a resource-limited setting that contained only candidate predictors that are routinely, programmatically obtained at hospital admission. We developed a research model for which all predefined candidate predictors were considered for inclusion. We considered predefined candidate predictors for inclusion in both models on the basis of clinical knowledge, previous studies,[8–17] and availability of variables collected at baseline in both trials (appendix p 4). No causal framework was used.

We only considered variables if they were available from at least 60% of participants.[22] We chose variables for inclusion separately for the basic and research models using backward selection, based on the Akaike Information Criterion (AIC).[23] In this process, variables from the full model were iteratively removed and the AIC was computed, which includes a penalty for the number of predictor parameters. Predictors in the model with lowest AIC were retained. We modelled continuous variables using restricted cubic splines with 3 knots to assess non-linear associations. Predictors retained in more than 50% of multiply-imputed sets were retained in the final models.

We evaluated the models using internal–external cross-validation (appendix p 7), in which participants from one country were iteratively left out of the model-development dataset and used for validation.[22,24,25] This method evaluates the potential generalisability of a model between settings by examining between-setting heterogeneity in performance. To do this, the models were re-trained in the remaining countries in the development dataset and validated in the omitted country by quantifying discrimination and calibration. Discrimination assesses how well a model differentiates between individuals who do and do not meet the outcome, measured as the C statistic. Calibration evaluates how well predicted risk matches observed risk. We measured the calibration slope (slopes <1 suggest overfitting, whereas slopes >1 indicate predictions are too conservative), calibration in the large (<0 indicates predictions are too high overall, whereas >0 indicates systematic underprediction), and visualised calibration plots.[26] We then used a random-effects meta-analysis to calculate pooled measures of discrimination and calibration across countries in the development dataset.[24] We conducted re-calibration by country by re-estimating model intercepts. The final models were trained on the full development dataset before further validation in the held-out dataset (ie, the Malawi site in AMBITION-cm).

We also sought to test whether a statistical model using logistic regression could be improved upon via machine learning. We re-trained the research model with the same chosen variables using XGBoost[27] as a commonly applied and well performing machine learning approach for predicting a binary outcome (appendix p 3).[19]

Held-out validation of the statistical and machine learning models was done in the AMBITION-cm Malawi dataset by quantifying the C statistic, calibration slope, calibration in the large, and visualisation of calibration plots. We benchmarked performance to single univariable predictors and other HIV-associated cryptococcal meningitis prognostic models for which constituent variables were available in more than 60% of the full

of Medicine, Kamuzu University of Health Sciences, Blantyre, Malawi (Prof H C Mwandumba); **Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK** (Prof H C Mwandumba, Prof S Mfinanga); **Department of Medicine and Wellcome Centre for Infectious Diseases Research in Africa, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa** (Prof G Meintjes PhD); **Department of Internal Medicine, Faculty of Medicine, Mbarara University of Science and Technology, Mbarara, Uganda** (C Muzoora MD); **Department of Internal Medicine, University of Botswana, Gaborone, Botswana** (Prof M Mosepele); **Internal Medicine Unit, Faculty of Medicine and Health Sciences, University of Zimbabwe, Harare, Zimbabwe** (Prof C E Ndhlovu FRCP)

Correspondence to:
Dr Rishi K Gupta, UCL Respiratory Research Department, Division of Medicine, University College London, London WC1E 6JF, UK
r.gupta@ucl.ac.uk

See **Online** for appendix

	Overall (n=1488)	Alive (n=1245)	Died (n=243)
Study site			
Botswana	84 (5·6%)	72 (5·8%)	12 (4·9%)
Cameroon	107 (7·2%)	80 (6·4%)	27 (11·1%)
Malawi	682 (45·8%)	574 (46·1%)	108 (44·4%)
South Africa	107 (7·2%)	97 (7·8%)	10 (4·1%)
Tanzania	52 (3·5%)	38 (3·1%)	14 (5·8%)
Uganda	327 (22·0%)	274 (22·0%)	53 (21·8%)
Zambia	58 (3·9%)	47 (3·8%)	11 (4·5%)
Zimbabwe	71 (4·8%)	63 (5·1%)	8 (3·3%)
Trial			
ACTA	674 (45·3%)	535 (43·0%)	139 (57·2%)
AMBITION-cm	814 (54·7%)	710 (57·0%)	104 (42·8%)
Treatment			
Single, high-dose liposomal amphotericin B plus 14 days of flucytosine plus fluconazole	407 (27·4%)	354 (28·4%)	53 (21·8%)
1 week of amphotericin B plus flucytosine	518 (34·8%)	454 (36·5%)	64 (26·3%)
1 week of amphotericin B plus fluconazole	111 (7·5%)	75 (6·0%)	36 (14·8%)
2 weeks of amphotericin B plus flucytosine	115 (7·7%)	91 (7·3%)	24 (9·9%)
2 weeks of amphotericin B plus fluconazole	112 (7·5%)	87 (7·0%)	25 (10·3%)
Fluconazole plus flucytosine for 14 days	225 (15·1%)	184 (14·8%)	41 (16·9%)
Age, years	37 (32–43)	37 (32–43)	37 (32–44)
Sex			
Female	612 (41·1%)	505 (40·6%)	107 (44·0%)
Male	876 (58·9%)	740 (59·4%)	136 (56·0%)
Weight, kg	52 (47–60)	53 (47–60)	50 (45–60)
Missing	15 (1·0%)	10 (0·8%)	5 (2·1%)
Seizures	204 (13·7%)	146 (11·7%)	58 (23·9%)
Missing	4 (0·3%)	4 (0·3%)	0
Glasgow Coma Scale score			
15	1095 (73·6%)	969 (77·8%)	126 (51·9%)
11–14	327 (22·0%)	241 (19·4%)	86 (35·4%)
≤10	66 (4·4%)	35 (2·8%)	31 (12·8%)
Eastern Cooperative Oncology Group performance status			
Normal	64 (4·3%)	63 (5·1%)	1 (0·4%)
Restricted activity	256 (17·2%)	241 (19·4%)	15 (6·2%)
Ambulatory	339 (22·8%)	301 (24·2%)	38 (15·6%)
Limited self-care	512 (34·4%)	438 (35·2%)	74 (30·5%)
Bedbound	316 (21·2%)	201 (16·1%)	115 (47·3%)
Missing	1 (0·1%)	1 (0·1%)	0
White cell count, $\times10^9$/L	4·20 (3·10–5·60)	4·10 (3·00–5·50)	4·80 (3·58–6·60)
Missing	12 (0·8%)	9 (0·7%)	3 (1·2%)
Neutrophil count, $\times10^9$/L	2·50 (1·66–3·80)	2·40 (1·60–3·51)	3·40 (2·16–4·67)
Missing	31 (2·1%)	23 (1·8%)	8 (3·3%)
Haemoglobin, g/L	110 (96–126)	111 (97–126)	105 (90–123)
Missing	10 (0·7%)	8 (0·6%)	2 (0·8%)
CD4 count, $\times10^6$/L	27 (10–62)	27 (11–63)	21 (9–48)
Missing	88 (5·9%)	61 (4·9%)	27 (11·1%)

(Table 1 continues on next page)

dataset and model reconstruction was possible from reported manuscripts.[9,28] We also conducted decision-curve analysis[28] in the validation dataset to quantify overall net benefit of implementing the models to inform clinical decisions compared with a treat-all approach, a treat-none approach, the best performing univariable predictor, and other identified pre-existing models.[9]

We assessed the performance of the final logistic-regression models in predicting mortality at 10 weeks by evaluating discrimination for this outcome in the held-out validation dataset. Moreover, we developed separate models specifically to predict the 10-week outcome using the same methods as for the primary models, to assess whether these models further improved performance for predicting the 10-week outcome.

Statistical analysis

Based on our sample size, we estimated that 28 parameters could be considered for inclusion in the final model (appendix p 3). Missing data were handled via multiple imputation (appendix p 3). Uncertainty was quantified as 95% CIs and random effects meta-analyses were performed during internal–external cross-validation to account for between-setting heterogeneity.

In exploratory analyses, we assessed whether single, high-dose liposomal amphotericin B plus 14 days of flucytosine plus fluconazole (AMBITION-cm) or fluconazole plus flucytosine for 14 days (ACTA) were more effective than 1 week of amphotericin B plus flucytosine (as the control in AMBITION-cm and the best performing regimen in ACTA)[4,5] among participants at low risk of mortality, in line with the Predictive Approaches to Treatment effect Heterogeneity statement.[29] To do this, we generated predictions for 2-week mortality using the final basic and research models in the pooled dataset. Here, we assigned all participants to 1 week of amphotericin B plus flucytosine treatment to prevent treatment regimen from affecting generated predictions. We divided participants into low-risk, medium-risk, and high-risk strata on the basis of tertiles of risk predicted by the models. We then examined the treatment efficacy of these regimens compared with 1 week of amphotericin B plus flucytosine for 10 week mortality separately within each trial, with interaction terms by risk stratum. We quantified absolute risk differences between treatments by risk stratum using an identity link-function and quantified relative risk as hazard ratios using Cox regression.

We also evaluated whether treatment efficacy was modified by predicted risk as a continuous variable using Cox regression, with an interaction term between treatment and predicted 2-week mortality risk (including restricted cubic spline transformations to account for non-linear associations). We then visualised treatment effects as hazard ratios against predicted risk.

All analyses were conducted and reported in accordance with TRIPOD standards[7] and were done in R version 4.3.2.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

We included 674 eligible participants from ACTA (excluding four participants who were lost to follow-up) and 814 from AMBITION-cm in the pooled analysis (total sample size 1488). 1263 participants were included in model development, with 225 from the Malawi site in AMBITION-cm held out for validation (table 1; appendix p 10). 222 (17·6%) of 1263 participants in the development set and 21 (9·3%) of 225 participants in the validation set met the primary model outcome of 2-week mortality.

In variable selection, we retained five predictors in the basic model and seven in the research model. Predictors in both models were Glasgow Coma Scale (GCS) score, Eastern Cooperative Oncology Group (ECOG) performance status, haemoglobin, blood neutrophil count, and treatment. Additional predictors in the research model were CSF opening pressure and $\log_{10}$CSF quantitative cryptococcal culture. Predictor–outcome associations were similar in both models (figure 1; appendix pp 11–14).

Discrimination was relatively consistent between study sites for both models (pooled C statistic 0·75 [95% CI 0·68–0·82] for the basic model and 0·78 [0·75–0·82] for the research model; figure 2). Calibration was more heterogeneous, with calibration in the large varying by study site in both models, likely reflecting variation in baseline risk between sites (figure 2). Pooled calibration plots by study site showed systematic overestimation of risk in Malawi and underestimation in Uganda (appendix p 15). Recalibration to each study site by re-estimation of the model intercept led to improvement in model calibration (appendix p 15).

In held-out validation, discrimination of both models was slightly higher than estimates from internal–external cross-validation (C statistic 0·78 [95% CI 0·70–0·87] in the basic model and 0·85 [0·79–0·92] in the research model; table 2). Calibration assessment suggested overestimation of risk, particularly in the high-risk range, where data were sparse (figure 3A, B). Discrimination did not vary systematically by sex or age (appendix p 16). ECOG performance status was the strongest univariable predictor for 2-week mortality (0·78 [0·71–0·85]; appendix p 17) but had lower discrimination than the research model (C statistic difference 0·07 [95% CI 0·00–0·15]; p=0·048; appendix p 19). Discrimination of single predictors varied by study site (appendix p 18).

Among identified pre-existing models, only the model created by Zhao and colleagues[9] could be reconstructed from available data. Discrimination in the validation dataset, calculated with their reported point score, was worse than both primary models (C statistic 0·69

	Overall (n=1488)	Alive (n=1245)	Died (n=243)
(Continued from previous page)			
CSF opening pressure, cm H20	22 (13–33)	21 (13–32)	25 (16–40)
Missing	54 (3·6%)	50 (4·0%)	4 (1·6%)
CSF cell count, white blood cells per mm^3	4 (1–37)	5 (1–40)	4 (0–16)
Missing	59 (4·0%)	48 (3·9%)	11 (4·5%)
$\log_{10}$CSF quantitative culture, CFU/mL	4·79 (3·08–5·66)	4·63 (2·91–5·54)	5·48 (4·49–6·15)
Missing	40 (2·7%)	31 (2·5%)	9 (3·7%)

Data are n (%) or median (IQR). CSF=cerebrospinal fluid.

Table 1: **Baseline characteristics of participants in the ACTA[4] and AMBITION-cm[5] trials, stratified by 2-week mortality outcome**

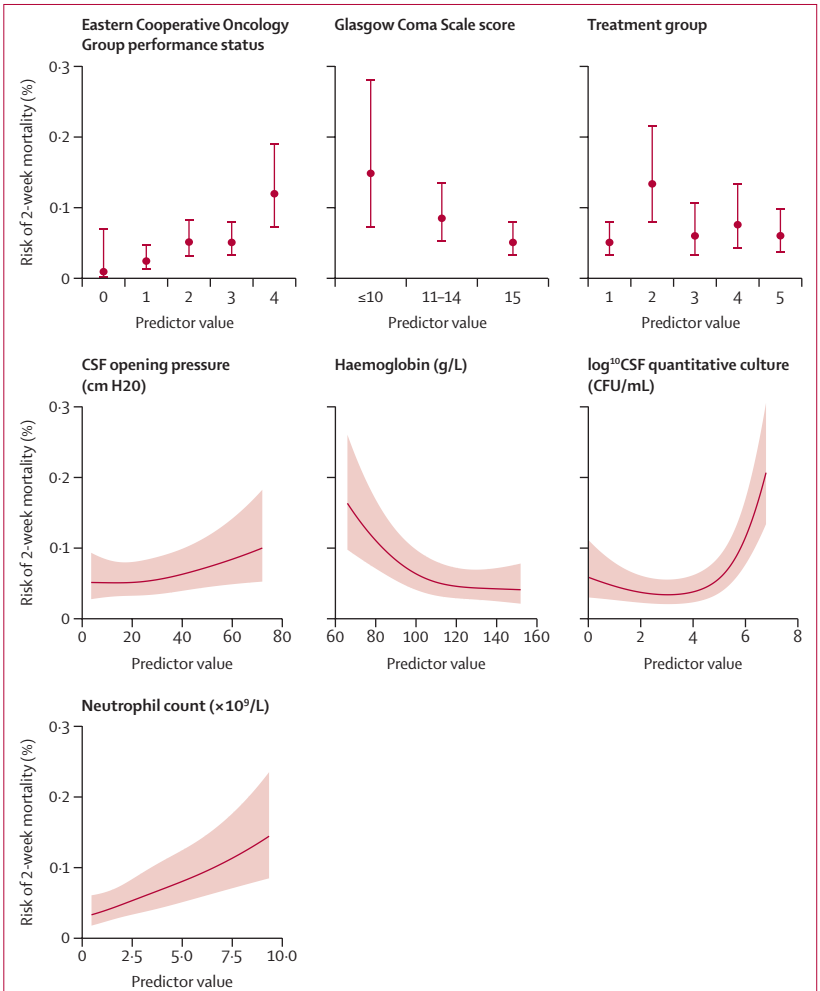

Figure 1: **Multivariable associations between specific predictors and outcomes in the primary research model**
We modelled continuous variables using restricted cubic splines. Final model parameters were pooled across multiple imputed datasets (total sample size 1263 participants; appendix p 12). The treatment groups are as follows: group 1 is single, high-dose liposomal-amphotericin-B regimen and the 1-week amphotericin B plus flucytosine groups from both ACTA and AMBITION-cm trials; group 2 is 1-week amphotericin B plus fluconazole; group 3 is 2 weeks amphotericin B plus flucytosine; group 4 is 2 weeks amphotericin B plus fluconazole; and group 5 is flucytosine plus fluconazole for 14 days (oral combination regimen). For categorical variables, solid dots show point estimates and lines show 95% CIs. For continuous variables, lines show point estimates and shaded areas show 95% CIs. CSF=cerebrospinal fluid.

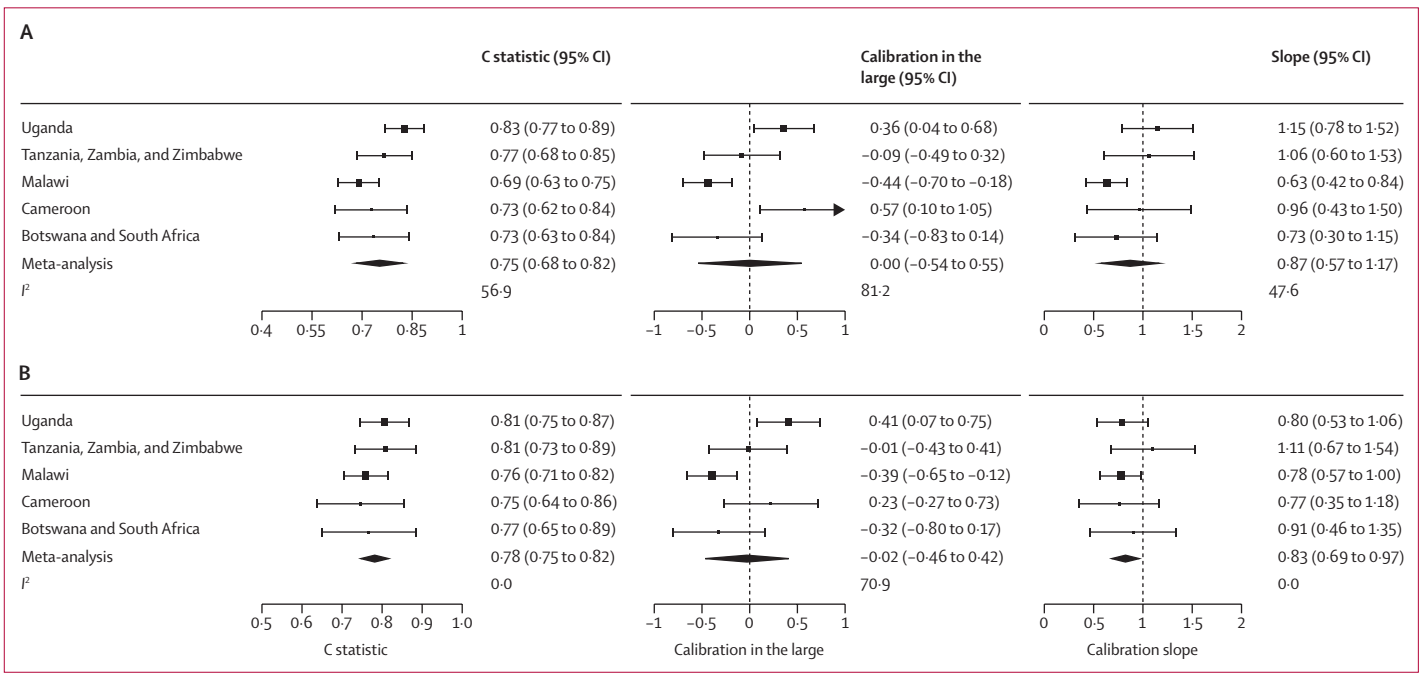

Figure 2: **Internal–external cross-validation of the basic model (A) and research model (B), by country**
We calculated pooled estimates through a random-effects meta-analysis (total sample size 1263 participants). Countries with <100 participants or <20 deaths were amalgamated and grouped by similarity of health-care environment. Dashed lines indicate lines of perfect calibration in the large and perfect slope. Black squares indicate point estimates and bars indicate 95% CIs. Diamonds indicate pooled random-effects meta-analysis estimates.

	Validation outcome	C statistic (95% CI)	Slope (95% CI)	Calibration in the large (95% CI)
Primary models trained for 2-week mortality				
Basic	2 weeks	0·78 (0·70 to 0·87)	1·04 (0·54 to 1·55)	−0·55 (−1·02 to −0·07)
Basic	10 weeks	0·74 (0·66 to 0·81)	..	..
Research	2 weeks	0·85 (0·79 to 0·92)	1·14 (0·69 to 1·60)	−0·57 (−1·06 to −0·07)
Research	10 weeks	0·77 (0·70 to 0·85)	..	..
Secondary models trained for 10-week mortality				
Basic	10 weeks	0·77 (0·71 to 0·84)	1·31 (0·85 to 1·77)	−0·16 (−0·48 to 0·16)
Research	10 weeks	0·78 (0·71 to 0·85)	1·10 (0·72 to 1·49)	−0·07 (−0·40 to 0·25)
Machine learning model trained for 2-week mortality				
XGBoost	2 weeks	0·83 (0·76 to 0·89)	0·87 (0·49 to 1·26)	−0·63 (−1·12 to −0·13)
Model by Zhao and colleagues[9] trained for 28-day mortality	28 days	0·69 (0·60 to 0·78)	0·35 (0·14 to 0·56)	..

Data for slope and calibration in the large are not reported for the 10-week validation outcome of the primary model as it was developed for the 2-week mortality outcome. Data for calibration in the large are not reported for the model by Zhao and colleagues as the model intercept was recalibrated to our data to reconstruct the model and assess calibration.

Table 2: **Model performance in held-out validation data**

[95% CI 0·60–0·78]). To report calibration, we recalibrated the model intercept to our validation dataset as no intercept was reported. Calibration assessment suggested that predictions were too extreme, with a slope of 0·35 (95% CI 0·14–0·56; table 2; appendix p 16).

Predictor–outcome associations with XGBoost were similar to the logistic-regression approach, with no evidence of two-way interactions (appendix p 21). In the validation dataset, XGBoost discrimination and calibration metrics were also similar to the logistic-regression model (table 2; figure 3). The XGBoost calibration plot showed miscalibration with predictions being too extreme, reflecting overfitting and leading to a slope of less than 1 (figure 3C). Overall, the machine learning model showed no improvement over the logistic-regression approach.

Decision-curve analysis in the validation dataset showed greater net benefit for our basic and research models than the model by Zhao and colleagues,[9] ECOG alone, the treat-all approach, and the treat-none approach across a range of threshold probabilities, for which the weighting of the false positives varied (appendix pp 19–20). The research model showed superior net benefit overall to guide management (appendix pp 19–20).

411 (32·5%) of 1263 participants met the 10-week mortality outcome in the development dataset and 58 (25·8%) of 225 participants met the 10-week mortality outcome in the validation dataset. We re-trained the logistic-regression models, including variable selection, to assess whether models trained for the 10-week outcome would improve prediction for the extended time period. Seven predictors were retained in the re-trained basic model and eight were retained in the research

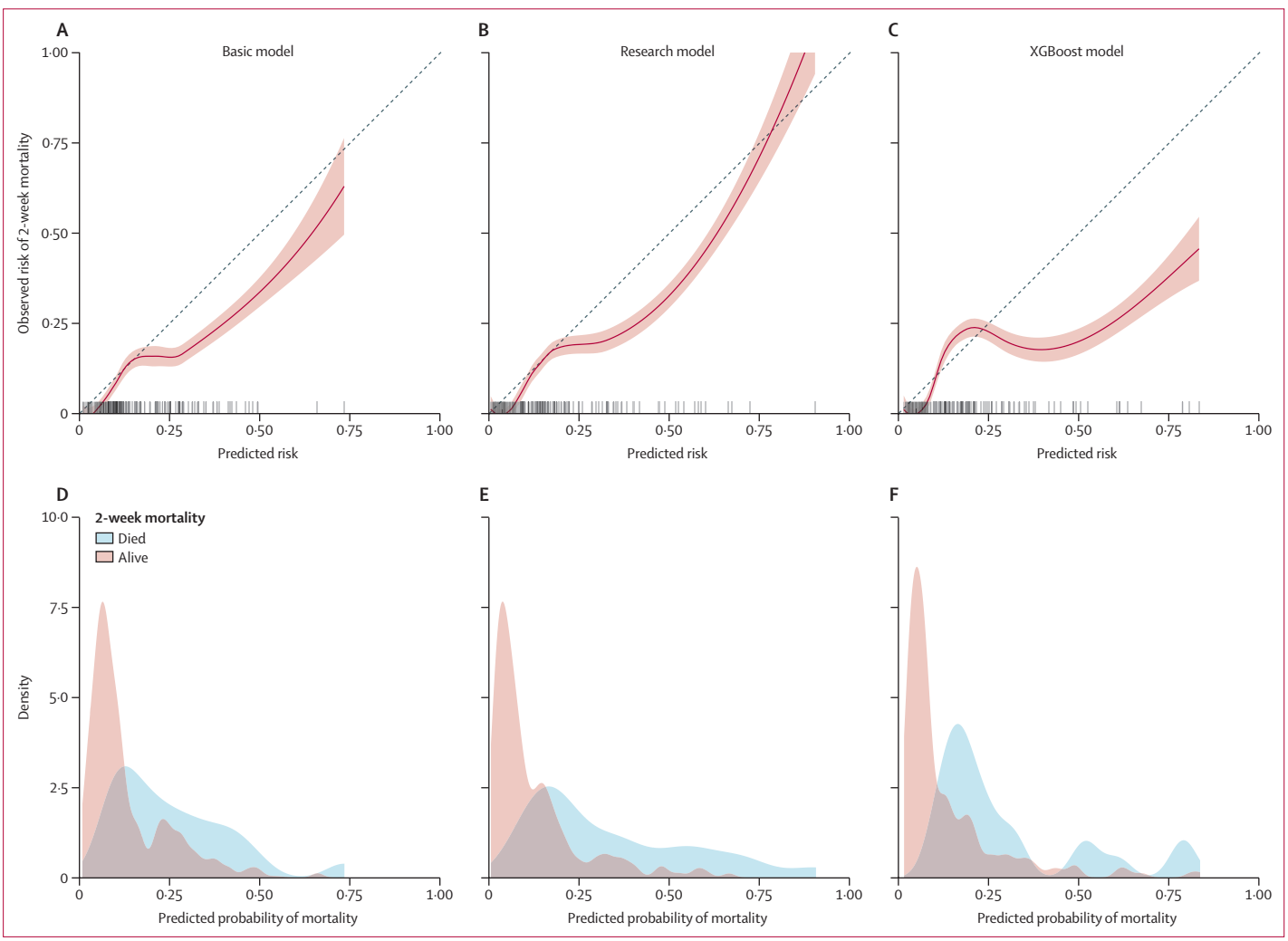

Figure 3: **Model calibration and prediction density in held-out validation data**
(A) Calibration of the basic 2-week mortality model. (B) Calibration of the research 2-week mortality model. (C) Calibration of the XGBoost machine learning 2-week mortality model. (D) Density plot for 2-week mortality predictions made by the basic model, stratified by 2-week mortality outcome. (E) Density plot for 2-week mortality predictions made by the research model, stratified by 2-week mortality outcome. (F) Density plot for 2-week mortality predictions made by the XGBoost machine learning model, stratified by 2-week mortality outcome. Calibration is shown via a loess smoother. Solid lines show point estimates and shaded areas show 95% CIs. Rug plots on x axes show the distribution of predicted risk of mortality. Dashed grey diagonal lines show perfect calibration.

model for the 10-week outcome. In the basic model, age, bodyweight, and presence of seizures at hospital admission were included and GCS was excluded. In the research model, age and CSF cell count were included and CSF opening pressure was excluded (appendix pp 22, 25–28). Overall, for both the basic and research models, discrimination for the 10-week outcome was similar for the re-trained 10-week models and the primary 2-week models (table 2). For both 10-week models, the calibration in the large was closer to 0 for the 10-week outcome compared with the primary 2-week models for the 2-week outcome (appendix pp 28–29), suggesting better calibration for the 10-week models.

We hypothesised that treatment effects within each trial might vary according to predicted 2-week mortality risk. Predicted risk for both models was positively skewed, with modal risks of 7·6% for the basic model and 4·2% for the research model, lower than the overall cohort mortality of 16·3% (appendix p 30). When comparing single, high-dose liposomal amphotericin B plus 14 days of flucytosine plus fluconazole with 1 week of amphotericin B plus flucytosine in AMBITION-cm, hazard ratios were 0·50 (95% CI 0·26–0·97) in the low-risk stratum and 0·96 (0·67–1·37) in the high-risk stratum for the basic model, and 0·61 (0·31–1·18) in the low-risk stratum and 1·03 (0·72–1·47) in the high-risk stratum for the research model (figure 4; appendix pp 31–32). When comparing fluconazole plus flucytosine for 14 days with 1 week of amphotericin B plus flucytosine in ACTA, hazard ratios were 1·12 (0·40–3·15) in the low-risk stratum and 1·44 (0·73–2·85) in the high-risk stratum for the basic model, and 1·65 (0·57–4·77) in the

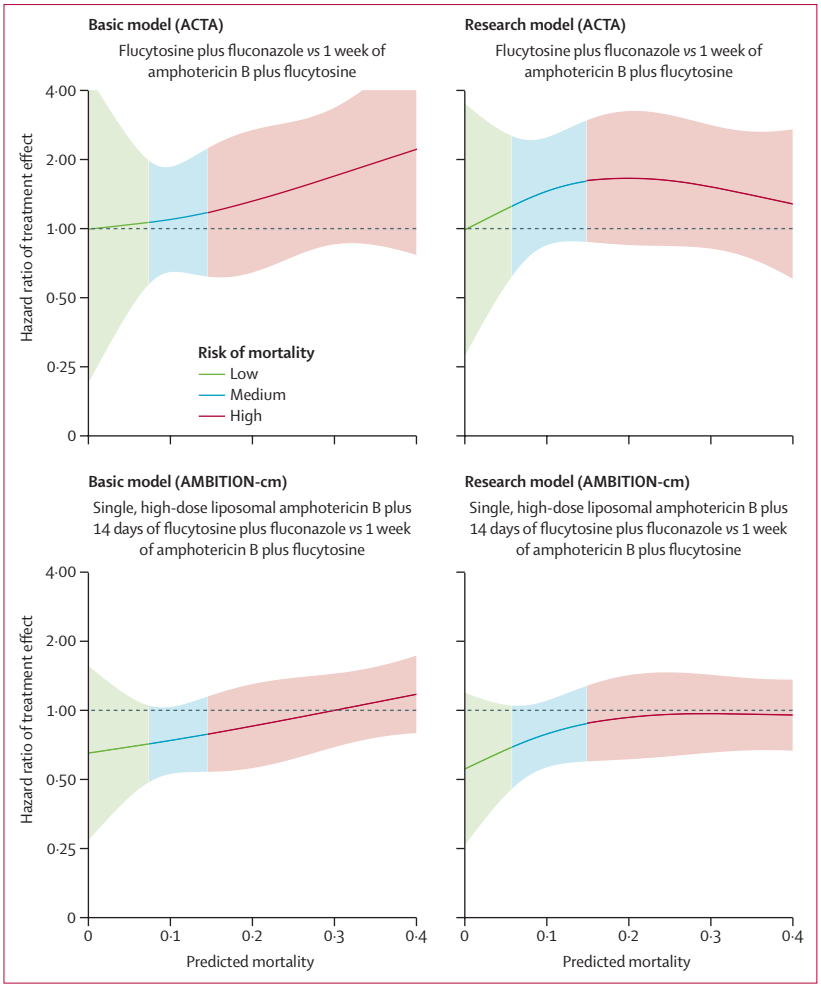

Figure 4: **Treatment effects across mortality predicted by models**
Treatment effect has been plotted logarithmically to base 2 on y axes. We foreshortened x axes to a predicted mortality of 0·4, which encompasses >95% of underlying data. Shaded areas show 95% CIs, solid lines show point estimates, and dashed grey lines indicate a hazard ratio of 1.

low-risk stratum and 1·67 (0·92–3·01) in the high-risk stratum for the research model (figure 4; appendix pp 31–32).

Discussion

We developed models to predict 2-week and 10-week mortality using a large, high-quality dataset containing the two largest randomised controlled trials conducted in sub-Saharan Africa to date for the treatment of cryptococcal meningitis. Our basic model included five predictor variables that are routinely measured programmatically, allowing for use in resource-limited settings. Our research model included two further predictor variables for use in settings where they are available. Both models showed consistent performance across all eight included countries, indicating their generalisability. Predicted risk was skewed, with modal mortality risk being far less than overall risk, suggesting that the majority of included people had individual-level risk much less than the mean and thus supporting an individualised approach to treatment.

Our results also provide proof-of-concept data for treatment stratification. Single, high-dose liposomal amphotericin B plus 14 days of flucytosine plus fluconazole was non-inferior to 1 week of amphotericin B plus flucytosine in AMBITION-cm overall.[5] In the current study, we found lower point estimates for mortality risk with the single, high-dose liposomal amphotericin B regimen among participants with lower predicted mortality risk, when using both categorical and continuous risk predictions, but this was not statistically significant for the research model. Treatment regimens for people with cryptococcal meningitis are intensive and use of conventional amphotericin B deoxycholate is associated with toxicity risk.[30] Treatment toxicity might be a particularly important factor in establishing outcomes among participants at low risk of mortality. When evaluating fluconazole plus flucytosine for 14 days in the ACTA trial, it was similar to the efficacy of 1 week of amphotericin B plus flucytosine in the participants at lowest risk of mortality. Our results suggest that our models could be used to direct stratified treatment approaches in future trials, with people at lower predicted risk of mortality receiving less toxic or intensive therapy and being considered for earlier hospital discharge than people at higher predicted risk. This effect was observed in both models, suggesting the basic model could be used if measurement of $\log_{10}$CSF quantitative cryptococcal culture and CSF opening pressure are not possible, broadening its use and practicality in low-resource settings.

Our models showed consistent discrimination across different study sites with some variation in calibration, likely reflecting site-level differences in baseline mortality risk that we did not account for in the model. The model showed systematic overprediction of risk in Malawi, indicating that people had less than expected mortality, and underprediction of risk in Uganda, indicating that people had greater than expected mortality from predictions from the models. Multiple explanations are plausible, including differences in health-care provision, genetic predispositions to outcomes, and other unmeasured determinants of outcome. Recalibration of the model intercepts to individual study sites resolved much of this difference, suggesting differences in baseline risk, rather than differing relationships between mortality and measured variables by site. Recalibration of the model intercept is a possible way to adapt the models to different settings in future, if required. In the held-out validation data, discrimination and calibration of both models improved on the most similar existing cryptococcal meningitis model for people living with HIV[9] and was consistent with that predicting COVID-19 mortality.[22] Our two models showed greater net benefit than alternative approaches to inform decision making in decision-curve analysis, without

recalibration. Notably, most of the observed miscalibration was at predicted risks of 25% or more, for which data were sparse. However, as this risk range is more than the likely threshold probability for most stratified interventions, it is unlikely to affect clinical use.

There were clinically plausible predictor–outcome associations for all included variables in the models and non-linear associations of continuous variables with mortality were explored. A GCS score of less than 15 and increased CSF fungal burden have been associated with increased mortality in previous studies from multiple settings.[8,9,12–15,20,21,31] Notably, CSF fungal burden at baseline, measured with quantitative cultures, had a non-linear association with mortality risk with risk increasing steeply above the threshold of $\log_{10}4$ CFU/mL, but little increased risk of mortality at lower levels of CSF quantitative culture. Although measurement of CSF fungal burden with quantitative cultures is not a routine laboratory procedure, results from point-of-care, semi-quantitative cryptococcal antigen tests or other approaches to quantify organism load, such as qPCR,[32] could be useful to increase the practicality of the research model in resource-limited settings.

ECOG performance status, reflecting functional status, was the most discriminative single predictor for mortality. As it is easy to measure at presentation, there is a strong rationale for considering this predictor both in research studies and programmatically. Increased intracranial pressure contributes to morbidity and mortality in HIV-associated cryptococcal meningitis[9,33,34] and is common at presentation. The risk associated with 2-week mortality increased slightly, with increasing CSF opening pressure in our study. As increased intracranial pressure was managed through protocolised therapeutic-lumbar puncture in ACTA and AMBITION-cm, the effect of unmanaged increased intracranial pressure on mortality was likely attenuated.[8] Thus, the association observed in our models should be considered to reflect intracranial pressure with protocolised management. Anaemia and high blood neutrophil count were also associated with increased 2-week mortality in our models, as observed for other infectious diseases, and likely reflect severity of systemic illness.[8] Most 2-week mortality risk likely relates to cryptococcal meningitis-related pathology, whereas over 10 weeks a larger proportion of observed risk could be attributable to underlying HIV and comorbidities. As 10-week mortality was greater than 2-week mortality, there were also more events for the 10-week outcome, which could have led to the inclusion of more variables and slightly better calibration observed for our 10-week models, compared with the primary 2-week models.

Although previous studies have sought to compare the performance of statistical and machine learning prediction models, most comparisons were classified at high risk of bias in a systematic review due to suboptimal statistical or machine learning methodology.[35] We found that performance of the machine learning model in our validation dataset showed no improvement in performance over the logistic-regression model, suggesting that, in datasets with low dimensions, traditional regression approaches could offer equivalent performance while remaining computationally less intensive and methodologically more transparent than machine learning approaches. Notably, our observed predictor–outcome associations from the XGBoost approach largely mirrored the associations found in the regression model.

Strengths of our study include our use of best practices for statistical and machine learning prediction model development and validation, including TRIPOD-standard reporting, use of multiple imputation to deal with missing data, and retaining continuous variables without arbitrary categorisation to avoid loss of information while accounting for non-linear associations. We also used the largest dataset to date to develop and validate two models for mortality risk and mitigated the risk of overfitting our models by defining candidate predictors a priori, in line with best practice sample-size guidance.[36] Furthermore, our pooled dataset included eight countries, which we used to explore generalisability and geographical heterogeneity in model performance through internal–external cross-validation. Finally, we explored heterogeneity of treatment effect by predicted risk, in line with the Predictive Approaches to Treatment effect Heterogeneity statement.[29]

Our study also has limitations. First, we developed and validated our models with data from Africa, where there is the greatest burden of cryptococcal meningitis.[1] Evaluation in other world regions where HIV-associated cryptococcal meningitis is prevalent is needed to further explore generalisability. Second, although developing the models on data from randomised controlled trials was a strength in terms of data quality, real-world outcomes for cryptococcal meningitis might be inferior to those observed in the trials due to factors such as trial exclusion criteria and improved standards of care in clinical trials.[37] Further validation in programmatic cohorts is therefore also required. Third, although our treatment-effect analysis provides early proof-of-concept evidence for treatment stratification in people with cryptococcal meningitis, caution is required due to these analyses being exploratory. Fourth, although the models were not trained to predict differential treatment effects, their discrimination and calibration to predict 2-week mortality were likely optimistic across participants in the development dataset. Future studies are required to further evaluate differential treatment effects when stratified by our models, including randomised controlled trials incorporating approaches to treatment stratification.

In summary, we developed prognostic models for 2-week and 10-week mortality in people with HIV-associated cryptococcal meningitis and showed consistent performance across eight sub-Saharan African

For **all models** see https://www.periskope.org/cm/

countries. The models used commonly available predictors and are freely available, to direct future treatment-stratification approaches in clinical trials.

Contributors

SFM, TSH, JNJ, and RKG conceptualised the analysis. THAS, SFM, TSH, JNJ, and RKG wrote the analysis plan. SFM curated the pooled dataset. THAS and RKG conducted the analysis. THAS, SFM, and RKG wrote the first draft of the manuscript, with input from TSH and JNJ. RKG built the Rshiny app. WSL developed the PERISKOPE website. All authors reviewed and approved the manuscript, had full access to all the data in the study if they wished, and had final responsibility for the decision to submit for publication.

Declaration of interests

TSH receives speaker fees from Gilead Sciences. DBM receives speaker fees from Gilead Sciences and GSK. All other authors declare no competing interests.

Data sharing

De-identified individual-level data from trials included in the pooled analysis for this modelling study are available on reasonable request (defined as a written request to the corresponding author of the original trial,[4,5] including a scientific rationale and analysis plan, and subject to a data sharing agreement) with no date restrictions. Full details of the trials, including study protocols, are available in the original trials.[4,5] The analysis code is available at https://github.com/thasamuels/Periskope-CM.

Acknowledgments

The trials included in the pooled analysis for this modelling study were funded by the UK Medical Research Council (100504), the French National Agency for AIDS Research (ANRS12275), the Swedish International Development Cooperation Agency (TRIA2015–1092), Wellcome–Medical Research Council–UKAID joint global health trials (MR/P006922/1), and the European and Developing Countries Clinical Trials Partnership. JNJ is funded by a National Institute for Health and Care Research Global Health Research Professorship (RP-2017–08-ST2–012). This research was also funded by grants from the National Institute for Health and Care Research (NIHR302829 and NIHR303184) and by National Institute for Health and Care Research Biomedical Research Funding, paid to RKG's institution. We thank all site staff and participants contributing to the trials included in the pooled analysis.

References

1 Rajasingham R, Govender NP, Jordan A, et al. The global burden of HIV-associated cryptococcal infection in adults in 2020: a modelling analysis. *Lancet Infect Dis* 2022; **22:** 1748–55.

2 Mwaba P, Mwansa J, Chintu C, et al. Clinical presentation, natural history, and cumulative death rates of 230 adults with primary cryptococcal meningitis in Zambian AIDS patients treated under local conditions. *Postgrad Med J* 2001; **77:** 769–73.

3 Longley N, Muzoora C, Taseera K, et al. Dose response effect of high-dose fluconazole for HIV-associated cryptococcal meningitis in southwestern Uganda. *Clin Infect Dis* 2008; **47:** 1556–61.

4 Molloy SF, Kanyama C, Heyderman RS, et al. Antifungal combinations for treatment of cryptococcal meningitis in Africa. *N Engl J Med* 2018; **378:** 1004–17.

5 Jarvis JN, Lawrence DS, Meya DB, et al. Single-dose liposomal amphotericin B treatment for cryptococcal meningitis. *N Engl J Med* 2022; **386:** 1109–20.

6 Gaskell KM, Rothe C, Gnanadurai R, et al. A prospective study of mortality from cryptococcal meningitis following treatment induction with 1200 mg oral fluconazole in Blantyre, Malawi. *PLoS One* 2014; **9:** e110285.

7 Jackson AT, Nussbaum JC, Phulusa J, et al. A phase II randomized controlled trial adding oral flucytosine to high-dose fluconazole, with short-course amphotericin B, for cryptococcal meningitis. *AIDS* 2012; **26:** 1363–70.

8 Jarvis JN, Bicanic T, Loyse A, et al. Determinants of mortality in a combined cohort of 501 patients with HIV-associated cryptococcal meningitis: implications for improving outcomes. *Clin Infect Dis* 2014; **58:** 736–45.

9 Zhao T, Xu XL, Nie JM, et al. Establishment of a novel scoring model for mortality risk prediction in HIV-infected patients with cryptococcal meningitis. *BMC Infect Dis* 2021; **21:** 786.

10 de Oliveira L, Melhem MSC, Buccheri R, Chagas OJ, Vidal JE, Diaz-Quijano FA. Early clinical and microbiological predictors of outcome in hospitalized patients with cryptococcal meningitis. *BMC Infect Dis* 2022; **22:** 138.

11 Kendi C, Penner J, Koech J, et al. Predictors of outcome in routine care for cryptococcal meningitis in western Kenya: lessons for HIV outpatient care in resource-limited settings. *Postgrad Med J* 2013; **89:** 73–77.

12 Anekthananon T, Manosuthi W, Chetchotisakd P, et al. Predictors of poor clinical outcome of cryptococcal meningitis in HIV-infected patients. *Int J STD AIDS* 2011; **22:** 665–70.

13 Saag MS, Powderly WG, Cloud GA, et al. Comparison of amphotericin B with fluconazole in the treatment of acute AIDS-associated cryptococcal meningitis. *N Engl J Med* 1992; **326:** 83–89.

14 Dismukes WE, Cloud G, Gallis HA, et al. Treatment of cryptococcal meningitis with combination amphotericin B and flucytosine for four as compared with six weeks. *N Engl J Med* 1987; **317:** 334–41.

15 Day JN, Chau TTH, Wolbers M, et al. Combination antifungal therapy for cryptococcal meningitis. *N Engl J Med* 2013; **368:** 1291–302.

16 Qu J, Jiang J, Lv X. The utility of cerebrospinal fluid white cell count during the prognostic assessment for cryptococcal meningitis patients: a retrospective study. *BMC Infect Dis* 2020; **20:** 571.

17 Chesdachai S, Engen NW, Rhein J, et al. Baseline serum C-reactive protein level predicts mortality in cryptococcal meningitis. *Open Forum Infect Dis* 2020; **7:** ofaa530.

18 de Jong VMT, Rousset RZ, Antonio-Villa NE, et al. Clinical prediction models for mortality in patients with COVID-19: external validation and individual participant data meta-analysis. *BMJ* 2022; **378:** e069881.

19 Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ* 2020; **370:** m3339.

20 Zhang C, He Z, Tan Z, Tian F. The clinic-based predictive modeling for prognosis of patients with cryptococcal meningitis. *BMC Infect Dis* 2023; **23:** 352.

21 Liu J, Lu Y, Liu J, et al. Development and validation of a machine learning model to predict prognosis in HIV-negative cryptococcal meningitis patients: a multicenter study. *Eur J Clin Microbiol Infect Dis* 2023; **42:** 1183–94.

22 Gupta RK, Harrison EM, Ho A, et al. Development and validation of the ISARIC 4C deterioration model for adults hospitalised with COVID-19: a prospective cohort study. *Lancet Respir Med* 2021; **9:** 349–59.

23 Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control* 1974; **19:** 716–23.

24 Debray TPA, Riley RD, Rovers MM, Reitsma JB, Moons KGM. Individual participant data (IPD) meta-analyses of diagnostic and prognostic modeling studies: guidance on their use. *PLoS Med* 2015; **12:** e1001886.

25 Gupta RK, Calderwood CJ, Yavlinsky A, et al. Discovery and validation of a personalized risk predictor for incident tuberculosis in low transmission settings. *Nat Med* 2020; **26:** 1941–49.

26 Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014; **35:** 1925–31.

27 Chen T, He T. XGBoost: extreme gradient boosting. 2024. https://cran.r-project.org/web/packages/xgboost/vignettes/xgboost.pdf (accessed May 17, 2024).

28 Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016; **352:** i6.

29 Kent DM, Paulus JK, van Klaveren D, et al. The Predictive Approaches to Treatment Effect Heterogeneity (PATH) statement. *Ann Intern Med* 2020; **172:** 35–45.

30 Bicanic T, Bottomley C, Loyse A, et al. Toxicity of amphotericin B deoxycholate-based induction therapy in patients with HIV-associated cryptococcal meningitis. *Antimicrob Agents Chemother* 2015; **59:** 7224–31.

31 Majumder S, Mandal SK, Bandyopadhyay D. Prognostic markers in AIDS-related cryptococcal meningitis. *J Assoc Physicians India* 2011; **59:** 152–54.

32 Mbangiwa T, Sturny-Leclère A, Lechiile K, et al. Development and validation of quantitative PCR assays for HIV-associated cryptococcal meningitis in sub-Saharan Africa: a diagnostic accuracy study. *Lancet Microbe* 2024; **5:** e261–71.

33 Graybill JR, Sobel J, Saag M, et al. Diagnosis and management of increased intracranial pressure in patients with AIDS and cryptococcal meningitis. The NIAID Mycoses Study Group and AIDS Cooperative Treatment Groups. *Clin Infect Dis* 2000; **30:** 47–54.

34 van der Horst CM, Saag MS, Cloud GA, et al. Treatment of cryptococcal meningitis associated with the acquired immunodeficiency syndrome. National Institute of Allergy and Infectious Diseases Mycoses Study Group and AIDS Clinical Trials Group. *N Engl J Med* 1997; **337:** 15–21.

35 Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 2019; **110:** 12–22.

36 Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ* 2020; **368:** m441.

37 Lawrence DS, Leeme T, Mosepele M, Harrison TS, Seeley J, Jarvis JN. Equity in clinical trials for HIV-associated cryptococcal meningitis: a systematic review of global representation and inclusion of patients and researchers. *PLoS Negl Trop Dis* 2021; **15:** e0009376.