**Statistical methods**

For the sizes of samples, the significance level was 0.05 and the power required was 90%. At present, there is no value of area under the receiver operating characteristic (ROC) curve (AUC) on rosacea prediction by this diagnostic method in the literature, so it is decided to use the poor situation to estimate, that is, the AUC value of rosacea prediction by this method is 0.650. The proportion of patients with rosacea and sensitive skin is expected to be 3:1. Based on the above assumptions, the sample size estimation shows that the number of subjects recruited in rosacea group is 145, and the number of subjects recruited in skin sensitive group is 48.

Continuous variables are expressed as the median (interquartile range), and categorical variables are expressed as the number and proportion, as appropriate. Inter-group comparisons of clinical characteristics were performed using a Wilcoxon's rank-sum test for continuous variables and chi-squared test for categorical variables. *P*-values <0.05 were considered statistically significant. For screening diagnostic factors in the training cohort, univariate logistic regression analysis was applied to examine the relationship between rosacea as a dependent variable and each clinical parameter and biomarker as an independent variable. Multivariable diagnostic models were constructed with multivariate logistic regression. We chose clinical parameters and biomarkers that had a *P*-value <0.05 on univariate logistic analysis. ROC curve analysis was performed to calculate the AUC to evaluate the diagnostic performance of the models. We computed the AUC with a 95% confidence interval using 1000 bootstrap re-sampling. Calculations of specificity, sensitivity, and likelihood ratios (LRs) were performed online (statis-ticscalculatorwww.vassarstats.net). A positive LR >1 indicated an association of the feature with rosacea. A positive LR of 1-5 showed a weak association. To evaluate the performance of the diagnostic models, decision curve analysis (DCA) was performed. This method is based on the principle that the relative harms of false positives and false negatives can be expressed in terms of a probability threshold. The net benefit is obtained by subtracting the proportion of patients who showed false-positive results from the proportion who showed true-positive results and then weighing the relative harm of false-positive and false-negative results. All statistical analyses were conducted using STATA 14.0. (Stata Corp., College Station, Texas, USA), including programs of DCA provided by Vickers.