Derivation and external validation of a model to predict 2-year mortality risk of patients with advanced schistosomiasis after discharge

Guo Li a, Shanshan Huang a, Lifei Lian a, Xiaoyan Song a, Wenzhe Sun a, Jinfeng Miao a, Bohan Li b, Yong Yuan b, Shengfan Wu b, Xiaoyan Liua, Wenzhe Sun a

a Department of Neurology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 Jiefang Avenue, Wuhan, Hubei 430030, China
b College of Mathematics and Statistics, Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan, Hubei 430074, China

Abstract

To date, no risk prediction tools have been developed to identify high mortality risk of patients with advanced schistosomiasis within 2 years after discharge. We aim to derive and validate a risk prediction model to be applied in clinical practice. The risk prediction model was derived from 1487 patients from Jingzhou and externally validated by 723 patients of Huangshi, two prefecture-level cities in Hubei province, China (from September 2014 to January 2015, with follow-up to January 2017). The baseline variables were collected. The mean age [SD] was 62.89 [10.38] years for the derivation cohort and 62.95 [12.22] years for the external validation cohort. The females accounted for 36.3% and 43.7% of the derivation and validation cohorts, respectively. 8.27% patients [SD] suffered from advanced schistosomiasis [8]. These cases were registered by the Centers for Disease Control and Prevention in China in 2017, and managed independently, since the disease results in high levels of mortality and disability as well as poor quality of life.

1. Introduction

Schistosomiasis is caused by infection with worms of the trematode genus Schistosoma, which is a neglected tropical disease affecting up to 250 million people in 76 countries [1]. The number of disability-adjusted life years (DALYs) lost caused by this disease was estimated up to 3.5 million in 2015 [2,3]. In China, schistosomiasis is caused by Schistosoma japonicum. The disease has long captured the attention of the Chinese authorities who have undertaken remarkably successful control programs to substantially reduced the schistosomiasis disease burden over the past 70 years [4].

Advanced schistosomiasis is a chronic disease caused by repeated infection of Schistosomiasis japonicum, which is associated with portal hypertension, splenomegaly, ascites, and gastro-esophageal variceal bleeding, or with granulomatous disease of the colon or severe growth retardation [5]. It has been classified into four clinical sub-types based on patients' major symptoms, namely ascites, splenomegaly, colonic tumour proliferation, and dwarfism. Over the past 70 years, significant achievements on schistosomiasis control have been attained through ongoing national control programs in China. At present, dwarfism and colonic tumour proliferation are rarely found [6,7]. However, ascites and splenomegaly are still common, typically in foci of high transmission intensity [5]. Based on the national schistosomiasis report issued by the Centers for Disease Control and Prevention in China in 2017, there were 37,601 cases of S. japonicum infection and 29,407 patients suffered from advanced schistosomiasis [8]. These cases were registered and managed independently, since the disease results in high levels of mortality and disability as well as poor quality of life.

Because of the serious consequences of advanced schistosomiasis, a more evidence-based management approach has been advocated by clinical experts [9]. The current consensus among the experts is that patients should be followed for 1 year at least to determine whether their
symptoms such as ascites or hemorrhage of upper digestive tract should be controlled and whether they need further treatment [9,10].

However, many patients with advanced schistosomiasis were not sufficiently followed up and were unable to adequately access medical care when their condition exacerbated, which resulted in inadequate opportunities to timely intervene and improve long-term prognosis. On the other hand, not all patients with advanced schistosomiasis progressed to death in 2 years after discharge. The follow-up of all advanced schistosomiasis patients with high 2-year mortality, the analysis also has the potential to improve the long term outcome of them.

Research in context

Evidence before this study

We searched the PubMed database according to the terms (“prediction” OR “risk prediction” OR “prediction model” OR “predictive” OR “predictive modeling”) AND (“advanced schistosomiasis”) among English-language articles before July 13th, 2019. We identified one study using three data mining models to predict the prognosis of patients with advanced schistosomiasis and another study using D-dimer to predict the occurrence of ascites. None of these studies have attempted to predict the 2-year mortality risk of patients with advanced schistosomiasis after discharge based on commonly used clinical indicators at admission, or discussed the performance between different mortality risk prediction models. We hypothesized that applying some commonly used clinical indicators at admission into a risk prediction model can predict 2-year mortality of patients with advanced schistosomiasis and improve the long term outcome of them.

Added value of this study

Our study constructed and compared 4 different models based on some commonly used clinical parameters to predict the 2-year mortality risk of patients with advanced schistosomiasis after discharge. We validated our models in internal and external cohorts by IDI values, NRI values, C statistics and ROC curves. A 7-variable model showed the best performance in the internal and external validation cohorts. Such risk prediction model could help guide follow-up, aid prognostic assessment and inform resource allocation since the disease carries a high risk of complications.

Implications of all the available evidence

To solve the problem of mortality risk prediction model shortages in patients with advanced schistosomiasis, we built a prediction model using age, clinical classification and some routine laboratory test indices at admission to predict the 2-year mortality risk after discharge. The model had shown good predictive performance.

2. Methods

We constructed multivariate prediction models that followed the TRIPOD statement [13] (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis). All procedures performed in this study involving human participants were in accordance with the 1964 Helsinki declaration and its later amendments. Patients’ confidentiality was protected by ensuring that the data were addressed in anonymous mode with personal information appropriately de-identified. The study was approved by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology. It met the definition of “minimal risk” and a waiver of informed consent was granted. Since this is a retrospective research involving no interventions, the waiver of consent will not adversely affect any rights and welfare of the subjects.

2.1. Study population and outcome

We formed a derivation and a validation cohort using a previously constructed database of patients with advanced schistosomiasis from Hubei Province, China. The database was a population-based repository that included blood biochemical measurements from ~4000 cases at admission [11]. Patients were diagnosed according to the Diagnostic Criteria for Schistosomiasis (WS261–2006) issued by China’s Ministry of Health which covered: (1) the patient who had ever lived in S. japonicum endemic region and has contact history with S. japonicum; (2) involvement of ascites, splenomegaly, portal hypertension and gastro-esophageal variceal bleeding, or with granulomatous lesion of the colon and rectum or severe growth retardation; (3) anti-S. japonicum antibody could be detected using at least one of the following tests: ELISA, IHA, DDI, COPT or DIFA; and (4) the result of stool examination or rectal biopsy was positive.

The discharge time window for the patients was from September 2014 to January 2015. These patients were followed to January 2017. All patients were advised to receive regular follow-ups. Patients were generally followed up every 3 months in the first 2 years and annually in the following years. Patients who did not visit our hospital as scheduled were telephoned for follow-ups to obtain the treatment information and vital status. The primary outcome was 2-year all-cause death after discharge. The study period was defined as the time from the discharge to the date of the last follow-up or death. The all-cause death events which occurred after January 2017 were considered nonevents for our study.

There are several standard treatment plans based on clinical classification, specific symptoms and disease severity. Splenectomy was provided to the splenomegaly type patients if hypersplenism symptom existed. Symptomatic treatment and liver protection measures were applied for ascites type patients. The praziquantel (PZD) treatment was administrated to patients after a stable period of 6 months during which they were in a good general condition (e.g. no ascites or hemorrhage symptoms).

2.2. Inclusion and exclusion criteria in derivation and validation cohorts

The patients with advanced schistosomiasis from Jingzhou served as the model derivation cohort. Approximately 50% of them were randomly selected as the internal validation cohort. The external validation cohort consisted of patients from Huangshi over the same time period.

The patients with advanced schistosomiasis in the derivation cohort were from Jingzhou that covers 9 counties. The inclusion criteria were as follows: (1) diagnosed as advanced schistosomiasis; (2) patients
with a long-term repeated history of infected water exposure or a definitive treatment history; (3) patients hospitalized to receive surgical or medical treatment. We excluded the patients with incomplete information, including lack of survival outcome, information on clinical characteristics, or laboratory test data. The number of patients with dwarfism and colonic tumoroid proliferation was too small to have any impact on the final analysis and were excluded in the analyses. The inclusion and exclusion criteria in external validation were the same as the derivation cohort.

Using the data from the derivation cohort, we sought to identify risk predictors of 2-year mortality after discharge through univariate and multivariate Cox proportional hazards models. The performance of the models was evaluated in the internal and external cohorts.

2.3. Candidate predictor variables

We chose 16 baseline candidate variables based on clinical practice and literature review. Candidate variables should be easily obtained in clinical settings and thus served as convenient predictors to be translated into future studies [14]. These variables included: age, gender, nourishment status, history of splenectomy, other diseases (cardiovascular diseases, digestive diseases and others), clinical classification, course of disease, frequency of ascites, serum total bilirubin (TBil), direct bilirubin (DBil), alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), albumin (ALB), Hepatitis B surface antigen (HBsAg) and alpha fetoprotein (AFP).

Patients received routine laboratory tests at admission. Serum samples were collected and clotted at room temperature, then centrifuged at 3500 r/min for 10 min, which could be used to estimate the levels of serum biomarkers. The baseline information, including age, gender, history of splenectomy, other diseases, course of disease and frequencies of ascites were extracted using standard sociodemographic and epidemiological questionnaires. The nourishment status of these patients was evaluated by measuring the triceps skinfold (TSF). Normal adult TSF values were defined as 12.5 mm in males and 16.5 mm in females. Well nourishment status referred to the measured value >90% of the normal adult TSF value. General nourishment status referred to the measured value equivalent to 80% – 90% of the normal adult TSF value. Poor nourishment status referred to the measured value <80% of the normal adult TSF value. Clinical classification was also identified according to WS261–2006.

2.4. Model derivation

Firstly, restricted cubic splines (RCS) were applied to detect the possible nonlinear dependency of the relationship between 6 continuous serum biochemical variables (TBil, DBil, ALT, AST, ALP and ALB) and the mortality risk, using 4 knots at the 5th, 25th, 75th, and 95th percentiles of corresponding variable (Table S1, Fig. S1). There were potential threshold associations between serum albumin (ALB) and the 2-year mortality risk of advanced schistosomiasis patients after discharge (P value for nonlinear=0.001). We have categorized the serum ALB to ≤45 g/L and >45 g/L groups based on the RCS curve. The other 5 serum biochemical indices (TBil, DBil, ALT, AST and ALP) were categorized to normal and abnormal groups based on the medical reference value, respectively.

The descriptive statistics was used to analyze the baseline information in derivation, internal validation and external validation cohorts. Univariate and multivariable Cox proportional hazards models were then used for variables selection. Each variable was first screened in the univariate model. The potential interactions between selected covariates after univariate analysis were also examined by multiple collinearity diagnosis using variance inflation factors (VIF) test before multivariate analysis. The existence of multi-collinearity between co-variates were determined by the VIF value (VIF value > 5) and tolerances (tolerance < 0.2). Multivariate Cox proportional hazards model was applied for all the statistically significant variables in univariate analysis, using backwards elimination. Variables with statistically significant differences in the multivariate Cox proportional hazard model (log-rank test, P < .05) were chosen to build the mortality risk prediction model. The results were presented as hazard ratios (HRs) and 95% confidence intervals (CIs).

The selected variables from the derivation cohort were included as potential covariates in multivariate logistic regression models. We then fit a series of reduced models by removing 2–3 variables at a time according to the specific clinical significance and compared the simplified models with the full multivariable model, using C statistic (equivalent to AUC value), net reclassification improvement (NRI) and integrated discrimination improvement (IDI). Reduced models were composed of variables more easily obtained during hospitalization. To further support the clinical use of this model, we developed a simple scale for measuring patients’ mortality risk in form of a risk index. The performance of the risk index was determined by its accuracy in predicting patient’s death outcome.

2.5. Prediction model performance

The regression coefficients from the logistic regression models in the derivation sample were fixed. We applied the fitted model to the internal and external validation cohorts. The Akaike Information Criteria (AIC) were applied to compare the overall goodness-of-fit of these models.

The discrimination abilities of the models were compared by IDI values, C statistics and Receiver operating characteristic (ROC) curves [15,16].

The IDI value reflects the percent improvement in discrimination attained by the full multivariate model over the reduced model with fewer variables.

\[
\text{IDI} = a - b \tag{1}
\]

\[a = (\text{mean predicted probability in patients with death outcome} - \text{mean predicted probability in patients without death outcome}) \text{ in the full multivariate model, b = (mean predicted probability in patients with death outcome} - \text{mean predicted probability in patients without death outcome}) \text{ in the reduced model with fewer variables.}

Model calibration was assessed by the calibration slope, the calibration intercept and locally weighted scatterplot smoothing (LOESS) plots of observed predicted probabilities of the outcome. We applied the NRI values to compare the ability of models to reclassify patients into lower-risk category. Similarly, we also regarded the worsening of classification as movement of non-events into a higher-risk category and non-events into a lower-risk category. The model with better performance can predict higher event rates for those with the events.

The NRI value was calculated by the sum of the percent improvement for events and nonevents, using the full multivariate model versus the reduced model with fewer variables [17].

\[
NRI = (a + b)_{\text{events}} + (b-a)_{\text{nonevents}} \tag{2}
\]

\[a = \text{proportion in higher-risk category based on predicted values from the full multivariate model} \text{ rather than the reduced model with fewer variables. b = proportion in lower-risk category based on predicted values from the full multivariate model rather than the reduced model with fewer variables.}
We also applied the continuous NRI values to compare the reclassification capacities of various models across a continuous range of risk thresholds. P value < 0.05 was considered statistically significant using 2-sided testing. Statistical analyses were performed using SPSS version 22.0 software (SPSS Inc., Chicago, IL, USA) and R version 3.5.2 software (The R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org).

The R packages “PredictABEL”, “rms”, “survival”, “survminer”, “cmprsk”, “Hmisc”, “lattice”, “Formula” and “grid” were applied in data analyses. ROC analyses were performed using MedCalc software version 15.8 (MedCalc Software, Ostend, Belgium) by DeLong function.

3. Results

3.1. Characteristics of the cohorts

A total of 1487 advanced schistosomiasis patients from Jingzhou were included in the derivation cohort, 700 of which were randomly selected as the internal validation cohort. 723 patients from Huangshi were included in the external validation cohort (Fig. 1). The mean age [SD] was 62.89 [10.38] years for the derivation cohort and 62.95 [12.22] years for the external validation cohort. The females accounted for 36.3% and 43.7% of the derivation and validation cohorts, respectively. General nourishment status accounted for 70.07% and 97.79% of the derivation and validation cohorts, respectively. Older age, splenomegaly clinical classification, abnormal serum DBil, AST, ALP, positive HBsAg and AFP (Table 2) were associated with a higher 2-year mortality risk in the derivation cohort. There was no potential interactions between co-variates based on the variance inflation factor (all VIF values < 5) and tolerance (all tolerances > 0.2) (Table S2).

We constructed a full multivariate model and three reduced models according to specific clinical significance. We took age and clinical classification as general demographic information. Abnormal serum DBil, AST and ALP were considered as indexes of liver function deterioration. Positive HBsAg and AFP were regarded as indexes reflecting coinfection with HBV and schistosomiasis japonica and potential risk of evolving to primary hepatic carcinoma. Therefore, model 1 included age, clinical classification, serum DBil, AST, ALP, HBsAg and AFP; model 2 included age, clinical classification, serum DBil, AST, ALP; model 3 included age, clinical classification, HBsAg and AFP; model 4 included serum DBil, AST, ALP, HBsAg and AFP.

The 7-variable model (model 1, Table 3) had the highest C statistic (0.785; 95% CI, 0.744–0.825) and lowest Akaike Information Criterion (AIC = 739.16). Discrimination capacities (based on the C statistic, ROC curves and IDI) were lower in the reduced models with fewer variables. The ROC curves, IDI, continuous NRI and categorization capacities of various models across a continuous range of risk thresholds. P value < 0.05 was considered statistically significant using 2-sided testing. Statistical analyses were performed using SPSS version 22.0 software (SPSS Inc., Chicago, IL, USA) and R version 3.5.2 software (The R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org).

3.2. Predict 2-year mortality risk in the derivation cohort

Seven variables were selected by univariate and multivariate Cox proportional hazards model analysis using backwards elimination. Older age, splenomegaly clinical classification, abnormal serum DBil, AST, ALP, positive HBsAg and AFP (Table 2) were associated with a higher 2-year mortality risk in the derivation cohort. There was no potential interactions between co-variates based on the variance inflation factor (all VIF values < 5) and tolerance (all tolerances > 0.2) (Table S2).

We constructed a full multivariate model and three reduced models according to specific clinical significance. We took age and clinical classification as general demographic information. Abnormal serum DBil, AST and ALP were considered as indexes of liver function deterioration. Positive HBsAg and AFP were regarded as indexes reflecting coinfection with HBV and schistosomiasis japonica and potential risk of evolving to primary hepatic carcinoma. Therefore, model 1 included age, clinical classification, serum DBil, AST, ALP, HBsAg and AFP; model 2 included age, clinical classification, serum DBil, AST, ALP; model 3 included age, clinical classification, HBsAg and AFP; model 4 included serum DBil, AST, ALP, HBsAg and AFP.

The 7-variable model (model 1, Table 3) had the highest C statistic (0.785; 95% CI, 0.744–0.825) and lowest Akaike Information Criterion (AIC = 739.16). Discrimination capacities (based on the C statistic, ROC curves and IDI) were lower in the reduced models with fewer variables. The ROC curves, IDI, continuous NRI and categorization capacities of various models across a continuous range of risk thresholds. P value < 0.05 was considered statistically significant using 2-sided testing. Statistical analyses were performed using SPSS version 22.0 software (SPSS Inc., Chicago, IL, USA) and R version 3.5.2 software (The R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org).

3.3. Predict 2-year mortality risk in the internal validation cohort

In the internal validation cohort, the 7-variable model also had a higher C statistic (0.763; 95% CI, 0.676–0.838) than the reduced models with fewer variables. The ROC curves, IDI, continuous NRI and categorical NRI favored the 7-variable model compared with the reduced models. For example, significantly improved classification into risk

Advanced schistosomiasis disease database from HuBei Province, P.R. China

756 Patients from Huangshi District with confirmed diagnosis of advanced schistosomiasis (September, 2014–January 2015)

33 Excluded
11 Insufficient follow-up information
21 Lack of serum biochemicals data
1 Colonic tumoroid proliferation type

723 Included (follow up to January, 2017)

1487 Included (follow up to January, 2017)

1545 Patients from Jingzhou District with confirmed diagnosis of advanced schistosomiasis (September, 2014–January 2015)

58 Excluded
20 Insufficient follow-up information
35 Lack of serum biochemicals data
3 Colonic tumoroid proliferation type

1487 Included (follow up to January, 2017)

Derivation Cohort
1487 Included
123 with death outcome in 2 years

Internal Validation Cohort
760 cases were randomly selected from derivation cohort
55 with death outcome in 2 years

Fig. 1. Formation of the Derivation, Internal Validation, and External Validation Cohorts.
categories was observed (Table 3, Fig. 2) in the 7-variable model 1 compared to model 2 (Continuous NRI, 28.25%; 95% CI, 2.73–53.80%; \( P = 0.03 \)), model 3 (Continuous NRI, 56.70%; 95% CI, 30.0–83.40%; \( P < 0.001 \)). Categorical NRI, 18.54%; 95% CI, 4.34–32.56%; \( P = 0.01 \)) and model 4 (Continuous NRI, 52.70%; 95% CI, 25.80–79.60%; \( P < 0.001 \)). Categorical NRI, 23.13%; 95% CI, 9.20–37.10%; \( P = 0.0011 \). The IDI index also obtained improvement in the model 1 (all \( P < 0.05 \)) compared to model 2–4.

The risk index (Fig. 3) was developed from the 7-variable model. Fig. 4A demonstrated observed versus predicted risks based on the risk index.

### 3.4. Predict 2-year mortality risk in the external validation cohort

In the external validation cohort, the 7-variable model 1 (0.717; 95% CI, 0.646–0.788) and 5-variable model 2 (0.717; 95% CI, 0.647–0.787) also had a higher C statistic than model 3 and model 4. The ROC curves, IDI, continuous NRI and categorical NRI favored the 7-variable model compared with the reduced models. For example, significantly improved classification into risk categories was observed (Table 3, Fig. 2) in the 7-variable model 1 compared to model 2 (Categorical NRI, 1.35%; 95% CI, 0.38–2.32%; \( P = 0.0064 \)), model 3 (Continuous NRI, 53.90%; 95% CI, 28.0–79.80%; \( P < 0.001 \)). Categorical NRI, 17.32%; 95% CI, 0.23–34.42%; \( P = 0.047 \)) and model 4 (Continuous NRI, 27.40%; 95% CI, 0.22–54.60%; \( P = 0.0482 \)). Categorical NRI, 11.46%; 95% CI, 1.68–21.25%; \( P = 0.022 \)). The IDI value also obtained improvement in model 1 (all \( P < 0.05 \)) compared to model 2–4. (Table 4)

### 4. Discussion

The AUC values of model 3–4 were significantly less than model 1 in derivation cohort. The AUC value of model 1 was not significantly higher than model 2 in derivation cohort. However, significantly improved classification abilities were observed in model 1 compared to model 2 in internal and external validation cohort. The IDI value, which represented the improvement of model’s overall performance, had also shown improvement in model 1 (all \( P < 0.05 \)) when compared to model 2–4 in internal and external cohorts. Furthermore, the selected 7 variables can be obtained readily in routine clinical practice with low cost economic. Therefore, we recommended model 1 with age, clinical classification, baseline serum DBil, AST, ALP, HBsAg and AFP at admission as the optimal 2-year mortality risk prediction model for patients with advanced schistosomiasis after discharge.

The predictor variables incorporated in this model and the risk index suggest a new plausible tool for clinicians during the discharge planning process to screen advanced schistosomiasis patients at high mortality risk. Such risk stratification tool can help guide follow-up, aid prognostic evaluation and inform resource allocation, since advanced schistosomiasis disease carries a high risk of complications and leads to high resource requirements to provide medical care. The predictor variables in the risk models have been studied before. Similar to previous studies, the highest schistosomiasis-related mortality was observed in older age groups [18]. This can be explained by the chronic nature of the disease, evolving into severe clinical forms, as well as more frequent chronic comorbidities in the elderly, may possibly increase the mortality risk [19]. With regard to clinical classification, splenomegaly type patients have complex disease conditions and many complications after splenectomy, especially combined with hepatic encephalopathy and rebleeding, such as esophageal varices bleeding, which may lead to higher mortality risk than ascites type patients [20,21].

Previous studies also suggested that HBV infection exacerbates liver damage due to schistosomiasis, for coinfection with HBV and S. japonicum is associated with accelerated deterioration in hepatic function, which can lead to more severe fibrosis and inflammatory activity in the liver [22]. In the advanced stage of schistosomiasis, the fibrosis and serious fibro-obstructive pathology lead to portal hypertension, ascites, hepatosplenomegaly and eventually, fatal hematemesis.

Furthermore, a portion of hepatitis B patients will develop chronic carriers, including cirrhosis and primary hepatic carcinoma, none of the most serious diseases threatening human health due to its high mortality. Similarly, the alpha-fetoprotein (AFP) has been also developed as a biomarker for the early diagnosis of primary liver cancer [23]. The HBsAg and AFP indices can be integrated into the mortality prediction model to express the combined risk of primary hepatic carcinoma.

Bilirubin has previously been proven to be a marker of liver injury and is incorporated in several prognostic scoring models, such as the model of end-stage liver disease (MELD) and the Child-Pugh (CP) [24]. In previous study, abnormal DBil was suggested to be independently associated with an increased risk of increased fibrosis indices [25]. Glucuronyl conjugation of bilirubin and biliary excretion of DBil are distinctly impaired in advanced cirrhosis [26], which was closely associated with poor prognosis of advanced schistosomiasis. Therefore, the abnormal serum DBil level can be used as a good prognostic marker for advanced schistosomiasis patients accompanied with decompensated liver cirrhosis, an apparent life-threatening event. There is a high risk of developing further fatal complications of hepatic cirrhosis, such as spontaneous bacterial peritonitis and hepatorenal syndrome (HRS) which are the most common decompensating events in patients with advanced schistosomiasis [27].
Elevated AST activity is a surrogate marker for cirrhosis because of reduced plasma clearance of AST secondary to impaired function of sinusoidal cells [28]. The aspartate aminotransferases (AST) to platelet ratio index (APRI) score is used to predict the degree of liver fibrosis caused by schistosome infection can be fatal. Of the routinely tested markers significantly associated with fibrosis, AST level was found to be a reliable and sensitive marker for differentiating significant hepatic fibrosis in patients with advanced schistosomiasis japonica in previous studies, including cases co-infected with HBV [31].

ALP is a hydrolytic enzyme that dephosphorylates and transphosphorylates molecules including nucleotides (adenosine triphosphate, adenosine biphosphate), pathogen-associated molecule patterns and danger-associated molecule patterns [32]. The serum ALP level increases under some pathological conditions including liver injury, hepatocellular carcinoma (HCC), cholangiocarcinoma and biliary cirrhosis.

| Variables | Overall survival | Univariate analysis | Multivariate analysis |
|-----------|-----------------|--------------------|---------------------|
| Age (years) |     | 1.062 (1.042–1.083) | 1.068 (1.046–1.089) |
| Gender | Male/Female | 0.881 (0.866–1.282) | 0.881 (0.866–1.282) |
| Nourishment status | Well/Poor | 1.066 (0.756–1.254) | 0.653 (0.438–0.973) |
| History of splenectomy | None/Yes | 1.033 (0.942–1.134) | 0.653 (0.438–0.973) |
| Other disease | None/Cardiovascular/Digestive/Other | 1.033 (0.942–1.134) | 0.653 (0.438–0.973) |
| Clinical classification | Splenomegaly/Ascites | 1.778 (1.130–2.798) | 1.783 (1.120–2.840) |
| Course of disease | ≤4/≤5 | 0.856 (0.600–1.220) | 0.856 (0.600–1.220) |
| Frequencies of ascites (times) | <5/5 | 0.837 (0.582–1.202) | 0.837 (0.582–1.202) |
| TBil | Normal/Abnormal | 3.056 (2.143–4.358) | 3.056 (2.143–4.358) |
| DBil | Normal/Abnormal | 3.674 (2.577–5.237) | 3.674 (2.577–5.237) |
| AST | Normal/Abnormal | 2.458 (1.711–3.311) | 2.458 (1.711–3.311) |
| ALP | Normal/Abnormal | 2.753 (1.883–4.025) | 2.753 (1.883–4.025) |
| HBsAg | Negative/Positive | 1.750 (1.178–2.591) | 1.750 (1.178–2.591) |
| AFP | Negative/Positive | 3.653 (1.853–7.201) | 3.653 (1.853–7.201) |

| Variables | Overall survival | Univariate analysis | Multivariate analysis |
|-----------|-----------------|--------------------|---------------------|
| Age (years) |     | 1.062 (1.042–1.083) | 1.068 (1.046–1.089) |
| Gender | Male/Female | 0.881 (0.866–1.282) | 0.881 (0.866–1.282) |
| Nourishment status | Well/Poor | 1.066 (0.756–1.254) | 0.653 (0.438–0.973) |
| History of splenectomy | None/Yes | 1.033 (0.942–1.134) | 0.653 (0.438–0.973) |
| Other disease | None/Cardiovascular/Digestive/Other | 1.033 (0.942–1.134) | 0.653 (0.438–0.973) |
| Clinical classification | Splenomegaly/Ascites | 1.778 (1.130–2.798) | 1.783 (1.120–2.840) |
| Course of disease | ≤4/≤5 | 0.856 (0.600–1.220) | 0.856 (0.600–1.220) |
| Frequencies of ascites (times) | <5/5 | 0.837 (0.582–1.202) | 0.837 (0.582–1.202) |
| TBil | Normal/Abnormal | 3.056 (2.143–4.358) | 3.056 (2.143–4.358) |
| DBil | Normal/Abnormal | 3.674 (2.577–5.237) | 3.674 (2.577–5.237) |
| AST | Normal/Abnormal | 2.458 (1.711–3.311) | 2.458 (1.711–3.311) |
| ALP | Normal/Abnormal | 2.753 (1.883–4.025) | 2.753 (1.883–4.025) |
| HBsAg | Negative/Positive | 1.750 (1.178–2.591) | 1.750 (1.178–2.591) |
| AFP | Negative/Positive | 3.653 (1.853–7.201) | 3.653 (1.853–7.201) |

Table 2
Univariate and multivariate Cox proportional hazards regression analysis for variables selection.

Table 3
Death events predictors of advanced schistosomiasis and performance of models in the derivation cohort.

Predictors | Models 1, Age, Clinical classification, DBil, AST, ALP, HbsAg, AFP | Models 2, Age, Clinical classification, DBil, AST, ALP | Models 3, Age, Clinical classification, HbsAg, AFP | Models 4, DBil, AST, ALP, HbsAg, AFP
|-------------|-------------------------------------------------|----------------------|-----------------|------------------------|
| Age, per year increase | 1.075 (1.052–1.099) | 1.067 (1.045–1.091) | 1.077 (1.054–1.101) | 1.077 (1.054–1.101) |
| Clinical classification | Ascites [Reference] | 1.863 (1.097–3.165) | 1.793 (1.063–3.025) | 2.100 (1.266–3.482) |
| Splenomegaly | Normal [Reference] | 1.863 (1.097–3.165) | 1.793 (1.063–3.025) | 2.100 (1.266–3.482) |
| DBil | Normal [Reference] | 2.580 (1.703–3.907) | 2.809 (1.870–4.220) | 2.967 (1.990–4.422) |
| AST | Normal [Reference] | 1.875 (1.221–2.880) | 1.930 (1.263–2.949) | 1.630 (1.076–2.470) |
| ALP | Normal [Reference] | 1.832 (1.149–2.920) | 1.826 (1.155–2.888) | 1.959 (1.250–3.070) |
| HBsAg | Negative [Reference] | 2.169 (1.359–3.463) | 2.361 (1.506–3.702) | 1.640 (1.050–2.563) |
| AFP | Negative [Reference] | 2.246 (1.644–3.366) | 3.741 (1.634–8.564) | 1.939 (0.823–4.571) |

Model Performance Measures

| Models | HR 95% CI | HR 95% CI | AIC | C statistic | Difference | P value | IDI (95%CI), % | P value |
|--------|-----------|-----------|-----|-------------|------------|---------|----------------|---------|
| Model 1 | 739.16 | 0.785 (0.744–0.825) | 0.768 (0.726–0.809) | -0.161 (0.003–0.305) | 0.096 | 0.007 | 0.0158 (0.0043–0.0272) | 0.0533 (0.0344–0.0722) |
| Model 2 | 749.85 | 0.785 (0.744–0.825) | 0.768 (0.726–0.809) | -0.161 (0.003–0.305) | 0.096 | 0.007 | 0.0158 (0.0043–0.0272) | 0.0533 (0.0344–0.0722) |

Notes:
- **P** value 0.05 indicate better performance for the full model 1 than for the reduced models.
- **P** value ≤ 0.05 indicate better performance for the full model 1 than for the reduced models.

**Abbreviations:** HR, hazard ratio; CI, confidence interval; NS, not significant; NI, not included.

*Of the 1487 patients in the derivation cohort, 123 patients had death events.

*Models 2 through 4 were each compared with model 1. For model comparisons, Integrated Discrimination Improvement (IDI) values and differences in a C statistic > 0 indicate better performance for the full model 1 than for the reduced models.*
The level of ALP was reported as an independent prognostic biomarker for patients with HCC [34] and as a variable integrated into Chinese University Prognostic Index (CUPI) system to predict survival outcome [35]. Some studies suggested that low AAPR, an index combined by ALB and ALP, may reflect a poor nutrition status, inactive immune reaction, exhausted liver function, all of which may lead to poor survival outcomes [36–38]. Moreover, ALP is an easy-to-get laboratory test variable for assessing liver function in clinical practice [37]. This work has integrated serum DBil, AST and ALP variables into multivariate risk models to reveal the deterioration of liver function. The above-mentioned 7 variables can be readily obtained and therefore integrated into mortality risk prediction models to estimate 2-year individualized mortality risk of patients with advanced schistosomiasis after discharge.

The predictive performance of 2-year mortality risk prediction models needs to be validated internally and externally. Both internal (C-statistic, 0.763; 95% CI, 0.696–0.830) and external validation results (C statistic of model 1, 0.717; 95% CI, 0.646–0.788) indicated that the models were likely not overfit and performed well. We identified the mortality risk models as useful clinical tool for decision making particularly when the C-statistic is higher than 0.70 [39]. Furthermore, the NRI of the full model compared with simpler models demonstrates that the model 1 could improve the accuracy of decision making compared with model 2–4 [15,16].

| Age years | Points |
|-----------|--------|
| <50       | 1      |
| 50–59     | 2      |
| 60–69     | 3      |
| 70–79     | 4      |
| ≥80       | 5      |

| Clinical classification | Points |
|------------------------|--------|
| Ascites                | 0      |
| Splenomegaly           | 1      |

| DBil Points |
|-------------|
| Normal      | 0       |
| Abnormal    | 1       |

| AST Points |
|------------|
| Normal     | 0       |
| Abnormal   | 1       |

| ALP Points |
|------------|
| Normal     | 0       |
| Abnormal   | 1       |

| HBsAg Points |
|--------------|
| Negative     | 0       |
| Positive     | 1       |

| AFP Points |
|------------|
| Normal     | 0       |
| Abnormal   | 1       |

| Total Risk Score |
|------------------|
| = (              ) |

Fig. 2. Reclassification Performances of Model 1–4 in the Different Cohorts. Model 1 included age, clinical classification, DBil, AST, ALP, HBsAg and AFP. Model 2 included age, clinical classification, DBil, AST, ALP. Model 3 included age, clinical classification, HBsAg and AFP. Model 4 included DBil, AST, ALP, HBsAg and AFP. a. ROC curves of the derivation cohort and AUC values (Model 1, 0.785; Model 2, 0.768; Model 3, 0.713; Model 4, 0.713). b. ROC curves of the internal validation cohort and AUC values (Model 1, 0.763; Model 2, 0.745; Model 3, 0.719; Model 4, 0.669). c. ROC curves of the external validation cohort and AUC values (Model 1, 0.717; Model 2, 0.717; Model 3, 0.669; Model 4, 0.684).

Fig. 3. Risk Index for Patients with Advanced Schistosomiasis after discharge. a. Point values for each variable (A patient’s total risk score can be obtained by summing points assigned to values of each variable, which can determine his/her corresponding predicted 2-year mortality risk). b. Predicted risk of 2-year mortality after discharge.
Strengths of this study include that the model was developed in a large, population-representative cohort. Since the Chinese government has embarked an effort to manage these advanced schistosomiasis cases independently and those registered in the Advanced Schistosomiasis Cases Management System receive a RMB 5000 subsidy yearly per capita for therapy, including anthelmintic therapy, antifibrosis treatment, diuresis, and hormonotherapy, the patients lost to follow-up were rare [40]. The predictor variables in these risk models...

### Table 4

Predictive Performance of Models for death events in Internal and External Validation Cohorts

| Models, Measures of Predictive Performance | Internal validation | External validation |
|------------------------------------------|---------------------|---------------------|
| Calibration intercept P value            | 0.233               | −0.396              |
| Calibration slope                        | 0.962               | 0.025               |
| P value                                  | 0.97                | 0.13                |
| C statistic (95%CI)                       | 0.763               | 0.717               |
| Difference in C statistic (95%CI)         | (0.696–0.83)        | (0.646–0.788)       |
| Difference in C statistic (95%CI)         | (0.677–0.812)       | (0.647–0.787)       |
| P value                                  | 0.018               | 0.048               |
| NRI (95%CI), %                           | 0.282               | 0.025               |
| P value                                  | 0.018               | 0.048               |
| NRI (95%CI), %                           | 0.282               | 0.025               |
| P value                                  | 0.018               | 0.048               |
| Calibration intercept P value            | 0.233               | −0.396              |
| Calibration slope                        | 0.962               | 0.025               |
| P value                                  | 0.97                | 0.13                |
| C statistic (95%CI)                       | 0.763               | 0.717               |
| Difference in C statistic (95%CI)         | (0.696–0.83)        | (0.646–0.788)       |
| Difference in C statistic (95%CI)         | (0.677–0.812)       | (0.647–0.787)       |
| P value                                  | 0.018               | 0.048               |
| NRI (95%CI), %                           | 0.282               | 0.025               |
| P value                                  | 0.018               | 0.048               |
| Calibration intercept P value            | 0.233               | −0.396              |
| Calibration slope                        | 0.962               | 0.025               |
| P value                                  | 0.97                | 0.13                |
| C statistic (95%CI)                       | 0.763               | 0.717               |
| Difference in C statistic (95%CI)         | (0.696–0.83)        | (0.646–0.788)       |
| Difference in C statistic (95%CI)         | (0.677–0.812)       | (0.647–0.787)       |
| P value                                  | 0.018               | 0.048               |
| NRI (95%CI), %                           | 0.282               | 0.025               |
| P value                                  | 0.018               | 0.048               |

Abbreviations: IDI, integrated discrimination improvement; NRI, net reclassification improvement; NR, not reportable (cells associated with 5 or fewer events are empty because of institute for Clinical Evaluation Science policy).

a Of the 700 patients in the internal validation cohort, 55 had death events, and 56 patients had death events of the 723 patients in the external cohort.

b Models 2 through 4 were each compared with model 1. For model comparisons differences in the C statistics, IDI, and NRI values > 0 indicate better performance for model 1 than the reduced models.

c Risk categories include patients with <10%, 10%–30%, 30% or higher risk of death.

d Net reclassification improvement events refer to occurrence of death, and NR nonevents refer to no occurrence of death.
consist of easily accessible patients’ demographics and laboratory test information, which could allow them to be implemented into clinical practice. The model was validated in a distinct, external cohort and had shown good predictive performance.

Our study also has several limitations. First, candidate variables were identified from secondary analysis of data, which does not include all potential risk factors, such as blood pressure, family nursing support and details of surgical or medical treatments. However, this study has included the most common and representative variables which were reported in other related literatures. Second, the participants did not have serum markers of liver fibrosis measured, such as Hyaluronic acid (HA), Laminin (LN), Collagen IV (CIV) and Procollagen III (PⅢ)，which may influence the performance of the models [41,42]. Likewise, the pathology evidence of advanced schistosomiasis can be directly obtained from ultrasound examination of the liver and spleen [43]. Unfortunately, we failed to collect the complete information about the patients’ ultrasonography examination results. However, models without serum markers of liver fibrosis and ultrasonography examination results of the liver and spleen also performed well, suggesting that the accurate risk prediction of the survival outcome could still be achieved without these indices. Third, models were derived and validated in cohorts from Hubei Province. The generalizability to patients of other regions needs further research.

This study may have implications for clinicians, patients and policy makers. Low rates of follow-up were demonstrated in these patients possibly because rural doctors lack the awareness of screening, or lack of health resources in these isolated rural villages. Hence, providing continuity of care between the hospital and community was inadequate in common clinical practice. The risk prediction model provides an accurate but simple method that can stratify patients into clinically meaningful risk groups at a time of hospital discharge and facilitate a further management in the local community and family.

In conclusion, we have identified the clinical predictors for 2-year mortality risk of patients with advanced schistosomiasis after discharge. We have also developed a prediction model that enables clinicians to estimate the 2-year mortality risk based on these variables. The prediction model has a good predictive performance through internal and external validation. The application of the model may materially improve the patients’ long-term outcome, provided it is followed by adequate treatment and follow-up.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ebiom.2019.08.028.

Funding sources

This work was financially supported by the National Key R&D Program of China [grant number 2017YFC1310000]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Z.Z. had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Authors’ contributions

Z.Z. and X.L. lead the study. G.L. performed the data analysis, implemented the methodology; X.S., W.S., J.M. collected the data; G.L. prepared the original draft; B.L., Y.Y. and S.W. helped to perfect the figures. S.H., L.L. and Z.Z. reviewed and edited the final manuscript.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgements

We would like to acknowledge all participants of this project and experts of local hospitals for collecting data. Also, we wish to thank Dr. Zhou Xiaorong for sharing data to support this study.

References

[1] Schistosomiasis: number of people receiving preventive chemotherapy in 2012, World Health Epidemol Rec 2014;89(2):21–8.
[2] Mitra AK, Mawson AR. Neglected tropical diseases: epidemiology and global burden. Trop Med Infect Dis 2017;2(3).
[3] Cai P, et al. Circulating miRNAs as footprints for liver fibrosis grading in schistosomiasis. ElBioMedicine 2018;37:334–43.
[4] McManus DP, et al. Schistosomiasis research in the dongting lake region and its impact on local and national treatment and control in China. PLoS Negl Trop Dis 2011;5(8):e1053.
[5] Jia TW, et al. Quantifying quality of life and disability of patients with advanced schistosomiasis japonica. PLoS Negl Trop Dis 2011;5(2):e696.
[6] Utzinger J, et al. Conquering schistosomiasis in China: the long march. Acta Trop 2005;96(2–3):69–96.
[7] Zhou XN, et al. The public health significance and control of schistosomiasis in China–then and now. Acta Trop 2005;96(2–3):97–105.
[8] Li-Jun Z, et al. Endemic status of schistosomiasis in People’s Republic of China in 2017. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi 2018;30(5):481–3.
[9] Zhong F, Liu C, Zhang X. Guideline adherence for the treatment of advanced schistosomiasis japonica in Hubei, China. Parasitol Res 2014;113(12):4535–41.
[10] Deng WC, et al. Diagnosis and treatment of schistosomiasis japonica–consensuses among experts in Hunan, Hubei and Jiangxi provinces. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi 2015;27(5):451–6.
[11] Li G, et al. Comparison of three data mining models for prediction of advanced schistosomiasis prognosis in the Hubei province. PLoS Negl Trop Dis 2012;6(2):e1000622.
[12] Wu X, et al. Plasma D-dimer can effectively predict the prospective occurrence of ascites in advanced schistosomiasis japonica patients. Korean J Parasitol 2017;55(2):167–74.
[13] Collins GS, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD). Ann Intern Med 2015;162(10):735–6.
[14] de Man-van Gijl, et al. In-hospital risk prediction for post-stroke depression: development and validation of the Post-stroke Depression Prediction Scale. Stroke 2013;44(9):2441–5.
[15] Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. Ann Intern Med 2009;150(11):795–802.
[16] Alba AC, et al. Discrimination and calibration of clinical prediction models: users’ guides to the medical literature. JAMA 2017;318(14):1377–84.
[17] Martins-Melo FR, et al. Trends in schistosomiasis-related mortality in Brazil, 2000–2011. Int J Parasitol 2014;44(14):1055–62.
[18] Renesandes AP, Souza-Santos R, Barbosa CS. Hospitalization and mortality from mansoni schistosomiasis in the state of Pernambuco, Brazil, 1992/2000. Cad Saude Publica 2005;21(5):1392–401.
[19] Jia-Xin L, et al. Investigation of direct medical expense for surgical patients with splenomegalic advanced schistosomiasis in Hunan Province from 2010 to 2014. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi 2016;28(8):365–9.
[20] Costa LC, et al. Schistosomal portal hypertension: randomized trial comparing endoscopic therapy alone or preceded by esophagogastric devascularization and splenectomy. Ann Hepatol 2016;15(5):738–46.
[21] Huang LH, et al. The efficacy of and safety of entecavir in patients with advanced schistosomiasis co-infected with hepatitis B virus. Int J Infect Dis 2013;17(8):e906–909.
[22] Ma H, et al. Multiplex immunochips for high-accuracy detection of AFP-L3x based on surface-enhanced Raman scattering: implications for early liver cancer diagnosis. Anal Chem 2017;89(17):8876–84.
[23] Shi QK, et al. Improvement of liver function and non-invasive fibrosis markers in hepatitis B virus-associated cirrhosis: 2 years of entecavir treatment. J Gastroenterol Hepatol 2015;30(12):1775–81.
[24] Du M, et al. The relationship between serum bilirubin and elevated fibrotic indices among HBV carriers: a cross-sectional study of a Chinese population. Int J Mol Sci 2016;17(12).
[25] Tsochatzis EA, Bosch J, Burroughs AK. Liver cirrhosis. Lancet 2014;383(9930):1740–61.
[26] Piano S, Tonon M, Angeli P. Management of ascites and hepatic encephalopathy. Hepatol Int 2018;12(Suppl. 1):122–34.
[27] Park GJ, et al. Aspartate aminotransferase: alanine aminotransferase ratio in chronic hepatitis C infection: is it a useful predictor of cirrhosis? J Gastroenterol Hepatol 2000;15(4):386–90.
[28] Wang CW, et al. Improvement of liver function and non-invasive fibrosis index in school-aged children living near a petrochemical complex. Environ Pollut 2015;209:648–56.
[29] Derbala M, et al. Aspartate transaminase to platelet ratio index in hepatitis B virus-associated cirrhosis: 2 years of entecavir treatment. J Gastroenterol Hepatol 2015;30(12):1775–81.
[30] Hou XY, et al. Diagnostic value of non-invasive bio-markers for stage-specific diagnosis of hepatic fibrosis in patients with advanced schistosomiasis japonica. Int J Parasitol 2011;41(3–4):325–32.
[31] Lalles JP. Intestinal alkaline phosphatase: multiple biological roles in maintenance of intestinal homeostasis and modulation by diet. Nutr Rev 2010;68(6):323–32.
[32] Williams-Koch K, Chapman RW. Editorial: further evidence for the role of serum alkaline phosphatase as a useful surrogate marker of prognosis in PSC. Aliment Pharmacol Ther 2015;41(1):149–51.
[33] Yu MC, et al. Alkaline phosphatase: does it have a role in predicting hepatocellular carcinoma recurrence? J Gastrointest Surg 2011;15(8):1440–9.
Leung TW, et al. Construction of the Chinese University Prognostic Index for hepatocellular carcinoma and comparison with the TNM staging system, the Okuda staging system, and the Cancer of the Liver Italian Program staging system: a study based on 926 patients. Cancer 2002;94(6):1760–9.

Cai X, et al. Albumin-to-alkaline phosphatase ratio as an independent prognostic factor for overall survival of advanced hepatocellular carcinoma patients without receiving standard anti-cancer therapies. J Cancer 2018;9(1):189–97.

Chan AW, et al. Albumin-to-alkaline phosphatase ratio: a novel prognostic index for hepatocellular carcinoma. Dis Markers 2015;2015:564057.

Nie M, et al. Albumin-to-alkaline phosphatase ratio: a novel prognostic index of overall survival in Cisplatin-based chemotherapy-treated patients with metastatic nasopharyngeal carcinoma. J Cancer 2017;8(5):809–15.

James MT, et al. Derivation and external validation of prediction models for advanced chronic kidney disease following acute kidney injury. JAMA 2017;318(18):1787–97.

Song L, et al. Assessment of the effect of treatment and assistance program on advanced patients with schistosomiasis japonica in China from 2009 to 2014. Parasitol Res 2016;115(11):4267–73.

Sinkala E, et al. Hepatosplenic schistosomiasis is characterised by high blood markers of translocation, inflammation and fibrosis. Liver Int 2016;36(1):145–50.

Wyszomirska RM, et al. High serum laminin and type IV collagen levels in schistosomiasis mansoni. Arch Gastroenterol 2005;42(4):221–5.

Li Y, et al. Severe hepatosplenic schistosomiasis: clinicopathologic study of 102 cases undergoing splenectomy. Hum Pathol 2011;42(1):111–9.