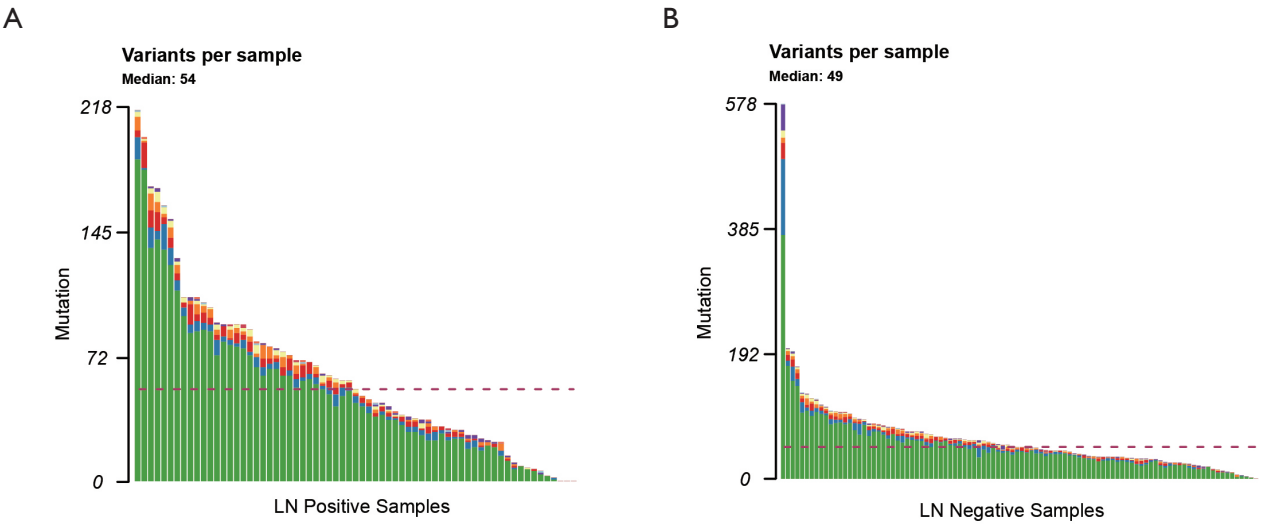

Figure S1 Mutation events in LN positive and LN negative patients. (A) The number of mutation events in each LN positive patient. Their median was 54. (B) The number of mutation events in each LN negative patient. Their median was 49.

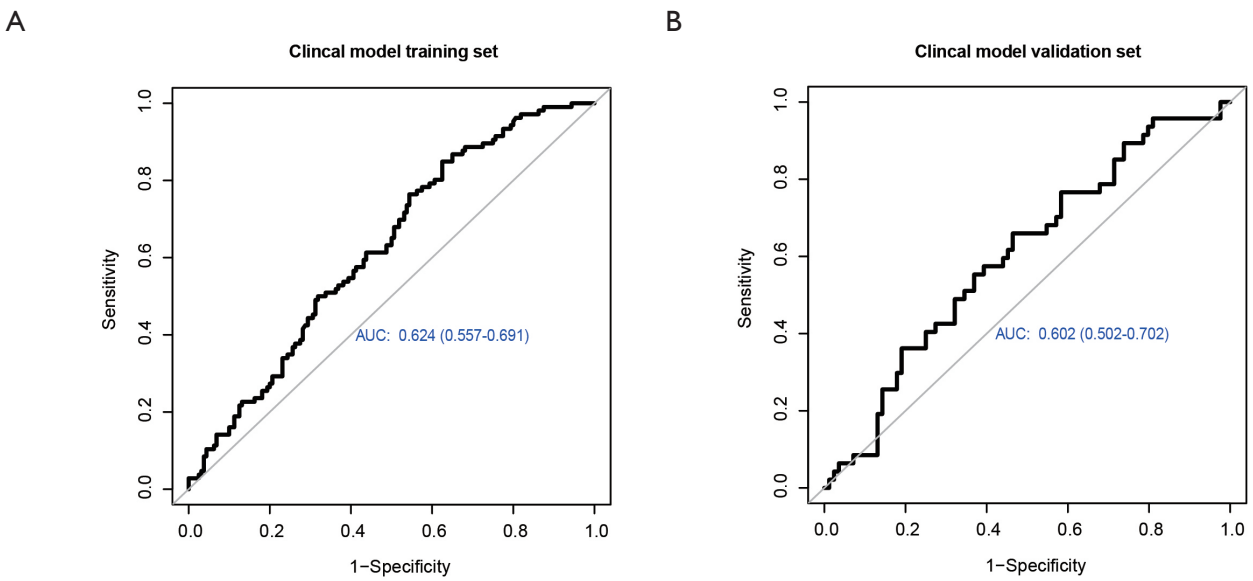

Figure S2 Details of the construction of the clinical model. (A) The area under the receiver operating characteristic (ROC) curve (AUC) of the clinical model in the training set. (B) The area under the curve (AUC) of the clinical model in the validation set.

https://dx.doi.org/https://dx.doi.org/10.21037/atm-22-277

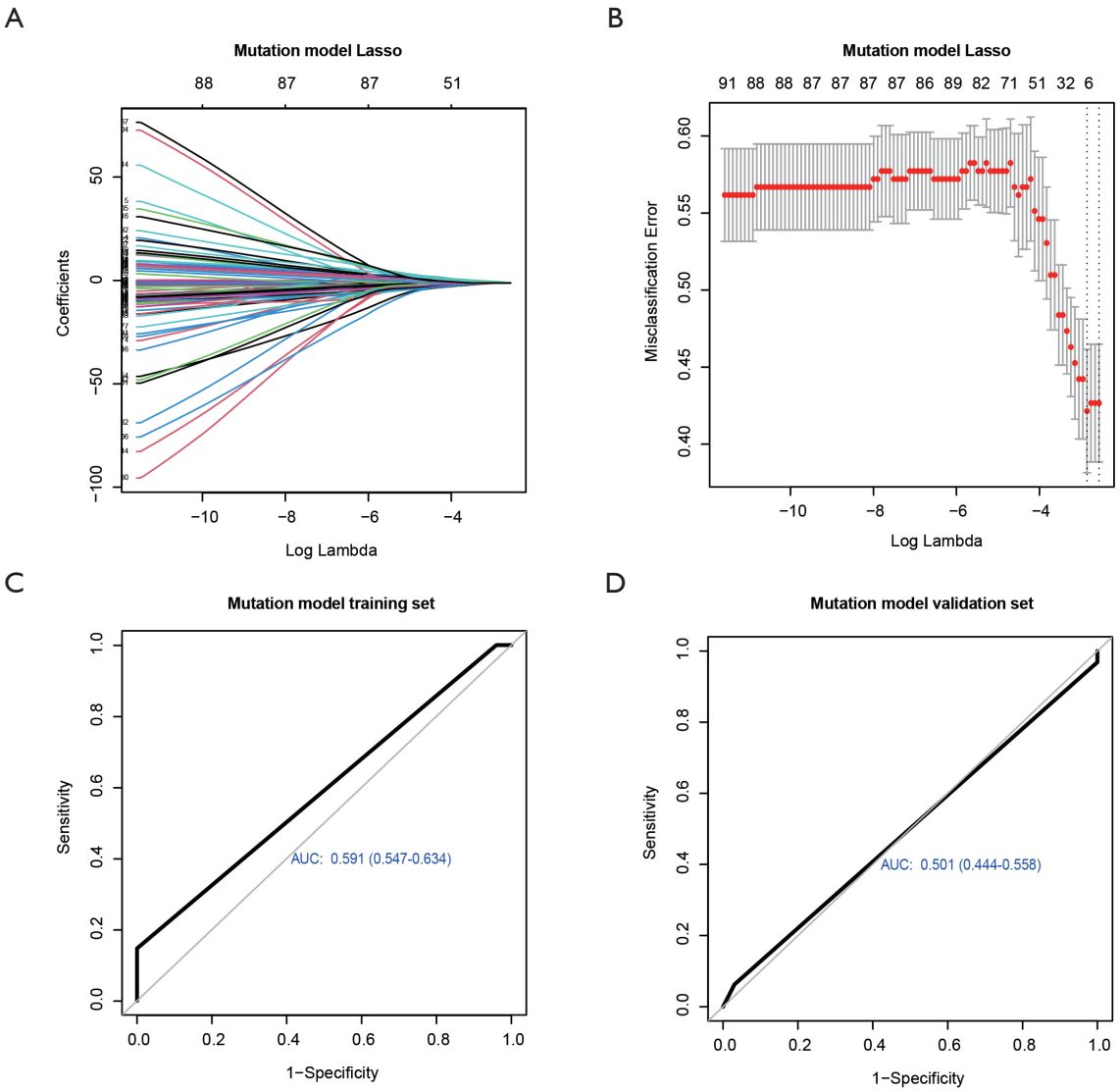

Figure S3 Details of the construction of the mutation model. (A) The least absolute shrinkage and selection operator (LASSO) coefficient profiles of all mutation signatures based on the training set. (B) LASSO algorithms were used to select optimal mutation signatures. (C) The area under the curve (AUC) of the mutation model in the training set. (D) The AUC of the mutation model in the validation set.

https://dx.doi.org/https://dx.doi.org/10.21037/atm-22-277

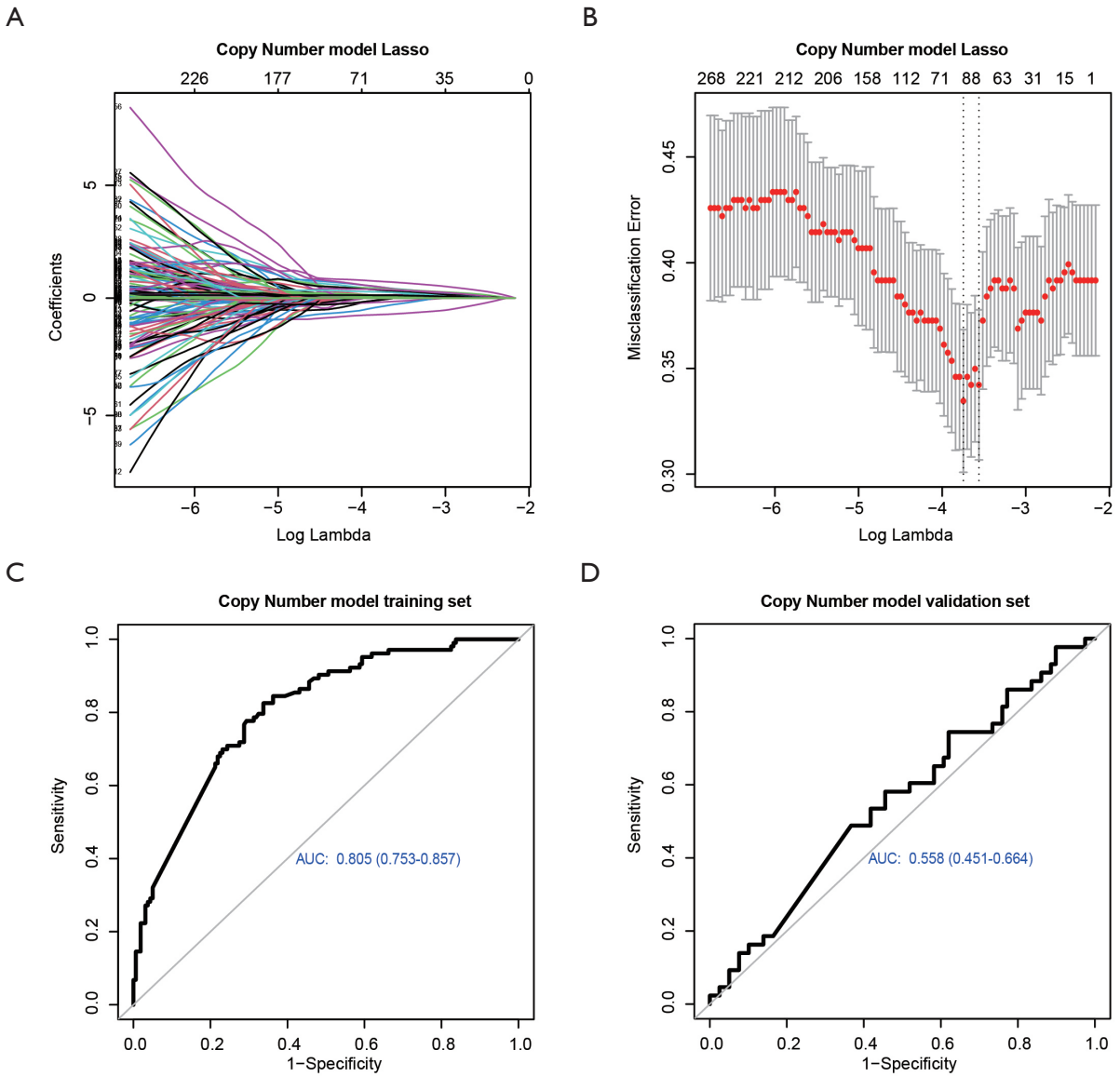

Figure S4 Details of the SCNA model construction. (A) The least absolute shrinkage and selection operator (LASSO) coefficient profiles of 1008 selected somatic copy number alterations (SCNAs) in genes based on the training set. (B) LASSO algorithms were used to select optimal SCNAs. (C) The AUC of the SCNA model in the training set. (D) The AUC of the SCNA model in the validation set.

https://dx.doi.org/https://dx.doi.org/10.21037/atm-22-277

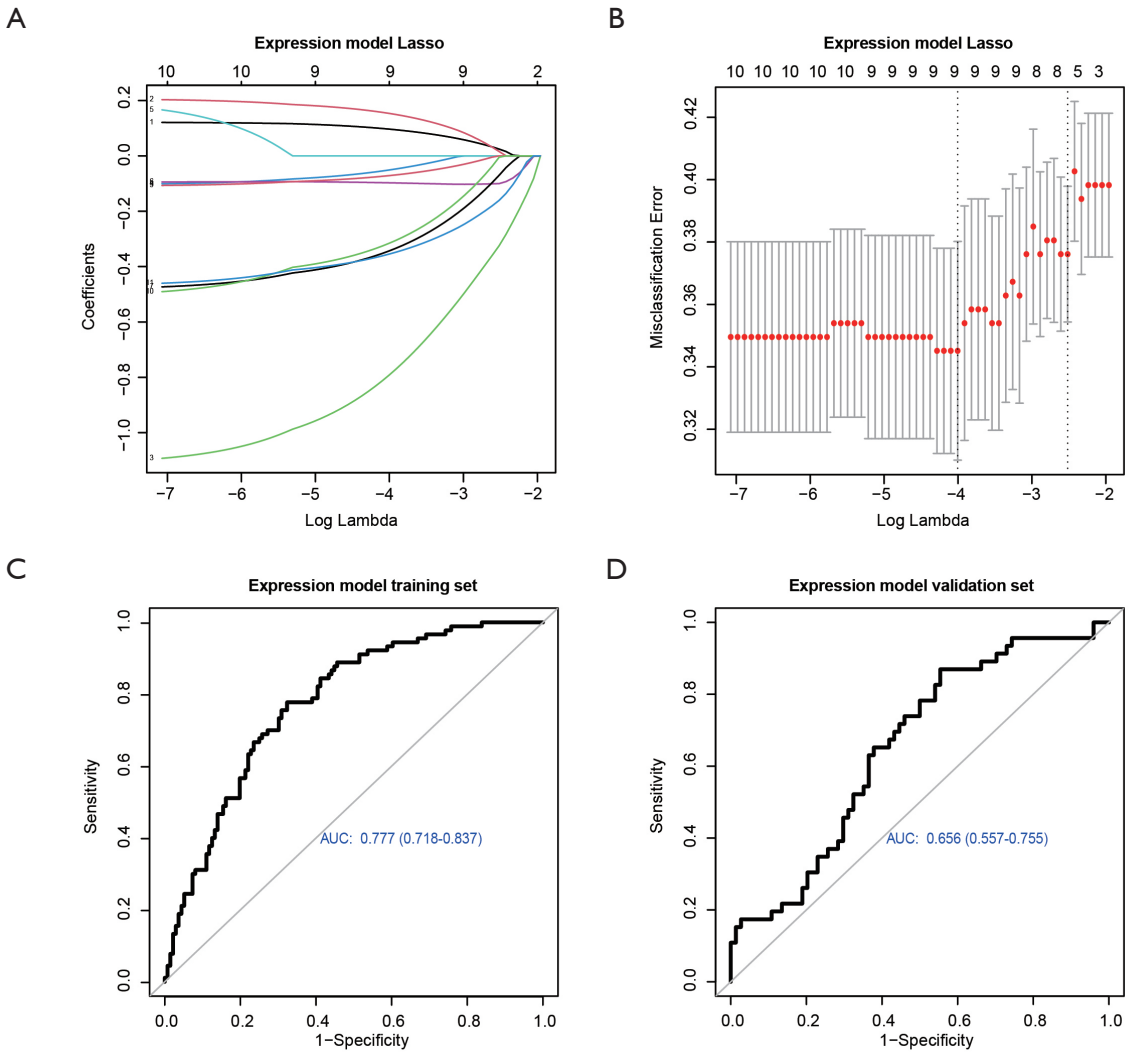

Figure S5 Details of the construction of the expression model. (A) The least absolute shrinkage and selection operator (LASSO) coefficient profiles of 11 selected expressions of genes based on the training set. (B) LASSO algorithms were used to select optimal gene expression levels. (C) The area under the curve (AUC) of the expression model in the training set. (D) The AUC of the expression model in the validation set.

https://dx.doi.org/https://dx.doi.org/10.21037/atm-22-277

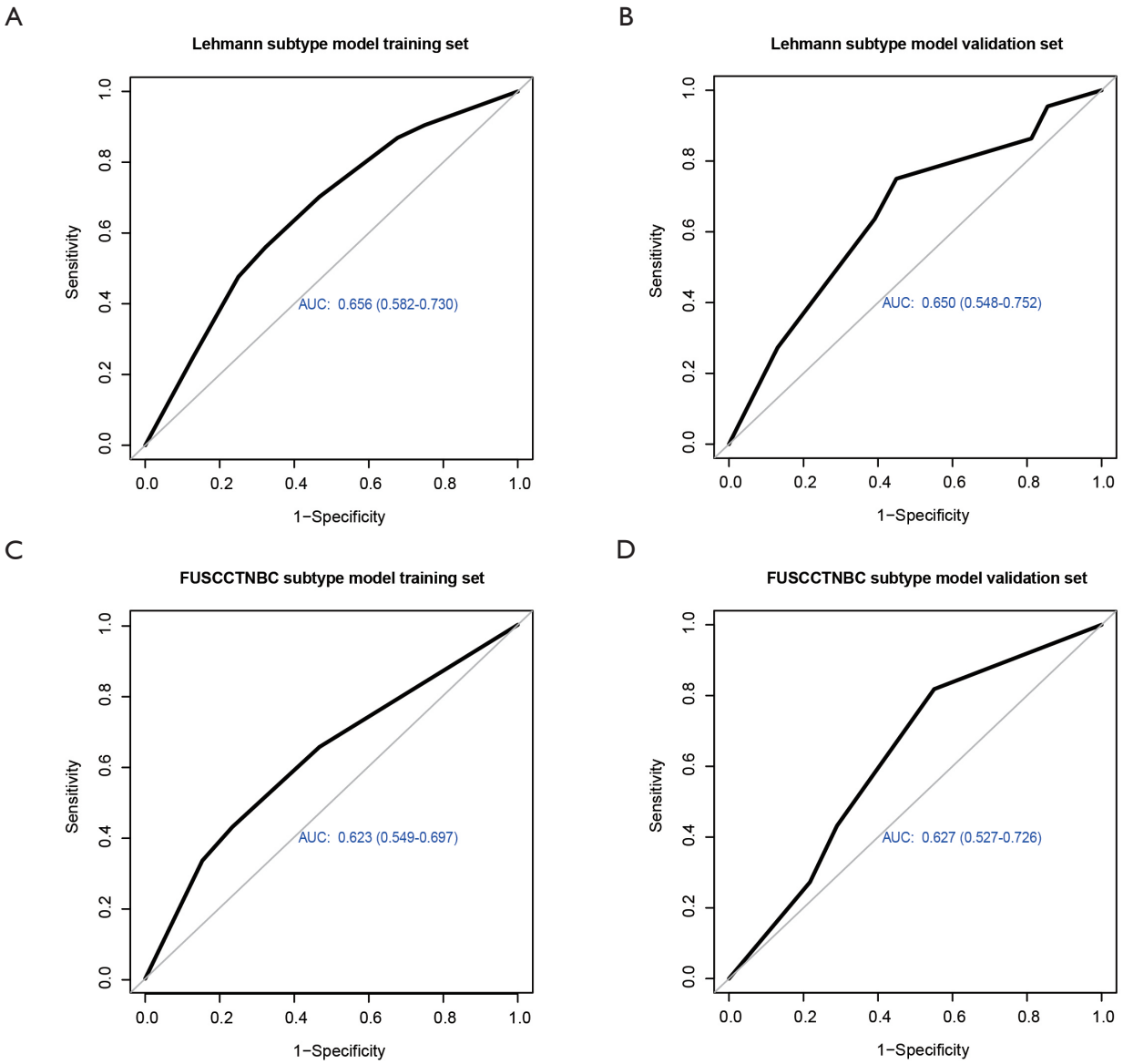

Figure S6 Details of the construction of the Lehmann and FUSCCTNBC subtype models. (A) The area under the curve (AUC) of the Lehmann subtype model in the training set. (B) The AUC of the Lehmann subtype model in the validation set. (C) The AUC of the Fudan University Shanghai Cancer Center Triple-negative Breast Cancer (FUSCCTNBC) subtype model in the training set. (D) The AUC of the FUSCCTNBC subtype model in the validation set.

https://dx.doi.org/https://dx.doi.org/10.21037/atm-22-277

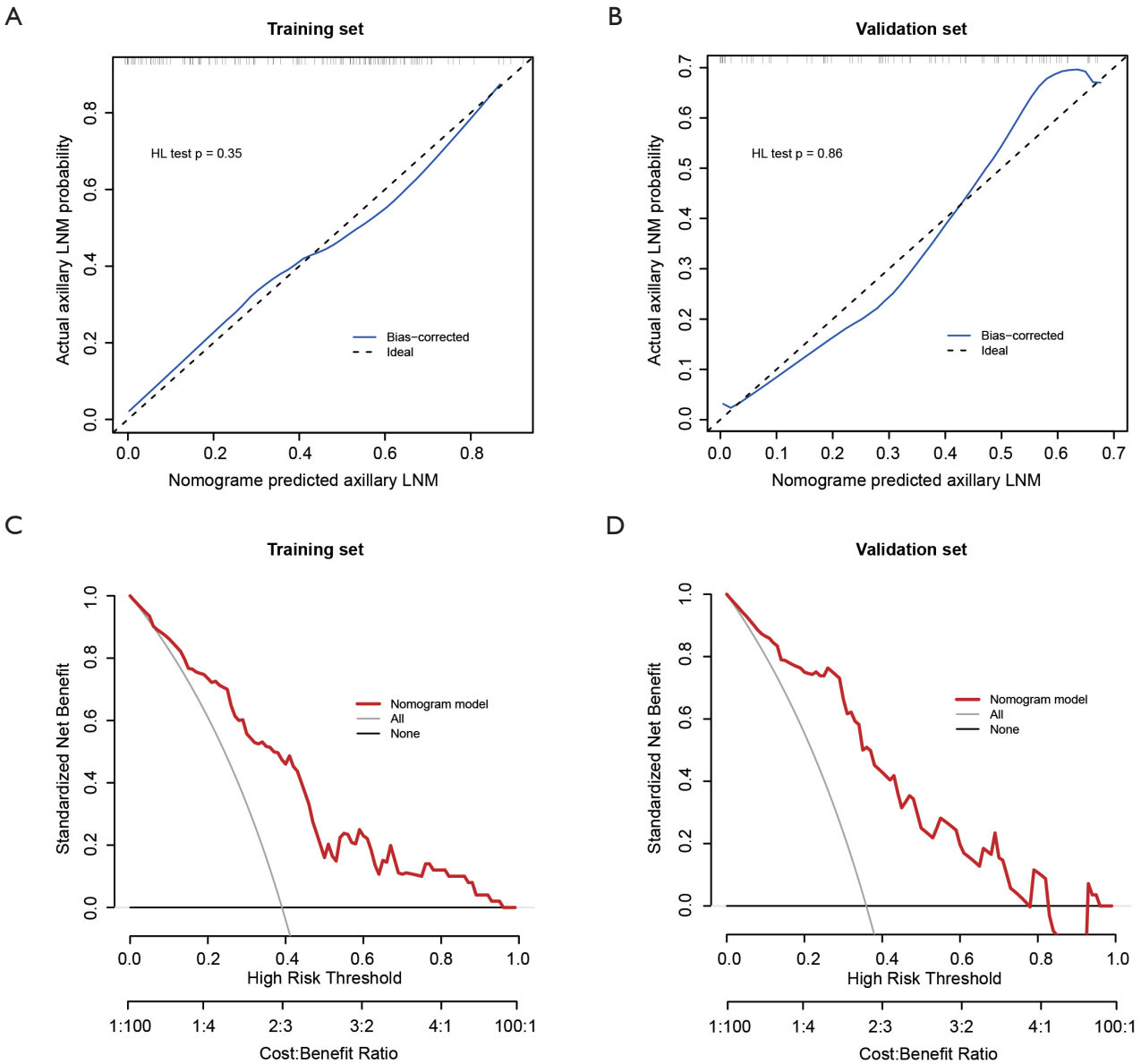

Figure S7 Evaluation of the multi-omics model. (A) The corresponding calibration curve of the multi-omics model in the training set. The 45-degree dotted black line represents a perfect prediction. The blue solid line represents the predictive performance of the multi-omics model in the training set. The Hosmer-Lemeshow (HL) test was used to compare the multi-omics model with its calibration in the training set. (B) The corresponding calibration curve of the multi-omics model in the validation set. The HL test was used to compare the multi-omics model with its calibration in the validation set. (C) The decision curve analysis (DCA) of the multi-omics model in the training set. The x-axis represents the threshold probability. The y-axis represents the standardized net benefit. The solid black line represents the net benefit when all patients are considered as not having axillary lymph node metastasis (LNM), while the gray line represents the net benefit when all patients are considered as having axillary LNM. The red line represents the net benefit when all patients are considered according to the multi-omics model in the training set. (D) The decision curve analysis (DCA) of the multi-omics model in the validation set.

https://dx.doi.org/https://dx.doi.org/10.21037/atm-22-277

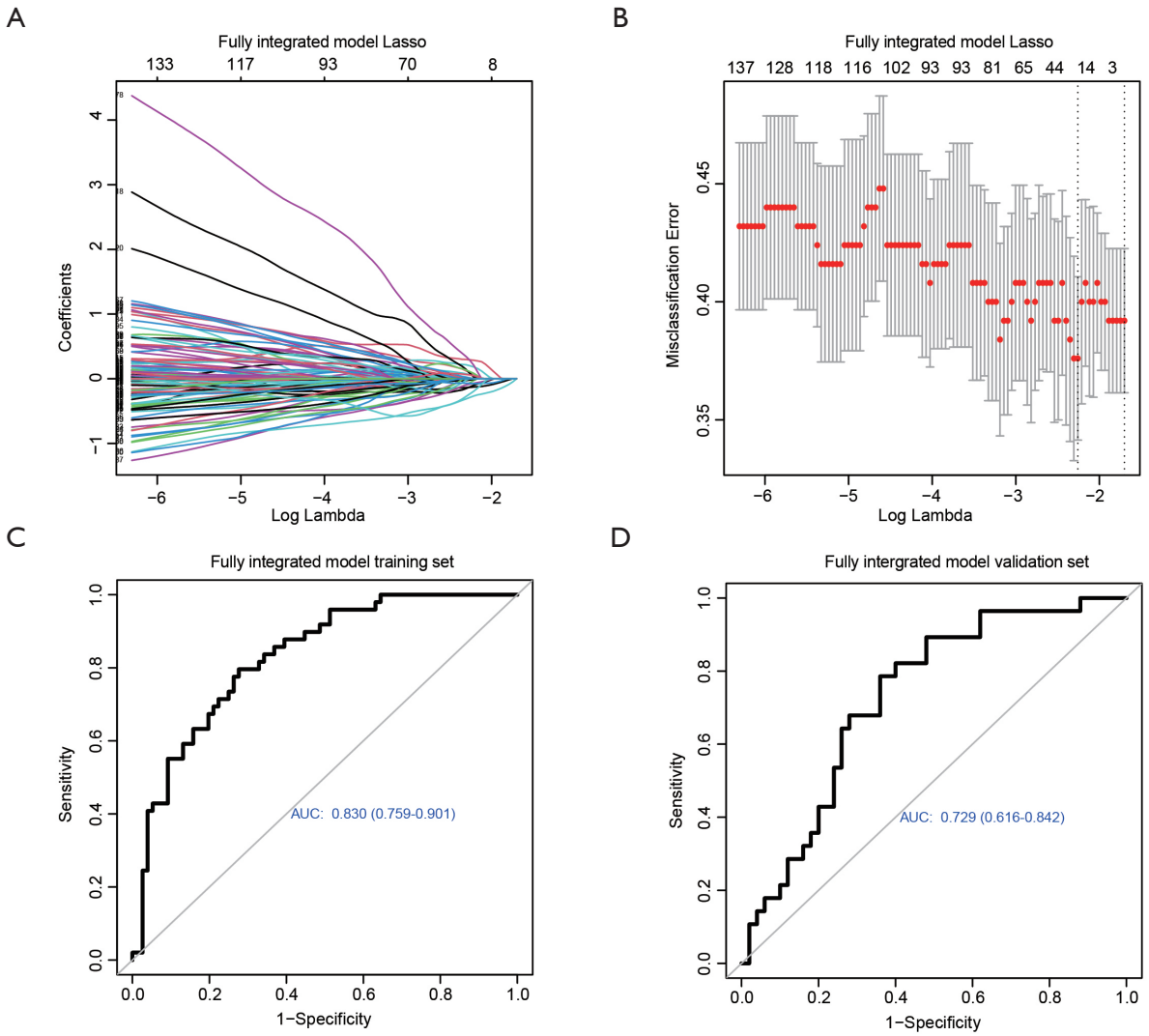

Figure S8 Details of the construction of fully integrated model. (A) The least absolute shrinkage and selection operator (LASSO) coefficient profiles of all factors in multi-omics based on the training set. (B) LASSO algorithms were used to select potential predictive factors. (C) The area under the curve (AUC) of the fully integrated model in the training set. (D) The AUC of the fully integrated model in the validation set.

https://dx.doi.org/https://dx.doi.org/10.21037/atm-22-277