The Turkish Journal of Pediatrics 2022; 64: 312-321
https://doi.org/10.24953/turkjped.2021.1644

Original Article

A feasibility study of risk prediction modelling for vaso-occlusive crisis in children with sickle cell disease

Merve Türkegün Şengül[1], Bahar Taşdelen[1], Selma Ünal[2], Veysi Akbey[3]

Departments of [1]Biostatistics and Medical Information and [2]Child Health and Diseases, Mersin University Faculty of Medicine, Mersin; [3]Mersin Su Hospital, Mersin, Turkey.

ABSTRACT

Background. The availability of a selection of biomarkers that includes information about disease risk is very important in the treatment of sickle cell disease (SCD). We used the predictiveness curve (PC), which classifies diseased individuals according to low- and high-risk thresholds, for this purpose. Our aim was to define this new statistical method and to determine the biomarkers that predict vaso-occlusive crisis (VOC) in children with SCD to guide preventive treatment.

Methods. Thirty-eight pediatric patients with SCD were included in this feasibility study. Leucocytes (WBC), C-reactive protein (CRP), interleukin-6 (IL-6), tumor necrosis factor (TNF-α), and YKL-40 were studied in patients with VOC and without VOC. The patient group with a low or high risk of VOC was assessed using the PC. Risk prediction and classification performance were evaluated using the PC and receiver operating characteristic (ROC) curve. Results. According to the PC, patients with a high risk of VOC could be detected via TNF-α, IL-6, and WBC, and TNF-α was the best risk prediction marker (TPF = 0.67).

Conclusions. The PC provides disease risk information by comparing more than one biomarker and can thereby help clinicians determine appropriate preventive treatments. This is the first study to evaluate biomarkers to predict VOC risk in SCD patients.

Key words: predictiveness curve, risk prediction, classifying, risk threshold, vaso-occlusion crisis.

In personalized medicine, clinicians can decide on the appropriate treatment for a patient based on the individual's demographic and genetic characteristics and/or biomarker measures so that the right treatment can be provided at the right time to the right patient.[1,2] Biomarker selection is currently the starting point for research on the diagnosis, identification, and treatment of many diseases.[3] Determining the biomarkers that provide the most beneficial information about disease risk is an important phase in the treatment process of complex diseases such as sickle cell disease (SCD).[4] Vaso-occlusive crisis (VOC) is the hallmark of SCD and is associated with various complications, including acute chest syndrome, multi-organ failure, and sudden death. VOC is both a cause of death and responsible for the majority of SCD patient hospitalizations.[5] It is important to understand the differences between VOC pain and other pain syndromes such as bone infarction, avascular necrosis, and leg ulcers in SCD patients so that the appropriate treatment can be determined.[6] No specific biomarker(s) are available for VOC diagnosis in SCD patients. This makes the treatment process challenging.[7] The prediction and classification of VOC risk would therefore be useful in treating these patients.

Biomarkers are used to classify or predict the risk according to the purpose of the research. Statistical processes are necessary to describe and select biomarkers.[8,9] The diagnostic classification performance of biomarkers is usually evaluated by a receiver operating

✉ Merve Türkegün Şengül
merveturkegun@gmail.com

Received 31st May 2021, revised 11th August 2021, accepted 5th October 2021.

characteristic (ROC) curve, sensitivity, and specificity criteria. These statistics can classify individuals as completely diseased or healthy; however, they do not provide information about future disease risk.[10] The risk of contracting a disease can be modelled as a function of the biomarkers using prediction models such as logistic or Cox regression.[11] Both methods are frequently used in biomarker selection in personalized medicine; however, individual differences should be considered when choosing a biomarker based on a diagnostic test because, main aim of personalized medicine is to determine which biomarker or biomarker combination can be used to separate patients into subgroups so that relevant treatment is offered to each patient.[1,2] The predictiveness curve (PC), which is a new prediction technique, offers an alternative to the classic statistical methods. PC visually presents population distribution of disease risk predicted by a continuous marker or risk model.[11]

In this study, we present a new graphic method, the PC, to calculate VOC risk. Our aim was to classify SCD patients based on their VOC risk. In doing so, we aimed to help clinicians develop patient-specific treatment plans and prevent the complications of this disease in the early stages so that treatment can be managed easily and fast.

Several candidate inflammatory biomarkers were identified in this feasibility study to predict the risk of VOC in a cohort of pediatric patients with SCD, namely, leucocytes (WBC) and C-reactive protein (CRP), which are routine follow-up markers, and the inflammation markers interleukin-6 (IL-6), tumor necrosis factor (TNF-α), and YKL-40 (also called chitinase-3-like-1)

Material and Methods

Predictiveness curve

The most common method for classifying patient risk is the ROC curve, which is obtained by pointing the false positive ratio (1–specificity) of the diagnostic test at different cut points against the true positive ratio (sensitivity). The minimum value for the sum of the false positive and false negative ratios is considered to be the appropriate threshold to indicate disease[12], but the obtained results only provide information only about disease status, not about disease risk. The PC plays an important role in the treatment management process for chronic diseases as it can be used to determine the population distribution of the predicted disease risk and the low-and high- risk groups.[13]

The PC models a biomarker's capacity of risk prediction and visually presents the distribution of the risk levels for the population from which the cohort is selected. It is used to determine disease risk for both the population and each individual and it illustrates absolute risk probabilities against the expected absolute disease risk values of individuals.[14] The most striking feature of the PC is the creation of a common scale that facilitates comparisons between biomarkers or risk models because the original scale is different and cannot be used for comparison purposes.[13]

The disease risk indicated by the biomarker (Y) is calculated using Equation 1, where D is the binary outcome variable related to disease (D=1 [present], D=0 [absent]) and Y is the biomarker value as the continuous variable:

$$Risk\ (Y) = P(D = 1 \mid Y = y) \qquad \text{Equation 1}$$

The PC is then drawn as the function of the disease risk versus cumulative risk percentage (v). The PC is a curve of R(v) versus to v and is indicated by Equation 2:[11]

$$R\ (v) = \mathrm{P}(D = 1 \mid Y = F^{-1}\ (v)) \qquad \text{Equation 2}$$

The inverse functions are easier to interpret than risk percentages. The inverse function $R^{-1}(p)$ of the risk percentage $R(v)$ gives a cumulative distribution of *risk* (Y) in the population, and it is the proportion of the population that has a risk less than or equal to p ($R^{-1}(p) = P[risk\ (Y) \leq p]$).[13]

To obtain the inverse functions, the low-and high-risk thresholds of the disease should be determined. These thresholds are detected by the trajectory of disease, the type of treatment, and the general prevalence of the disease. P_L and P_H represent the low- and high-risk thresholds, respectively. 1 - R^{-1} (p_H) is the ratio for a high-risk population, and R^{-1} (p_L) is the ratio for a low-risk population. This determines the percentage of the patient population that is below the low-risk threshold and above the high-risk threshold. The PC can be interpreted easily and quickly using the inverse functions. The percentage of the cohort between the low- and high-risk thresholds are unclear and are calculated using R^{-1} (p_H) − R^{-1} (p_L).[11,13,14] This risk class can be termed "the equivocal risk range."[15] We have defined this equivocal risk range as the "gray area".

Two or more risk models can be compared with the PC. When evaluating biomarkers, both the percentage of patients in the gray area and the true positive fraction (TPF) and false positive fraction (FPF) of the classification performance statistics are considered.[11,16]

Data collection

We designed a retrospective feasibility study and reviewed the medical records of hospitalized pediatric patients diagnosed with SCD who were admitted to the Department of Child Hematology, Mersin University Faculty of Medicine, Research Center Hospital, Mersin, between May 2016 and May 2017. The study group consisted of SCD patients between 3 and 17 years of ages. Patients who had been followed up regularly for a year, were not receiving erythrocyte transfusions, had not had VOC at last one month prior to the study, and did not have an inflammatory disease aside from SCD were included in the study.

The protocol was approved by the Mersin University Clinical Research Ethics Committee (protocol number of Mersin University-2018/258). All patient's parent was well informed about the protocols of study, and written informed consent was obtained.

The patients were divided into two groups: those who had had VOC for at least one year and the steady-state SCD patients who reported no VOC episodes for at least one year. The WBC, CRP, and inflammation parameters of IL-6, TNF-α, and serumYKL-40 were used to predict VOC risk in the steady-state patients. During the steady state period, 5 ml of blood samples was collected from the patients for a complete blood count and to measure the CRP, IL-6, TNF-α and serumYKL-40 levels. The blood samples were kept at −20°C until the study was complete.

Statistical methods

Univariate analysis

The Shapiro–Wilk test was used to check the normal distribution. We presented continuous variables as mean and standard deviation (mean ± SD). Frequencies (n) and percentages (%) were used for categorical data Student's *t*-test was used to compare age means between the steady-state and VOC patients. A chi-square test was performed to determine the relationships between the VOC and gender groups. McNemar's test was performed for the categorical data. Agreement between the new risk groups and the VOC groups was calculated using the Kappa statistic. Statistical significance was set at $p < 0.05$ for all the comparisons.

Multiple analysis and predictiveness curve

A multiple logistic regression (MLR) model was used to calculate age-and gender-adjusted risk thresholds for all the biomarkers. The risk predictions were recorded for each patient with the MLR analysis. The mean of the predicted values and the 95% confidence interval (95% CI) of the mean were then calculated. The lower and upper limits of the CI were defined as the low- and high- risk thresholds for the biomarkers, thereby establishing the risk groups. The low- and high-risk thresholds of VOC for a randomly selected patient were determined visually using the PC model. To select the biomarkers that could detect VOC risks in SCD patients, TPF ≥

0.60 and FPF ≤ 0.30 criteria and CIs not including 0 were considered statistically significant.

ROC analysis was conducted twice. First, we studied the performance of the biomarkers in discriminating between the VOC and steady-state patients. Second, we assessed the performance of the biomarkers in discriminating between the low- and high-risk groups. The area under the curve (AUC), accuracy, positive predictive value (PPV), negative predictive value (NPV) and 95% CIs of these values were then calculated.

Software

STATISTICA Version 13.5.0.17 (TIBCO Software Inc., Hillview Avenue Palo Alto, CA) was used for the univariate statistics. Stata/MP 11.0 (Stata Corp, College Station, TX) and its risk prediction package *predcurve* was used for PC analyses.[17,18]

Results

The diagnostic performances of biomarkers

In this study, the mean age of the patients was 12.6 ± 4.2 years. In the steady-state group, 55% of the patients were girls (n = 11), and 45% were boys (n = 9). In the VOC group, 27.8% of patients were girls (n = 5), and 72.2% were boys (n = 13). Both the gender (p = 0.090) and age (p = 0.941) distributions of the groups were similar (age means, 12.5 ± 4.2 years, 12.8 ± 4.3 years, respectively).

The AUC values of the WBC, IL-6, YKL-40 and TNF-α markers were calculated as less or equal to 0.50 (AUC= 0.434, 0.391, 0.535, 0.358, respectively). These biomarkers failed to distinguish the patients with VOC and those with steady-state disease. Although only CRP was greater than 0.50, it was not statistically significant (AUC=0.616, p=0.20).

Evaluation of the risk thresholds using prediction model

All the biomarkers were modeled together with adjusted age and gender to assess their risk prediction performance for VOC. Based on the MLR, the risk prediction model was:

(VOC risk) = 2.089 − 0.082 (WBC) + 0.130 (CRP) + 0.001 (YKL-40) − 0.008 (IL-6) − 0.028 (TNF-α) − 0.014 (age) − 1.43 (gender).

The risk predictions were recorded for each patient using MLR. The mean of the predicted values and the 95% CI of the mean were calculated. The mean of the risk predictions was calculated as 0.47 ± 0.28, and the 95% CI was 0.38 – 0.55. For VOC in SCD patients, the low-risk threshold was 0.38, and the high-risk threshold was 0.55. This meant that the patients whose risk estimation value was below 0.38 had a slight possibility of experiencing a VOC episode and those whose risk estimation value was above 0.55 had a higher possibility of experiencing a VOC episode. The contribution of biomarkers to risk estimation was evaluated visually by plotting the PC for each biomarker along with these thresholds. When the TNF-α, IL-6, and WBC markers were evaluated together with the covariate variables, we observed that they could be used to classify patients with a high risk of VOC. As shown in Table I, TNF-α had maximum TPF of 0.67 and minimum FPF of 0.20 Among the patients in the cohort, 31.6% were at or below the low-risk threshold. The patients with a VOC risk of 0.55 or greater comprised 39.5% of the cohort. TNF-α levels accurately predicted that 67% of the patients had a high risk of VOC (TPF = 0.67). 28.9% of the patients could not be classified into either the low- or high-risk groups, and allocated to the gray area based on their TNF-α levels (Table I).

The PC for TNF-α (Fig. 1) provided a more efficient predictive performance than the other biomarkers. The PC was interpreted in two ways without considering the risk thresholds. First, we evaluated the cumulative distributions using the x-axis. The results showed that 80% of the population had a VOC risk of 0.73 and lower, while 20% of the population had a VOC risk higher than 0.73. Second, we interpreted the cumulative distribution versus the risks on the y-axis. It could therefore be said that the

Table I. Summary statistics for the biomarkers according to the predictiveness curve.

	Low risk threshold (p_L)=0.38 (n=14)			High risk threshold (p_H)=0.55 (n=16)			GA (n=8)
	TPF (95% CI)	FPF (95% CI)	Risk percentile $R^{-1}(0.38)$	TPF (95% CI)	FPF (95% CI)	100-Risk percentile $1-R^{-1}(0.55)$	Risk percentile
TNF- α	0.89 (0.63-1.00)	0.55 (0.16-0.88)	31.60%	0.67 (0.23-0.91)	0.20 (0.04-0.50)	39.50%	*28.90%*
IL-6	0.71 (0.64-1.00)	0.60 (0.18-0.85)	34.20%	0.67 (0.00-0.93)	0.25 (0.00-0.58)	41.10%	23.70%
WBC	0.72 (0.53-1.100)	0.55 (0.21-1.00)	39.50%	0.66 (0.00-0.88)	0.30 (0.00-0.56)	44.70%	15.80%
CRP	0.78 (0.45-1.00)	0.60 (0.21-1.00)	34.20%	0.39 (0.07-0.91)	0.25 (0.00-0.57)	28.90%	36.90%
YKL-40	0.78 (0.50-1.00)	0.40 (0.20-1.00)	42.10%	0.72 (0.00-0.90)	0.45 (0.00-0.57)	55.30%	2.60%

WBC: leucocyte, TNF-α: tumor necrosis factor, IL-6: interleukin-6, CRP: C-reactive protein,YKL40:chitinase-3-like-1,pL: low-risk threshold, pH: high-risk threshold, $R^{-1}(0.38)$: the ratio for a low-risk population, $1-R^{-1}(0.55)$: the ratio for a high-risk population,GA: Grey Area; ($R^{-1}(0.55)- R^{-1}(0.38)$), TPF: true positive fraction, FPF: false positive fraction, 95% CI: 95% confidence interval for TPF and FPF.

TPF > 0.60 and FPF < 0.30 criteria and confidence intervals not including zero were considered statistically significant. For IL-6, TPF was 0.67 and FPF was 0.25 and 41.1 percent of patients were in the high-risk group or above for IL-6. Since TPF and FPF confidence intervals included zero, IL-6 was evaluated as not statistically significant.)

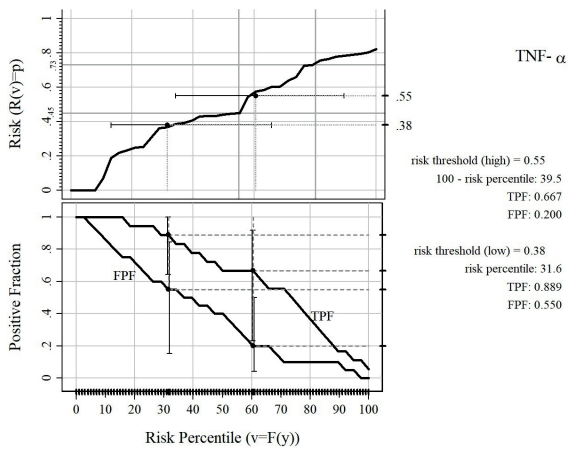

Fig. 1. Predictivenes curve for TNF- α

(TNF-α is predictive of low risk in $R^{-1}(0.38)$ = 31.6% of the population, of high risk in $1-R^{-1}(0.55)$ = 39.5% and it leaves 28.9% of patients in the gray area.)

patients whose VOC risk was 0.45 comprised 55% of the cohort.

For IL-6, the TPF was 0.67 and the FPF was 0.25 Accordingly, 41.1% of the patients were in the high-risk group or above. Since both CIs included 0, IL-6 was evaluated as not statistically significant. On the other hand, the other biomarkers were excluded from evaluation because of a low TPF, a high FPF, or a CI of 0 (Table I).

The biomarkers did not identify the low-risk patients as successfully as the high-risk patients. The TPF values were calculated as above 0.60 for all the biomarkers, while the FPF values were over 0.30 (Table I). Even though the low-risk threshold of the population was known, not all the biomarkers could accurately identify the patients at low risk of VOC.

Differences between the VOC and risk groups

According to the risk model, the risk predictions of 14 patients were below the low threshold, while the risk predictions of 16 patients were above the high threshold. Notwithstanding, eight patients' risk predictions were unclear (in the gray area), and these patients were not considered in the analysis. We constituted new risk groups based on the low- and high-risk thresholds using individual risks calculated by applying MLR.

The agreement between the new predicted risk groups and the observed VOC groups was investigated (Table II). No statistically significant differences were noted between the observed VOC and the predicted VOC risk groups when applying McNemar's test ($p > 0.05$) (Table II). This meant that the new prediction technique was agreement with the gold standard test results. The new technique can therefore be used for the prediction of VOC risk. The classification performance statistics were also calculated for these groups. The new risk groups distinguished the VOC and steady state patients with 70% accuracy (95% CI 50.60 - 85.26). The risk groups identified 62.5% of those at high risk of having a VOC episode (PPV = 62.5%) and 78.57% of those at low risk of having steady state (NPV = 78.57). Moderate agreement was also observed between the new risk groups and crisis groups' results (Kappa = 0.405 95% CI 0.086% – 0.725%).

The discrimination of power in the observed and predicted VOC groups was assessed for all the biomarkers using ROC analysis (Table III). The patients in the gray area were not included in this analysis. The discrimination power of TNF-α, IL-6, and WBC was uninterpretable for the VOC and steady-state patients in the observed VOC groups (AUC <0.50). CRP and YKL-40 had weak discrimination power and were not statistically significant. While TNF-α and IL-6 had excellent discrimination power in the low-and high-risk VOC groups, the discrimination power of WBC was acceptable in the predicted VOC risk groups.

Discussion

The most common reason of hospitalization among SCD patients is acute VOC episodes. Different risk factors trigger VOC episodes, which are sometimes unpredictable. VOC can cause a variety of complications, including acute chest syndrome, multi-organ failure and sudden death.[5] Treatments for VOC have focused on the symptomatic management of painful episode like to reduce pain, but it is currently not possible to prevent the painful episodes.[19] We therefore focused on developing

Table II. Differences between and observed and predicted groups.

Predicted VOC groups	Observed VOC groups		p
	VOC	Steady-state	
High risk (>0.55)	10 (76.9)	6 (35.3)	
Low risk (<0.38)	3 (23.1)	11 (64.7)	0.50
Total	13 (100)	17 (100)	

VOC: vaso-occlusion crisis. Data are shown as n (%) and were compared by McNemar test. PPV: positive predictive value, NPV: negative predictive value. 95% CI: 95% confidence interval for PPV and NPV. PPV is 62.50% (95% CI (45.05%-77.20%]) NPV is 78.57% (95% CI (56.13%-91.31%]). Accuracy is 70% (95% CI (50.60%-85.26%]))

Table III. ROC analysis results for observed and predicted groups for biomarkers.

	Observed crisis groups			Predicted risk groups		
	Crisis (n=13) vs steady state (n=17)			High risk (n=16) vs Low risk(n=14)		
	AUC (95%CI)	p	Power	AUC (95%CI)	p	Power
TNF-α	0.335 (0.138-0.531)	0.32	0.13	0.835 (0.685-0.985)	<0.001	0.94
IL-6	0.385 (0.179-0.591)	0.18	0.39	0.929 (0.773-0.990)	<0.001	0.99
WBC	0.439 (0.228-0.650)	0.08	0.57	0.710 (0.520-0.900)	0.03	0.53
CRP	0.683 (0.490-0.876)	0.38	0.09	0.650 (0.450-0.850)	0.14	0.30
YKL-40	0.606 (0.389-0.824)	0.16	0.32	0.585 (0.377-0.793)	0.42	0.11

AUC: area under the curve, 95% CI: 95% confidence interval for AUC, Power: statistical power for AUC. If risk is ≥ 0.55, it is high risk. If risk is ≤0.38, it is low risk.

a useable PC method to help clinicians prevent the complications of SCD in the early stages using predictive risk biomarkers and risk groups. Our study results are not generalizable to the entire SCD population because our sample size was small. However, we used our data to introduce a new prediction technique and obtained reasonable results. Nevertheless, risk models like PC should be calculated using a large cohort that is representative of the target population.[11] The use of a large cohort could provide greater confidence in the results and be generalizable. Patients could then be classified with high accuracy, and percentage of grey area patients would decrease. We aimed to demonstrate that this new prediction technique could be applicable for use in patients with this important chronic disease However, we did not consider the patients' genotypes and other important clinical factors. The simultaneous evaluation of genotypes and such clinical factors in a large sample size may have a considerable effect on the results of future studies.

The hematologic changes that appear in the steady state or during the VOC are very important for SCD treatment.[20] WBC, CRP, IL-6, and TNF-α levels are known to increase during VOC episodes in SCD patients.21 Although many biomarkers have been identified to assist in the diagnosis of SCD, none is a reliable and certain indicator of VOC risk.[7] Although high LDH, PCT, and WBC levels are statistically associated with VOC risk on admission, WBC and PCT are not used as risk markers for VOC. When red blood cell transfusion or exchange is considered during VOC episode but the clinical indication is unclear, LDH level may help in making a decision. But this result for LDH does not guarantee predict VOC risk.[22,23] Studies have shown a strong correlation between CRP and WBC levels and hospitalization with VOC episodes.[24,25] These studies have suggested that WBC and CRP levels could be used as predictive biomarkers for VOC by applying a univariate statistical analysis to compare the means of the biomarkers between groups or a correlation analysis. We aimed to determine the best biomarkers to predict VOC risk. Accordingly, we classified patients at risk of VOC based on their biomarker values using the new statistical technique, PC.

ROC analysis is insufficient for risk prediction, because ROC analysis only distinguishes people as diseased or healthy.[26,27] The proposed PC indicates the capacity of a biomarker to predict disease risk. It also provides information about the performance of these predictive models while estimating risk in the population.[13] The PC can further be used to display the population distribution of disease risk as predicted by the chosen biomarker.[15] The PC provides important information about risk, which cannot be shown using the ROC curve, by classifying patients who have low- and high disease risk based on the biomarker values. By applying thresholds to classify patients as high or low risk, the value of a continuous biomarker can be evaluated by estimating the PPV and the NPV.[28]

An acceptable low- and high-risk threshold or risk intervals should be determined in the research, development, and selection of studies for biomarkers.[11] Risk thresholds depend on the context of the disease and include weighing the expected costs against benefits associated with a high-risk findding.[13] However, alternative methods can be used to determine risk thresholds because such analyses are difficult and time-consuming.[27] In our study, we calculated the low- and high-risk thresholds using a MLR risk model. By applying the PC model and risk thresholds, we were able to obtain the VOC risks for children with SCD.

The probability of VOC is estimated according to the threshold value of a biomarker calculated using the PC and can thereby help to start, continue, or complete treatment for a clinical study.[29,30] A perfect marker is indicated if the biomarker has a TPF = 1 and a FPF = 0. Even if the TPF and the FPF are not equal to 1 and 0, respectively, or are less, the marker can still be beneficial. The criteria to decide which biomarker will be useful depends on the study purpose. The general approach is that the TPF

and the FPF should be balanced.[16] Biomarkers can be classified as fitting the optimal balance of TPF and FPF if the *risk value > risk threshold* rule, which is often used as a clinical decision criterion.[11] In our study, we considered the clinical status of the patients, calculated their risk thresholds, and then used this status to determine the biomarker selection criteria; TPF ≥ 0.60 was the ratio of detecting patients with VOC, while a FPF ≤ 0.30 was the ratio indicated to show the steady-state patients with VOC. Our results showed that all the biomarkers were insufficient to classify low-risk patients.

If the risk of the disease is close to 0 or 1, the drug dose, variety of treatment, and duration of treatment can easily be determined. However, if the risk of disease is in the gray area, the decision becomes more difficult.[11] If a biomarker assigns a large number of individuals to the low- and high-risk groups and a smaller number of individuals to the gray area, it is accepted that the risk biomarker is strong or best.[14,15] In our study, five patients with VOC and three steady-state patients were in the gray area. In such instances, the question of whether such individuals should be evaluated as low- or high-risk patients is important. The status of patients who cannot not be assigned to any risk class and remain in the gray area is an important topic as this could lead to these patients being over- or under-treated. For patients in the gray area, genetic and environmental risk factors and compliance with treatment should therefore be investigated retrospectively.

According to the PC results, TNF-α was the best biomarker for classifying patients at a high risk of VOC (Table I). TNF-α, IL-6, and WBC were found to discriminate between the low- and high-risk groups when the PC risk model was used with the ROC analysis (Table III). The PC and ROC analysis results were consistent. On the other hand, the classic statistic results showed that classifications based on the performance of the biomarkers for steady-state patients and those with VOC were not good. AUC values were above 0.60 for all biomarkers. These results mean that making comparisons according to risk categories using the PC method can provide more accurate results in critical diseases like SCD.

It is a feasibility study. So, we evaluated statistical power of study in the table III. The statistical power for AUCs was lower than 0.80 and an unacceptable level for the observed groups. Once the low- and high- risk groups had been predicted, AUC values and the statistical power increased for TNF-α, IL-6, and WBC. Notably, the statistical powers were higher than 0.80 for the TNF-α and IL-6 (Table III). Although our sample size was small, agreement was established between the predicted and observed results.

The findings of our study showed that classifying patients by their VOC risk group and managing their treatment accordingly are extremely important. Group comparisons of biomarkers should be done in line with the low and high VOC risk groups in SCD patients. Attention should be paid to the treatment of patients whose VOC risk is unclear.

Ethical approval

The protocol was approved by the Mersin University Clinical Research Ethics Committee by protocol number of MEU-2018/258.

Author contribution

The authors confirm contribution to the paper as follows: study conception and design: BT, MTŞ; data collection: SÜ, VA; analysis and interpretation of results: MTŞ, BT; draft manuscript preparation: MTŞ. All authors reviewed the results and approved the final version of the manuscript.

Source of funding

The authors declare the study received no funding.

Conflict of interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Zhao Y, Zeng D. Recent development on statistical methods for personalized medicine discovery. Front Med 2013; 7: 102-110. https://doi.org/10.1007/s11684-013-0245-7
2. Li W, Chen C, Li X, Beckman RA. Estimation of treatment effect in two-stage confirmatory oncology trials of personalized medicines. Stat Med 2017; 36: 1843-1861. https://doi.org/10.1002/sim.7272
3. Antoniou M, Jorgensen AL, Kolamunnage-Dona R. Biomarker-Guided adaptive trial designs in phase II and phase III: A methodological review. PLoS One 2016; 11: 1-30. https://doi.org/10.1371/journal.pone.0149803
4. Piel FB, Patil AP, Howes RE, et al. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. Lancet 2013; 381: 142-151. https://doi.org/10.1016/S0140-6736(12)61229-X
5. Ballas SK, Lusardi M. Hospital readmission for adult acute sickle cell painful episodes: frequency, etiology, and prognostic significance. Am J Hematol 2005; 79: 17-25. https://doi.org/10.1002/ajh.20336
6. Darbari DS, Sheehan VA, Ballas SK. The vaso-occlusive pain crisis in sickle cell disease: Definition, pathophysiology, and management. Eur J Haematol 2020; 105: 237-246. https://doi.org/10.1111/ejh.13430
7. Tumblin A, Tailor A, Hoehn GT, et al. Apolipoprotein A-I and serum amyloid a plasma levels are biomarkers of acute painful episodes in patients with sickle cell disease. Haematologica 2010; 95: 1467-1472. https://doi.org/10.3324/haematol.2009.018044
8. Lipkovich I, Dmitrienko A. Strategies for identifying predictive biomarkers and subgroups with enhanced treatment effect in clinical trials using SIDES. J Biopharm Stat 2014; 24: 130-153. https://doi.org/10.1080/10543406.2013.856024
9. Gosho M, Nagashima K, Sato Y. Study designs and statistical analyses for biomarker research. Sensors (Basel) 2012; 12: 8966-8986. https://doi.org/10.3390/s120708966
10. Huang Y, Pepe MS. Semiparametric methods for evaluating the covariate-specific predictiveness of continuous markers in matched case-control studies. J R Stat Soc Ser C Appl Stat 2010; 59: 437-456. https://doi.org/10.1111/j.1467-9876.2009.00707.x
11. Pepe MS, Feng Z, Huang Y, et al. Integrating the predictiveness of a marker with its performance as a classifier. Am J Epidemiol 2008; 167: 362-368. https://doi.org/10.1093/aje/kwm305
12. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29-36. https://doi.org/10.1148/radiology.143.1.7063747
13. Huang Y, Pepe MS, Feng Z. Evaluating the predictiveness of a continuous marker. Biometrics 2007; 63: 1181-1188. https://doi.org/10.1111/j.1541-0420.2007.00814.x
14. Huang Y, Pepe MS. Assessing risk prediction models in case-control studies using semiparametric and nonparametric methods. Stat Med 2010; 29: 1391-1410. https://doi.org/10.1002/sim.3876
15. Huang Y, Pepe MS. A parametric ROC model-based approach for evaluating the predictiveness of continuous markers in case-control studies. Biometrics 2009; 65: 1133-1144. https://doi.org/10.1111/j.1541-0420.2009.01201.x
16. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. Am J Epidemiol 2004; 159: 882-890. https://doi.org/10.1093/aje/kwh101
17. STATISTICA (data analysis software system). Version 13.5.0.17. Hillview Avenue, CA; USA: TIBCO Software Inc; 2017.
18. Stata/MP (Statistical Software) Release 11. College Station, Texas; USA: StataCorp LP; 2009.
19. Osunkwo I, Manwani D, Kanter J. Current and novel therapies for the prevention of vaso-occlusive crisis in sickle cell disease. The Adv Hematol 2020; 11: 1-15. https://doi.org/10.1177/2040620720955000
20. Zempsky WT. Evaluation and treatment of sickle cell pain in the emergency department: paths to a better future. Clin Pediatr Emerg Med 2010; 11: 265-273. https://doi.org/10.1016/j.cpem.2010.09.002
21. Rees DC, Gibson JS. Biomarkers in sickle cell disease. Br J Haematol 2011; 156: 433-445. https://doi.org/10.1111/j.1365-2141.2011.08961.x
22. Stojanovic KS, Steichen O, Lefevre G, et al. High lactate dehydrogenase levels at admission for painful vaso-occlusive crisis is associated with severe outcome in adult SCD patients. Clin Biochem 2012; 45: 1578-1582. https://doi.org/10.1016/j.clinbiochem.2012.07.114
23. Onyekwelu KC, Ufelle SA, Ikekpeazu JE, Ezeh RC, Uche CZ, Udoh IP. Changes during vaso-occlusive crisis (VOC) and normal state in sickle cell disease patients. Med Case Rep Rev 2019; 2: 1-4. https://doi.org/10.15761/MCRR.1000122

24. Krishnan S, Setty Y, Betal SG, et al. Increased levels of the inflammatory biomarker C-reactive protein at baseline are associated with childhood sickle cell vasocclusive crises. Br J Haematol 2010; 148: 797-804. https://doi.org/10.1111/j.1365-2141.2009.08013.x

25. Al-Basheer JG, Abdelgadir O, Mustafa AEM, Muddathir ARM, Abdelgadir RE. C - reactive protein level and WBC count as biomarkers for vaso-occlusive crisis among patients with sickle cell disease. American Journal of Medicine and Medical Sciences 2015; 5: 283-286.

26. Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. Biostatistics 2005; 6: 227-239. https://doi.org/10.1093/biostatistics/kxi005

27. Pepe MS, Feng Z, Gu JW. Comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by M. J. Pencina et al., Statistics in Medicine. Stat Med 2008; 27: 173-181. https://doi.org/10.1002/sim.2991

28. Sachs MC, Zhou XH. Partial summary measures of the predictiveness curve. Biom J 2013; 55: 589-602. https://doi.org/10.1002/bimj.201200146

29. Empereur-mot C, Guillemain H, Latouche A, Zagury JF, Viallon V, Montes M. Predictiveness curves in virtual screening. J Cheminform 2015; 7: 52. https://doi.org/10.1186/s13321-015-0100-8

30. Janes H, Pepe MS, Bossuyt PM, Barlow WE. Measuring the performance of markers for guiding treatment decisions. Ann Intern Med 2011; 154: 253-259. https://doi.org/10.7326/0003-4819-154-4-201102150-00006