Received: 28 October 2019 | Revised: 11 December 2019 | Accepted: 12 December 2019

DOI: 10.1002/jcla.23189

RESEARCH ARTICLE

WILEY

Novel classifiers with clinical laboratory parameters for early detection of osteosarcoma

Lin-Lin Cao[1] | Zhaoming Chen[2] | Zhihong Yue[1] | Lin Pei[1] | Mei Jia[1] | Hui Wang[1] | Tingting Li[2]

[1]Department of Clinical Laboratory, Peking University People's Hospital, Beijing, China

[2]Department of Biomedical Informatics, School of Basic Medical Sciences, Peking University Health Science Center, Beijing, China

Correspondence
Tingting Li, Department of Biomedical Informatics, School of Basic Medical Sciences, Peking University Health Science Center, #38 Xueyuan Road, Beijing 100191, China.
Email: litt@hsc.pku.edu.cn

Lin-Lin Cao, Department of Clinical Laboratory, Peking University People's Hospital, #11 Xizhimen South Street, Beijing 100044, China.
Email: caollpku@163.com

Funding information
National Key Research and Development Program of China, Grant/Award Number: 2018YFA0507504; National Natural Science Foundation of China, Grant/Award Number: 61773025 and 81702788

Abstract

Background: Osteosarcoma (OS) is one of the most common malignant bone tumors. It is essential to explore early diagnostic indicators with high sensitivity and specificity due to the rapid progression and metastasis of OS and the poor survival of metastatic OS patients. However, a few indicators of diagnostic significance have been described.

Methods: A total of 458 OS patients, 312 healthy individuals, and 228 patients with primary benign bone lesions were included. Logistic regression was performed on 46 clinical laboratory parameters to establish the diagnostic classifiers, which were evaluated by analysis of the receiver operating characteristic (ROC) curves.

Results: We established three diagnostic classifiers, called C_{os} for all ages, C_{los} for low ages, and C_{hos} for high ages, with clinical laboratory parameters to distinguish OS from healthy individuals. All classifiers showed better diagnostic performances than alkaline phosphatase (ALP) in the independent validation cohort. In addition, these classifiers had better ability than ALP to discriminate OS from primary benign bone lesions. Furthermore, C_{os}, $C_{los,}$ and C_{hos} had larger AUC than ALP to identify small-size and early-stage OS and could also detect ALP-negative OS effectively.

Conclusion: Our study suggests the potential of C_{os}, C_{los}, and C_{hos} as non-invasive biomarkers for early OS.

KEYWORDS

alkaline phosphatase, classifier, diagnosis, laboratory parameter, osteosarcoma

1 | INTRODUCTION

Osteosarcoma (OS), which arises from primitive bone-forming mesenchymal cells, is one of the most prevalent bone malignancies. It affects patients of all ages, particularly children and adolescents, accounting for approximately 56% of all pediatric bone tumors.[1,2] OS develops most commonly in areas where the bone is growing rapidly, such as the distal femur and proximal tibia. The typical presenting symptom is the onset of pain and swelling in the affected bone, and occasionally, patients present with pathologic fracture.[3] Approximately 20% of

Cao and Chen equally contributed to this paper.

OS patients have clinically detectable metastases at diagnosis, which occurs most commonly in the lung.[4] With the introduction of chemotherapy in the 1970s, substantial improvement in long-term survival has been achieved.[5] However, the prognosis of some OS patients is still poor due to the rapid progression and metastasis, and patients with metastatic disease showed much lower event-free survival rate than that those with localized disease.[3] Therefore, early diagnosis of OS is critical for improving the therapeutic effect of patients.

Tumor markers have been widely used in clinical management for many cancers. They are of great value in many clinical aspects, such as screening early malignancy, diagnosis, therapeutic effect monitoring, and prognosis evaluation.[6] Regarding OS, serum alkaline phosphatase (ALP) is probably one of the most well-known tumor markers. Numerous studies have shown that pre-treatment serum ALP, as well as its changes during therapeutic process, is a critical prognostic indicator for chemotherapy response, skeletal and lung metastases, and survival of OS patients.[7-10] In addition, it has been known for a long time that OS patients often have higher serum ALP than normal subjects and patients with benign bone lesions, and ALP might help to diagnose OS initially. However, there were still a large part (>40%) of OS patients with normal ALP levels, resulting in very low sensitivity for ALP in OS diagnosis.[10]

Some other clinical parameters have been identified and reported as potential tumor markers in OS besides ALP. For example, serum lactate dehydrogenase (LDH),[8,11] serum C-reactive protein,[12] lymphocyte-to-monocyte ratio (LMR),[13] neutrophil-to-lymphocyte ratio (NLR), and platelet-to-lymphocyte ratio (PLR)[14] were also described as predictors of clinical prognosis in OS patients. However, the diagnostic performances of these clinical parameters to identify OS are largely unknown. Although some new indicators have been found to be elevated in OS patients,[15,16] their role in clinical diagnosis still needs further evaluation. Therefore, there is still no highly sensitive and specific markers to detect OS at the early stage at present.

TABLE 1 Clinicopathological characteristics of included OS patients

	Training cohort		Validation cohort	
Clinical parameters	**≤15 y**	**>15 y**	**≤15 y**	**>15 y**
Gender				
Male	98	108	41	46
Female	67	49	28	21
Age (Mean ± SD)	11.00 ± 3.02	25.36 ± 11.97	11.14 ± 2.88	25.78 ± 12.47
Pathological fracture				
Yes	15	9	10	9
No	150	148	59	58
Clinical stage				
I-II	121	127	51	49
III	44	30	18	18
Pulmonary metastasis				
Yes	40	22	14	17
No	125	135	55	50
Tumor size				
≤6 cm	83	93	34	33
>6 cm	58	49	22	24
NA	24	15	13	10
Tumor location				
Tibia/femur	125	101	52	43
Elsewhere	40	56	17	24
Preoperative chemotherapy				
Yes	163	114	65	55
No	2	43	4	12
Huvos grade				
I-II	94	67	31	34
III-IV	48	39	25	12
NA	23	51	13	21

Abbreviations: OS, osteosarcoma; NA, not available; SD, standard deviation.

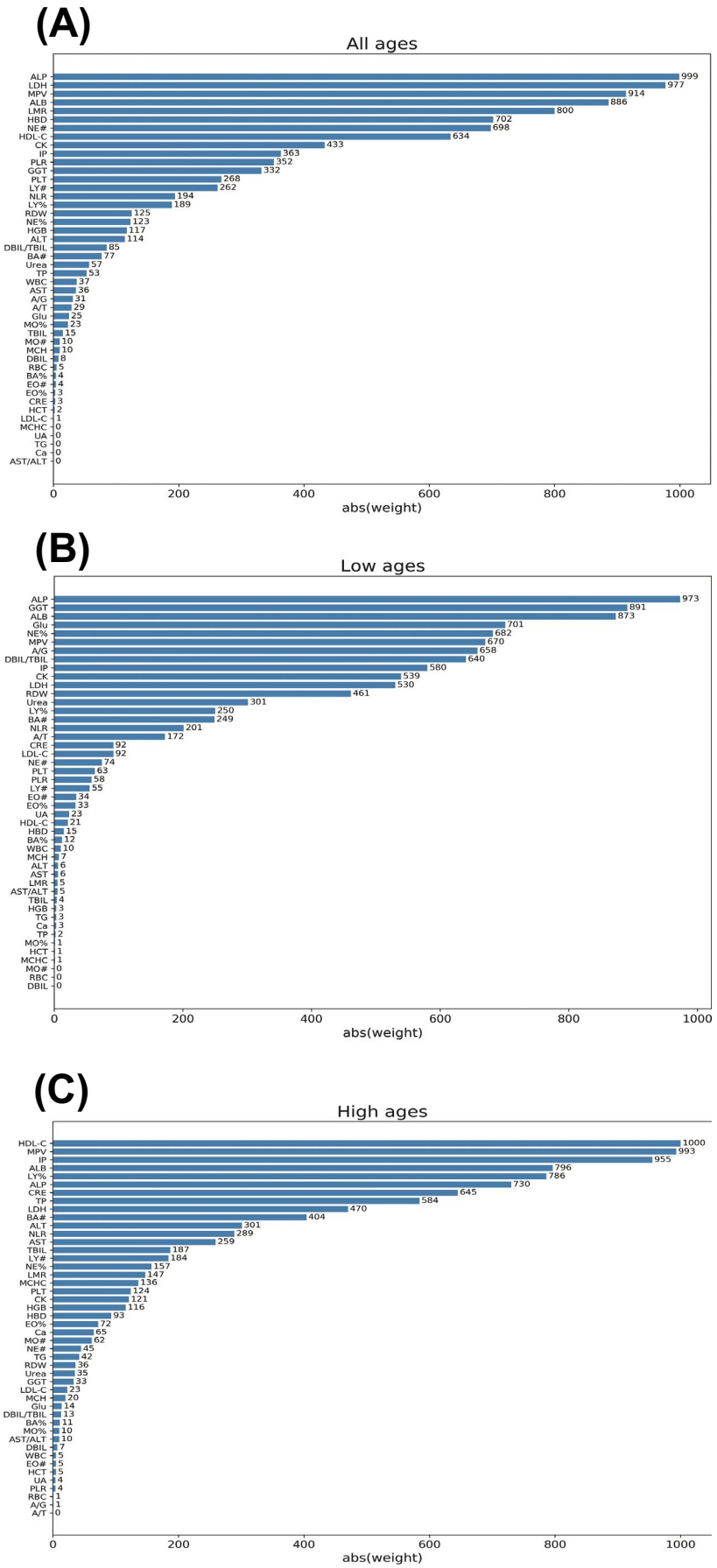

FIGURE 1 Importance matrix plot of logistic regression as described in Materials and Methods. The rankings of 46 clinical parameters for all ages (A), low ages (B), and high ages (C) were shown

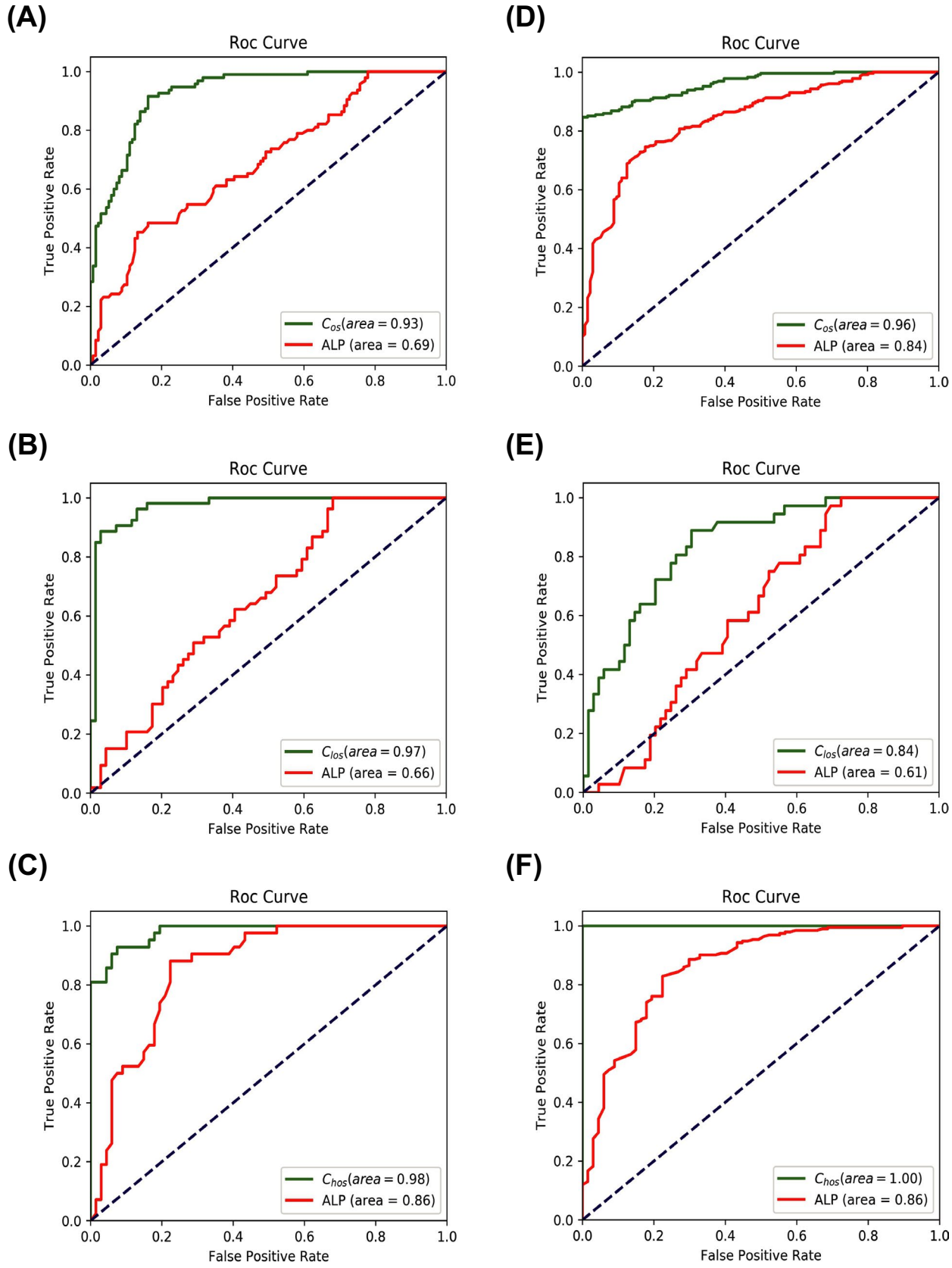

FIGURE 2 Performance of C_{os}, $C_{los,}$ and C_{hos} to detect osteosarcoma (OS) in the validation cohort. ROC curves were shown for distinguishing individuals with OS from healthy controls (HCs) in all ages (A), low ages (B), and high ages (C), respectively. Performance of C_{os}, $C_{los,}$ and C_{hos} to discriminate OS from primary benign bone lesions in all ages (D), low ages (E), and high ages (F) was shown as well

TABLE 2 Performance of C_{os}, C_{los}, $C_{hos,}$ and ALP for detection of osteosarcoma

	Accuracy % (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)
Osteosarcoma vs healthy control								
All ages								
C_{os}	84.0 (79.3-88.7)	84.6 (78.5-90.6)	83.2 (75.6-90.7)	87.8 (83.6-92.0)	79.0 (73.7-84.3)	5.02 (3.19-7.89)	0.19 (0.12-0.28)	27.04 (13.28-55.03)
ALP	68.8 (62.9-84.8)	83.1 (76.8-89.4)	48.4 (38.4-58.5)	69.8 (63.8-75.7)	66.7 (60.6-72.7)	1.61 (1.31-1.99)	0.35 (0.23-0.54)	4.61 (2.53-8.42)
≤15 y								
C_{los}	88.5 (82.9-94.2)	85.5 (77.2-93.8)	92.5 (85.3-99.6)	93.7 (89.3-98.0)	83.1 (76.4-89.7)	11.33 (4.39-29.22)	0.16 (0.09-0.28)	72.28 (21.34-244.78)
ALP	59.8 (51.1-68.5)	76.8 (66.9-86.8)	37.7 (24.7-50.8)	61.6 (53.0-70.3)	55.6 (46.7-64.4)	1.23 (0.96-1.58)	0.61 (0.35-1.07)	2.01 (0.91-4.42)
>15 y								
C_{hos}	91.7 (86.6-96.9)	88.1 (80.3-95.8)	97.6 (93.0-100.0)	98.3 (95.9-100.0)	83.7 (76.7-90.6)	36.99 (5.32-256.99)	0.12 (0.06-0.23)	302.38 (36.41-2510.94)
ALP	81.7 (74.4-88.9)	77.6 (67.6-87.6)	88.1 (78.3-97.9)	91.2 (85.9-96.5)	71.2 (62.6-79.7)	6.52 (2.84-14.99)	0.25 (0.16-0.40)	25.65 (8.57-76.79)
Osteosarcoma vs primary benign bone lesions								
All ages								
C_{os}	87.6 (84.3-91.0)	86.0 (80.2-91.9)	88.6 (84.5-92.7)	81.8 (77.9-85.8)	91.4 (88.5-94.3)	7.54 (5.22-10.90)	0.16 (0.10-0.24)	47.84 (25.38-90.18)
ALP	52.2 (47.1-57.3)	97.8 (95.3-100.0)	25.0 (19.4-30.6)	43.8 (38.7-48.8)	95.0 (92.8-97.2)	1.30 (1.20-1.41)	0.09 (0.03-0.28)	14.78 (4.53-48.23)
≤15 y								
C_{los}	76.2 (68.0-84.3)	76.8 (66.9-86.8)	75.0 (60.9-89.1)	85.5 (78.7-92.2)	62.8 (53.5-72.0)	3.07 (1.72-5.49)	0.31 (0.19-0.49)	9.94 (3.89-25.41)
ALP	61.0 (51.6-70.3)	71.0 (60.3-81.7)	41.7 (25.6-57.8)	70.0 (61.2-78.8)	42.9 (33.4-52.3)	1.22 (0.89-1.67)	0.70 (0.41-1.19)	1.75 (0.75-4.06)
>15 y								
C_{hos}	98.5 (97.0-100.0)	94.0 (88.4-99.7)	100.0	100.0	98.0 (96.2-99.7)	NA	0.06 (0.02-0.15)	NA
ALP	81.5 (76.7-86.2)	77.6 (67.6-87.6)	82.8 (77.5-88.1)	61.2 (55.2-67.1)	91.4 (88.0-94.8)	4.52 (3.23-6.32)	0.27 (0.17-0.42)	16.70 (8.41-33.17)

Abbreviation: NA, not applicable.

In this study, we established diagnosis models of OS based on large-scale clinical laboratory data and identify classifiers for different age groups that could differentiate individuals with OS from healthy individuals. Then, we determined whether these classifiers discriminated OS from primary benign bone lesions effectively. In addition, we investigated the performances of the established diagnostic classifiers for detection of small-size and early-stage OS. Finally, the capability of these classifiers to detect ALP-negative OS was also evaluated.

2 | MATERIALS AND METHODS

2.1 | Study population

A total of 458 patients initially diagnosed with OS, and 312 healthy individuals who had undergone a medical examination in Peking University People's Hospital from September 2012 to July 2018 were included in this study. All samples were randomly divided into training cohort and validation cohort. The training cohort included 322 patients and 217 healthy individuals, while the validation cohort included 136 patients and 95 healthy individuals.

To test whether the classifier can discriminated OS from primary benign bone lesions, 228 patients with primary benign bone lesions, including 10 with simple bone cyst, 5 with aneurysmal bone cyst, 42 with osteoma, 10 with osteoid osteoma, 100 with giant cell tumor of bone, 19 with hemangioma, 29 with osteochondroma, and 13 with enchondroma, were also used as test set. All patients with OS and benign bone lesions were pathologically confirmed. This study was conducted according to the principles of the Declaration of Helsinki and approved by the Ethics Committee of Peking University People's Hospital. The requirement to obtain informed consent was waived because of the retrospective nature of the study.

The medical records of the included patients were reviewed, and hematological and biochemical test results before treatment were obtained. The clinicopathological characteristics, including age, gender, pathological fracture, clinical stage, pulmonary metastasis, tumor size, tumor location, preoperative chemotherapy, and Huvos grade, of OS patients were extracted. In addition, the hematological and biochemical test results of the medical examination in healthy individuals were obtained as well. A total of 43 common hematological and biochemical parameters, and three ratios which were reported to be associated with OS, including LMR, NLR, and PLR,[13,14] were included in this study (Table S1). The reference interval of serum ALP was that: 0-15 y, 42-390 U/L; 16-18 y, 52-171 U/L; >19 y (male), 45-125 U/L; 19-49 y (female), 35-100 U/L; >50 y (female), 50-135 U/L. Serum ALP below the upper limit of the reference interval was considered as ALP-negative.

2.2 | Feature selection and the establishment of the diagnostic classifiers

Logistic regression (LR) was performed with Python 3.7.3 to establish the diagnostic classifiers. As a statistical model, LR uses a logistic function to model a binary dependent variable and learns the coefficients of each feature. Features used to establish the diagnostic model were selected based on the training cohort. In each round of feature selection process, 70% of OS patients and health individuals in training cohort were randomly selected to build a logistic regression model based on all the 46 features. In order to eliminate the differences between features, we rescaled the data by applying *z*-score transformation to each feature, which ensures that each feature obeys a standard normal distribution. Since our data have been standardized, the absolute value of the coefficients in the logistic regression model can reflect the importance of corresponding features. The above process was repeated 1000 times, and each time, the top-10 ranked features were recorded. Finally, all the 46 features were ranked by their frequencies in the records and the top-10 ranked features were selected to establish the diagnostic model. Analysis of the receiver operating characteristic (ROC) curves and the area under the ROC curves (AUC) were performed using Python 3.7.3.

2.3 | Statistical analysis

All statistical analyses were carried out with GraphPad Prime 5.01 or SPSS software. The clinical parameters are expressed as mean ± standard deviation. Results between OS patients and healthy individuals were compared using student's *t* test for values that followed a Gaussian distribution, and using Mann-Whitney *U* test for values that did not follow a Gaussian distribution. All the statistical tests were 2-tailed. A P value < .05 was considered statistically significant.

3 | RESULTS

3.1 | Establishment of diagnostic classifiers for osteosarcoma

We collected 770 participants, including 458 OS patients and 312 healthy controls (HCs), and randomly divided into training cohort and validation cohort (Figure S1). For each cohort, participants were divided into two age groups: low ages (≤15 y) and high ages (>15 y), because of the great difference in the reference interval of serum ALP between two age groups. The age and gender of patients with OS and HCs were well matched in each group (Table S2). Clinicopathological characteristics of OS patients in each group were displayed in Table 1. A total of 46 clinical laboratory parameters (Table S1) were used to develop the diagnostic models.

Classifiers were built for three age groups all ages (C_{os}), low ages (C_{los}), and high ages (C_{hos}), respectively. For each age group, top ten features were selected to build the diagnostic model (Figure 1). Details of the feature selection process were described in the materials and methods section. Then, for each age group, the classifier was built with the training set on the top selected features. Notably, all included parameters showed statistically significant differences between OS patients and HCs (Tables S3-S5).

The classifiers for all ages (C_{os}), low ages (C_{los}), and high ages (C_{hos}) were as follows. In order to make these classifiers easy to use, we reversed the z-score to the actual value of each feature and derived the final formulas of these classifiers. If the value of the corresponding classifier was lower than 0.5, the detected sample was predicted as OS; otherwise as non-OS.

$$C_{os} = -16.949 - 0.4604 \times \text{NE\#} + 0.6517 \times \text{MPV} - 0.0051 \times \text{ALP} - 0.0329 \times \text{LDH} + 0.0337 \times \text{HBD} + 0.0092 \times \text{CK} + 0.2345 \times \text{ALB} + 2.5718 \times \text{HDL\_C} - 0.2579 \times \text{IP} - 0.0249 \times \text{LMR}$$

$$C_{los} = -3.0437 - 10.2969 \times \text{NE\%} + 0.6227 \times \text{MPV} - 0.1086 \times \text{GGT} - 0.0051 \times \text{ALP} + 0.0104 \times \text{CK} + 0.2508 \times \text{ALB} - 0.9695 \times \text{Glu} - 2.004 \times \text{IP} + 2.0102 \times \text{A/G} - 5.3642 \times \text{DBIL/TBIL}$$

$$C_{hos} = -25.0011 + 10.0529 \times \text{LY\%} + 19.9989 \times \text{BA\#} + 0.7464 \times \text{MPV} - 0.024 \times \text{ALP} - 0.0117 \times \text{LDH} + 0.0837 \times \text{TP} + 0.2004 \times \text{ALB} + 0.0414 \times \text{CRE} + 4.8913 \times \text{HDL\_C} - 4.3709 \times \text{IP}$$

(Annotations: The features in the above formulas represent their actual values. Units: NE#, 10^9/L; MPV, fL; ALP, U/L; LDH, U/L; HBD, U/L; CK, U/L; ALB, g/L; HDL_C, mmol/L; IP, mmol/L; GGT, U/L; Glu, mmol/L; BA#, 10^9/L; TP, g/L; CRE, μmol/L).

ROC analyses were conducted to evaluate the performances of these classifiers in the validation cohort. As shown in Figure 2A and Table 2, C_{os} had an AUC 0.93 to discriminate all individuals with OS from HCs, and the accuracy, sensitivity, and specificity were 84.0%, 84.6%, and 83.2%, respectively. However, the AUC (0.69), accuracy (68.8%), sensitivity (83.1%), and specificity (48.4%) for ALP were much lower than those for C_{os} in all ages. For low ages, the C_{los} had an AUC 0.97 to discriminate relatively young individuals with OS from HCs, and the accuracy, sensitivity, and specificity were 88.5%, 85.5%, and 92.5%, respectively, which was also much better than ALP (AUC 0.66, accuracy 59.8%, sensitivity 76.8%, and specificity 37.7%) (Figure 2B, Table 2). For high ages, C_{hos} had an AUC 0.98 to discriminate relatively old individuals with OS from HCs, and the accuracy, sensitivity, and specificity were 91.7%, 88.1%, and 97.6%, respectively, while the AUC (0.86), accuracy (81.7%), sensitivity (77.6%), and specificity (88.1%) for ALP was lower than those for C_{hos} as well (Figure 2C, Table 2). Together, these results suggest that the performances of C_{os}, $C_{los,}$ and C_{hos} were greater than that of ALP on OS diagnosis.

3.2 | Performances of C_{os}, $C_{los,}$ and C_{hos} to discriminate OS from primary benign bone lesions

It is also critical for clinical practice to differentiate OS and primary benign bone lesions, and a good OS classifier should be sufficiently sensitive and specific to exclude benign bone diseases. To validate whether our models can discriminate OS from primary benign bone lesions, 228 patients of eight primary benign bone lesions were used to test the prediction performance of the established OS classifiers. Basic demographics of each disease were listed in Table S6. These patients were also divided into two age groups: low ages (≤15 y) and high ages (>15 y), and the age and gender of patients with OS and primary benign bone lesions were well matched in each group as well (Table S7).

Then, ROC analyses were performed, and the performances of C_{os}, $C_{los,}$ and C_{hos} were compared with that of ALP. Although the sensitivity of ALP was higher than that of C_{os} (97.8% vs 86.0%), C_{os} was more accurate (87.6% vs 52.2%) and specific (88.6% vs 25.0%) to discriminate OS from primary benign bone lesions, and the AUC for C_{os} was larger than that for ALP (0.96 vs 0.84) (Figure 2D and Table 2). In low ages, C_{los} showed much greater performance than ALP, with larger AUC (0.84 vs 0.61), and higher accuracy (76.2% vs 61.0%), sensitivity (76.8% vs 71.0%), and specificity (75.0% vs 41.7%) (Figure 2E and Table 2). In addition, the performance of C_{hos} in high ages was slightly better than that of ALP, and C_{hos} had a larger AUC (1.00 vs 0.86) and higher accuracy (98.5% vs 81.5%), sensitivity (94.0% vs 77.6%), and specificity (100.0% vs 82.8%) compared to ALP as well (Figure 2F and Table 2). These data suggest that these established classifiers showed great performances to discriminate OS from primary benign bone lesions as well.

3.3 | Performances of established classifiers to detect small-size and early-stage OS

We next determined whether the established C_{os}, $C_{los,}$ and C_{hos} were beneficial for early diagnosis of OS patients. ROC analyses were performed on small-size (tumor size ≤ 6 cm) and early-stage (clinical stage I-II) OS in the validation cohort. For small-size OS, C_{os} identified small-size tumors with larger AUC (0.93 vs 0.63), and higher accuracy (83.6% vs 57.2%), sensitivity (85.1% vs 62.7%), and specificity (82.6% vs 53.3%) in all ages (Figure 3A, Table 3). Similarly, C_{los} had larger AUC (0.95 vs 0.65) to diagnose low age individuals with small-size OS, and the accuracy (88.4%), sensitivity (94.1%), and specificity (84.6%) for C_{los} were higher than those for ALP (accuracy 59.3%, sensitivity 70.6%, specificity 51.9%) as well (Figure 3B, Table 3). Although C_{hos} was less specific than ALP (87.5% vs 90.0%) to distinguish patients with small-size tumors from HCs in high ages, it still showed better performance than ALP with larger AUC (0.98 vs 0.76), and higher accuracy (90.4% vs 76.7%) and sensitivity (93.9% vs 60.6%) (Figure 3C, Table 3).

For early-stage OS, C_{os} had a larger AUC to identify all ages with early-stage tumors (0.93 vs 0.65) with higher accuracy (84.9% vs 60.0%), sensitivity (87.0% vs 67.0%), and specificity (82.4% vs 51.8%) compared to ALP (Figure 3D, Table 3). Similarly, C_{los} also showed better performance than ALP to detect early-stage OS in low ages with larger AUC (0.96 vs 0.65), and higher accuracy (89.1% vs 59.4%), sensitivity (90.2% vs 62.7%), and specificity (88.0% vs 56.0%) (Figure 3E, Table 3). In addition, C_{hos} was more sensitive (89.8% vs 73.5%) and specific (94.3% vs 88.6%) than ALP, with larger AUC (0.98 vs 0.83) and higher accuracy (91.7% vs 79.8%) to discriminate patients with early-stage tumors from HCs in high ages (Figure 3F, Table 3). Collectively, these findings suggest that C_{os}, $C_{los,}$ and C_{hos} have important clinical significance for the early diagnosis of OS.

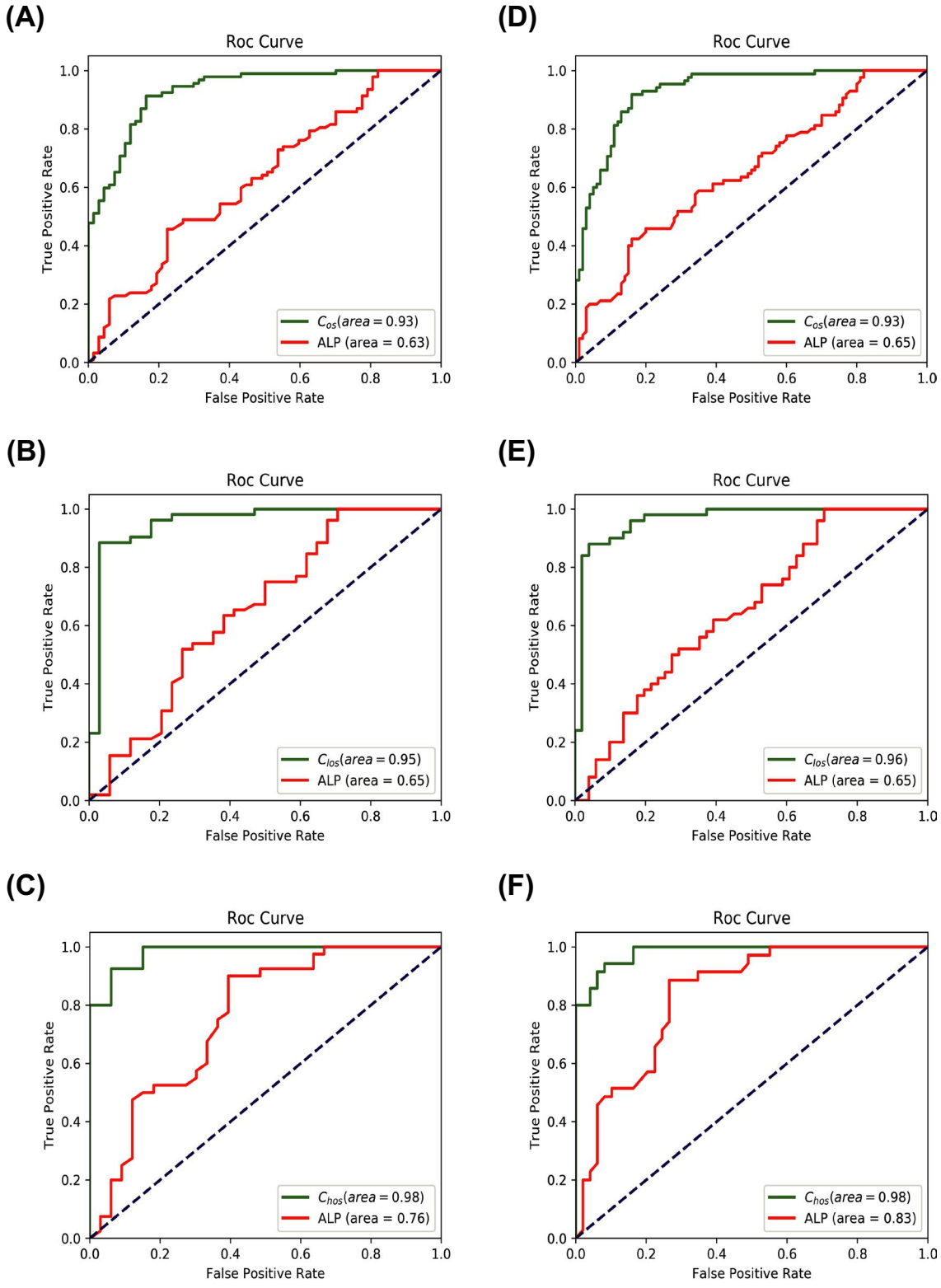

FIGURE 3 Performance of C_{os}, $C_{los,}$ and C_{hos} to detect small-size and early-stage osteosarcoma (OS). ROC curves from the validation cohort showed performance to distinguish individuals with small-size OS (tumor size ≤ 6 cm) from healthy controls (HCs) in all ages (A), low ages (B), and high ages (C), respectively. Performance to distinguish individuals with early-stage OS (clinical stage I-II) from HCs in all ages (D), low ages (E), and high ages (F) was shown as well

TABLE 3 Performance of C_{os}, C_{los}, $C_{hos,}$ and ALP to detect small-size and early-stage osteosarcoma

	Accuracy % (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)
Small-size tumor								
All ages								
C_{os}	83.6 (77.9-89.4)	85.1 (76.5-93.6)	82.6 (74.9-90.4)	78.1 (71.7-84.5)	88.4 (83.4-93.4)	4.89 (3.10-7.72)	0.18 (0.10-0.32)	27.08 (11.44-64.10)
ALP	57.2 (49.5-64.9)	62.7 (51.1-74.3)	53.3 (43.1-63.5)	49.4 (41.6-57.2)	66.2 (58.9-73.6)	1.34 (1.01-1.79)	0.70 (0.49-1.01)	1.91 (1.01-3.64)
≤15 y								
C_{los}	88.4 (81.6-95.1)	94.1 (86.2-100.0)	84.6 (74.8-94.4)	80.0 (71.5-88.5)	95.7 (91.3-100.0)	6.12 (3.22-11.64)	0.07 (0.02-0.27)	88.00 (17.50-442.47)
ALP	59.3 (48.9-69.7)	70.6 (55.3-85.9)	51.9 (38.3-65.5)	49.0 (38.4-59.5)	73.0 (63.6-82.4)	1.47 (1.03-2.10)	0.57 (0.32-1.01)	2.59 (1.04-6.48)
>15 y								
C_{hos}	90.4 (83.7-97.2)	93.9 (85.8-100.0)	87.5 (77.3-97.7)	86.1 (78.2-94.0)	94.6 (89.4-99.8)	7.52 (3.30-17.14)	0.07 (0.02-0.27)	108.50 (19.63-599.66)
ALP	76.7 (67.0-86.4)	60.6 (43.9-77.3)	90.0 (80.7-99.3)	83.3 (74.8-91.9)	73.5 (63.3-83.6)	6.06 (2.30-15.98)	0.44 (0.28-0.68)	13.85 (3.98-48.18)
Early-stage tumor								
All ages								
C_{os}	84.9 (79.7-90.0)	87.0 (80.4-93.6)	82.4 (74.2-90.5)	85.3 (80.2-90.4)	84.3 (79.1-89.6)	4.93 (3.10-7.85)	0.16 (0.09-0.26)	31.23 (13.94-69.97)
ALP	60.0 (52.9-67.1)	67.0 (57.8-76.2)	51.8 (41.1-62.4)	62.0 (55.0-69.0)	57.1 (50.0-64.3)	1.39 (1.07-1.80)	0.64 (0.45-0.90)	2.18 (1.20-3.95)
≤15 y								
C_{los}	89.1 (83.0-95.2)	90.2 (82.0-98.4)	88.0 (79.0-97.0)	88.5 (82.2-94.7)	89.8 (83.9-95.7)	7.52 (3.53-16.01)	0.11 (0.05-0.26)	67.47 (19.20-237.08)
ALP	59.4 (49.8-69.0)	62.7 (49.5-76.0)	56.0 (42.2-69.8)	59.3 (49.7-68.8)	59.6 (50.0-69.1)	1.43 (0.98-2.08)	0.67 (0.43-1.03)	2.14 (0.97-4.75)
>15 y								
C_{hos}	91.7 (85.8-97.6)	89.8 (81.3-98.3)	94.3 (86.6-100.0)	95.7 (91.3-100.0)	86.8 (79.6-94.1)	15.71 (4.08-60.56)	0.11 (0.05-0.25)	145.20 (26.50-795.48)
ALP	79.8 (71.2-88.4)	73.5 (61.1-85.8)	88.6 (78.0-99.1)	90.0 (83.6-96.4)	70.5 (60.7-80.2)	6.43 (2.52-16.42)	0.30 (0.19-0.48)	21.46 (6.34-72.64)

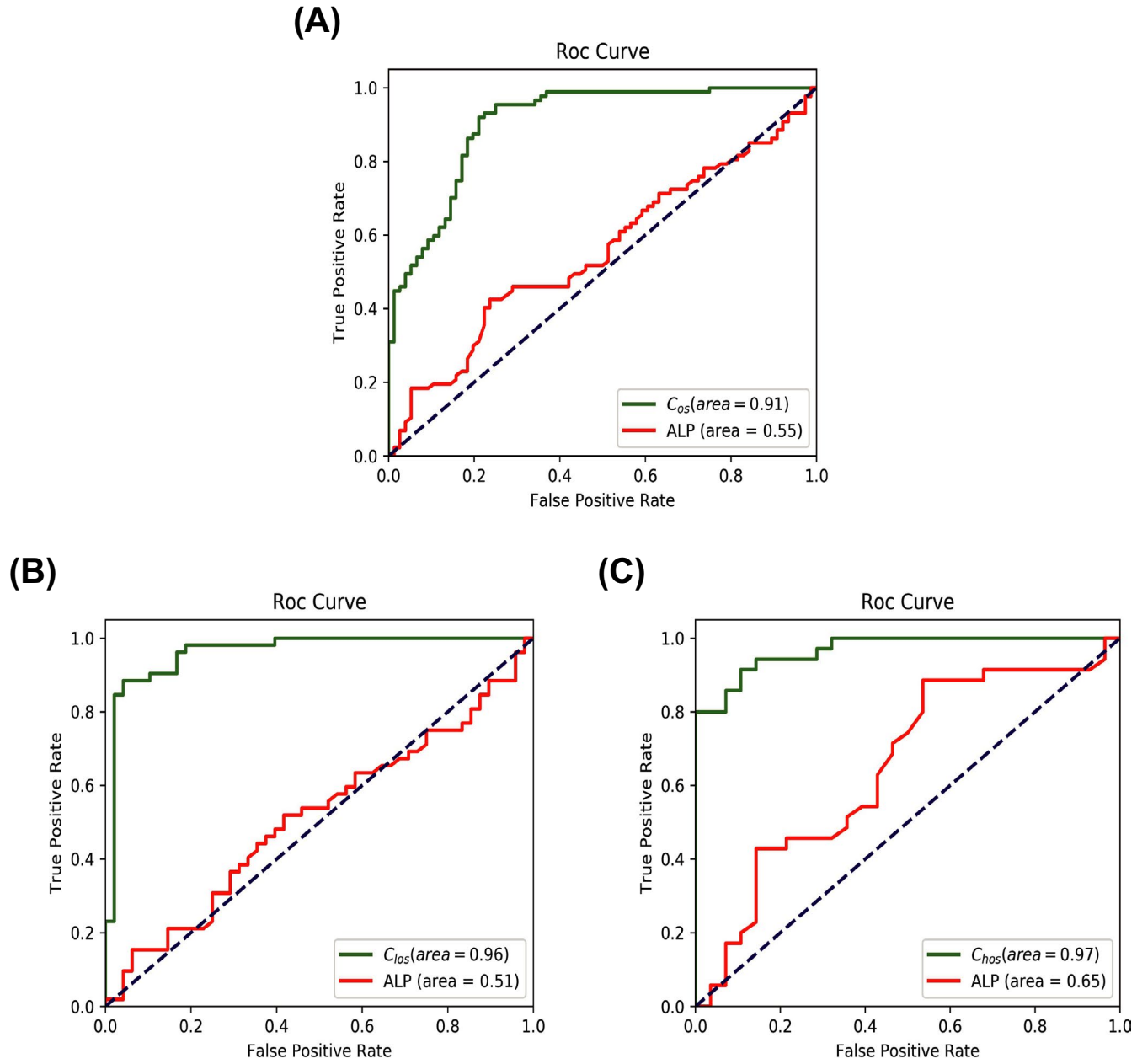

FIGURE 4 Performance of C_{os}, $C_{los,}$ and C_{hos} to detect ALP-negative osteosarcoma (OS). ROC curves from the validation cohort showed performance to distinguish individuals with ALP-negative OS (serum ALP below the upper limit of the reference interval) from healthy controls (HCs) in all ages (A), low ages (B), and high ages (C), respectively

3.4 | Performances of C_{os}, $C_{los,}$ and C_{hos} to detect ALP-negative OS

Over 40% of OS patients possessed normal/negative pre-treatment serum ALP as described before. Next, to evaluate the diagnostic performances of C_{os}, $C_{los,}$ and C_{hos} to identify ALP-negative OS, ROC analyses were performed in the validation cohort as well. As shown in Figure 4A and Table 4, C_{os} had a larger AUC to identify all ages with ALP-negative tumors (0.91 vs 0.55) with higher accuracy (82.2% vs 52.8%), sensitivity (81.6% vs 53.9%), and specificity (82.8% vs 51.7%) compared to ALP. In low ages, C_{los} had a great performance to detect ALP-negative tumors with an AUC 0.96, and the accuracy, sensitivity, and specificity were 88.0%, 83.3%, and 92.3%, respectively, which was much higher than those for ALP (AUC 0.51, accuracy 52.0%, sensitivity 66.7%, and specificity 38.5%) (Figure 4B and Table 4). Although ALP had a same sensitivity as C_{hos} (89.3% vs 89.3%) to diagnose ALP-negative tumors in high ages, C_{hos} was more accurate (88.9% vs 50.8%) and specific (88.6% vs 20.0%), and the AUC for C_{hos} was a little larger than that for ALP (0.97 vs 0.65) (Figure 4C and Table 4). These results clearly demonstrate that these established classifiers showed much better performances than ALP on ALP-negative OS detection.

TABLE 4 Performance of C_{os}, C_{los}, $C_{hos,}$ and ALP to detect ALP-negative osteosarcoma

	Accuracy % (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)
All ages								
C_{os}	82.2 (76.3-88.1)	81.6 (72.9-90.3)	82.8 (74.8-90.7)	80.5 (74.4-86.6)	83.7 (78.1-89.4)	4.73 (2.95-7.59)	0.22 (0.14-0.36)	21.26 (9.52-47.48)
ALP	52.8 (45.1-60.4)	53.9 (42.7-65.2)	51.7 (41.2-62.2)	49.4 (41.7-57.1)	56.3 (48.6-63.9)	1.12 (0.83-1.51)	0.89 (0.65-1.22)	1.26 (0.68-2.33)
≤15 y								
C_{los}	88.0 (81.6-94.4)	83.3 (72.8-93.9)	92.3 (85.1-99.6)	90.9 (85.3-96.5)	85.7 (78.9-92.6)	10.83 (4.19-28.01)	0.18 (0.10-0.34)	60.0 (16.82-213.97)
ALP	52.0 (42.2-61.8)	66.7 (53.3-80.0)	38.5 (25.2-51.7)	50.0 (40.2-59.8)	55.6 (45.8-65.3)	1.08 (0.81-1.45)	0.87 (0.51-1.47)	1.25 (0.55-2.84)
>15 y								
C_{hos}	88.9 (81.1-96.6)	89.3 (77.8-100.0)	88.6 (78.0-99.1)	86.2 (77.7-94.7)	91.2 (84.2-98.2)	7.81 (3.08-19.82)	0.12 (0.04-0.35)	64.58 (13.21-315.75)
ALP	50.8 (38.4-63.1)	89.3 (77.8-100.0)	20.0 (6.7-33.3)	47.2 (34.8-59.5)	70.0 (58.7-81.3)	1.12 (0.91-1.38)	0.54 (0.15-1.88)	2.08 (0.49-8.94)

4 | DISCUSSION

In this study, diagnostic classifiers for OS in different age groups were established with logistical regression model. Top ten clinical parameters in each weight ranking were included, respectively. All classifiers, including C_{os}, C_{los}, and C_{hos}, had higher sensitivity and specificity than ALP to distinguish individuals with OS from HCs. In addition, these classifiers could discriminate OS from primary benign bone lesions. Furthermore, C_{os}, C_{los}, and C_{hos} had better performances than ALP to detect small-size, early-stage, and ALP-negative OS. These findings highlight the potential of C_{os}, C_{los}, and C_{hos} as biomarkers for OS diagnosis at early clinical stages.

Despite of the importance of early diagnosis of OS, there is still no effective method at present. Early signs and symptoms, such as pain and pathological fracture, are not specific for OS, often resulting in initial misdiagnosis and delayed correct treatment for several months.[17] Biopsy is the gold standard for definitive diagnosis of OS,[18] but it is invasive and costly and requires an experienced pathologist. Our findings suggest that C_{os}, C_{los}, and C_{hos} are promising non-invasive assessments for early diagnosis of OS with high sensitivity and specificity. Notably, the diagnostic performances of these classifiers to detect small-size and large-size tumors (Figure 3, Table 3, Figure S2, and Table S8), as well as their ability to detect early-stage and late-stage tumors (Figure 3, Table 3, Figure S3, and Table S9), were different from each other, suggesting their performances are associated with tumor size and stage. In addition, although all three classifiers showed high sensitivity and specificity, there was still a little difference between the performances of C_{os} and C_{los} for low ages, as well as those of C_{os} and C_{hos} for high ages, indicating that C_{os} combined with C_{los} or C_{hos} might be better for OS diagnosis than a single one in clinical practice.

In this study, many parameters were found to be dysregulated in OS, including those included in the classifiers (Tables S3-S5). Different clinical parameters were included in the classifiers for all, low, and high ages according to the weight ranking of all parameters in the corresponding age group. The difference in weight ranking might be due to the distinction of dysregulated parameter patterns in different age groups. In addition to ALP, many other parameters made big contributions to the classifiers as well. For example, the weights of LDH and MPV were only slightly lower than that of ALP in all ages, whereas those of GGT and ALB were close to that of ALP in low ages. Interestingly, the weight of ALP was only ranked sixth, behind HDL-C, MPV, IP, ALB, and LY% in high ages. These suggested the necessity of parameter combination to better diagnose OS.

ALP is widely distributed, especially in bone and liver and can be released in the circulation.[19] Many pathological conditions, such as liver disease, bone disease, endocrine disease, neoplasia, and other disorders, result in increased serum ALP activity.[20,21] Therefore, the specificity of ALP was unsatisfactory for OS diagnosis. In addition, previous studies have described the low sensitivity of ALP to detect OS.[10,22,23] Consistently, our study discovered that ALP was less sensitive and specific than our established classifiers to distinguish

all OS patients from HCs and patients with primary benign bone lesions. For small-size and early-stage tumors, the performance of ALP was also poorer than C_{os}, C_{yos}, and C_{oos}. This study also extended the previous literature and evaluated the diagnostic value of ALP for ALP-negative OS patients. As expected, ALP has much worse performance than established classifiers to detect ALP-negative OS patients. In contrast, the sensitivity and specificity of ALP to detect ALP-positive OS were high in each age group (Figure S4 and Table S10). Therefore, this study further confirms the limited value of ALP in OS diagnosis.

Some diagnostic classifiers for OS diagnosis have been established in previous studies. For example, metabolomic data were used to classify healthy individuals and patients with benign tumor or OS, and the established classifiers have good performances to distinguish tumors and healthy controls (largest AUC 0.99 for either training or testing set), whereas the performances to discriminate OS from benign tumor were poor (largest AUC 0.60 for either training or testing set).[24] Another study has also described a classifier established with integrative metabolomic and transcriptomic profiles, and the classifier had an AUC 0.83 to identify OS from healthy control.[25] In addition, a proteomic classifier was constructed with plasma proteomic profiles to distinguish OS from osteochondroma patients, and it achieved a high sensitivity (97%) and specificity (80%).[26] Compared with these studies, our research has several strengths. First, to guarantee the robustness of our conclusion, we recruited approximately 1000 participants, including patients with OS and benign bone lesions, as well as HCs, while the above studies included less than 100 samples separately. Second, both patients with different benign bone lesions and HCs were introduced as controls to assess the specificity of our classifiers, while some other studies included only HCs or benign tumors. Third, our classifiers were established with clinical parameters, which were widely used in clinic, while other classifiers were constructed with metabolomic, transcriptomic, or proteomic profiles, which required complex methods to acquire the data and were not suitable for clinical application at present.

Our study has a few limitations. First, there were only 46 hematological and biochemical parameters included in this study. Many other clinical parameters, such as C-reactive protein, were not included, because it is not essential to test these parameters in routine medical examination, resulting in difficulty in data collection in the control group. Second, our participants were collected from only one clinical center in China and most cases of OS were Han Chinese. The study power might be increased when more clinical parameters and clinical centers are recruited to the retrospective study in the future, and the ability of C_{os}, C_{yos}, and C_{oos} to detect OS in other ethnicities merits investigation.

In summary, our study suggests that our diagnostic classifiers could discriminate OS from HC and primary benign bone lesions effectively, and it is also valuable for identifying small-size, early-stage, and ALP-negative OS with high sensitivity and specificity. The parameters included in the classifiers are routinely used in clinic, and the data are easy to acquire. Therefore, analysis of C_{os}, C_{los}, and C_{hos} in clinical practice should be feasible.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China grant 81702788 and 61773025, and National Key Research and Development Program of China [2018YFA0507504].

ORCID

Lin-Lin Cao https://orcid.org/0000-0003-2958-6263

REFERENCES

1. Mirabello L, Troisi RJ, Savage SA. International osteosarcoma incidence patterns in children and adolescents, middle ages and elderly persons. *Int J Cancer*. 2009;125:229-234.
2. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. *CA Cancer J Clin*. 2014;64:83-103.
3. Isakoff MS, Bielack SS, Meltzer P, Gorlick R. Osteosarcoma: current treatment and a collaborative pathway to success. *J Clin Oncol*. 2015;33:3029-3035.
4. Kaste SC, Pratt CB, Cain AM, Jones-Wallace DJ, Rao BN. Metastases detected at the time of diagnosis of primary pediatric extremity osteosarcoma at diagnosis: imaging features. *Cancer*. 1999;86:1602-1608.
5. Botter SM, Neri D, Fuchs B. Recent advances in osteosarcoma. *Curr Opin Pharmacol*. 2014;16:15-23.
6. Duffy MJ. Role of tumor markers in patients with solid cancers: a critical review. *Eur J Intern Med*. 2007;18:175-184.
7. Han J, Yong B, Luo C, Tan P, Peng T, Shen J. High serum alkaline phosphatase cooperating with MMP-9 predicts metastasis and poor prognosis in patients with primary osteosarcoma in Southern China. *World J Surg Oncol*. 2012;10:37.
8. Marais LC, Bertie J, Rodseth R, Sartorius B, Ferreira N. Pre-treatment serum lactate dehydrogenase and alkaline phosphatase as predictors of metastases in extremity osteosarcoma. *J Bone Oncol*. 2015;4:80-84.
9. Bramer JA, Abudu AA, Tillman RM, Carter SR, Sumathi VP, Grimer RJ. Pre- and post-chemotherapy alkaline phosphatase levels as prognostic indicators in adults with localised osteosarcoma. *Eur J Cancer*. 2005;41:2846-2852.
10. Kim SH, Shin KH, Moon SH, et al. Reassessment of alkaline phosphatase as serum tumor marker with high specificity in osteosarcoma. *Cancer Med*. 2017;6:1311-1322.
11. Chen J, Sun MX, Hua YQ, Cai ZD. Prognostic significance of serum lactate dehydrogenase level in osteosarcoma: a meta-analysis. *J Cancer Res Clin Oncol*. 2014;140:1205-1210.
12. Funovics PT, Edelhauser G, Funovics MA, et al. Pre-operative serum C-reactive protein as independent prognostic factor for survival but not infection in patients with high-grade osteosarcoma. *Int Orthop*. 2011;35:1529-1536.
13. Liu T, Fang XC, Ding Z, Sun ZG, Sun LM, Wang YL. Pre-operative lymphocyte-to-monocyte ratio as a predictor of overall survival in patients suffering from osteosarcoma. *FEBS Open Bio*. 2015;5:682-687.
14. Xia WK, Liu ZL, Shen D, Lin QF, Su J, Mao WD. Prognostic performance of pre-treatment NLR and PLR in patients suffering from osteosarcoma. *World J Surg Oncol*. 2016;14:127.
15. Lamora A, Talbot J, Bougras G, et al. Overexpression of smad7 blocks primary tumor growth and lung metastasis development in osteosarcoma. *Clin Cancer Res*. 2014;20:5097-5112.
16. Li Y, Nakka M, Kelly AJ, et al. p27 Is a candidate prognostic biomarker and metastatic promoter in osteosarcoma. *Cancer Res*. 2016;76:4002-4011.
17. Widhe B, Widhe T. Initial symptoms and clinical features in osteosarcoma and Ewing sarcoma. *J Bone Joint Surg Am*. 2000;82:667-674.

18. Bielack S, Carrle D, Casali PG. Osteosarcoma: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann Oncol*. 2009;20(Suppl 4):137-139.
19. Buchet R, Millan JL, Magne D. Multisystemic functions of alkaline phosphatases. *Methods Mol Biol*. 2013;1053:27-51.
20. Lee TH, Kim WR, Poterucha JJ. Evaluation of elevated liver enzymes. *Clin Liver Dis*. 2012;16:183-198.
21. Fernandez NJ, Kidney BA. Alkaline phosphatase: beyond the liver. *Vet Clin Pathol*. 2007;36:223-233.
22. Bacci G, Picci P, Ferrari S, et al. Prognostic significance of serum alkaline phosphatase measurements in patients with osteosarcoma treated with adjuvant or neoadjuvant chemotherapy. *Cancer*. 1993;71:1224-1230.
23. Liu PP, Leung KS, Kumta SM, Lee KM, Fung KP. Bone-specific alkaline phosphatase in plasma as tumour marker for osteosarcoma. *Oncology*. 1996;53:275-280.
24. Zhi L, Reza Soroushmehr SM, Yingqi H, Min M, Yunping Q, Najarian K. Classifying osteosarcoma patients using machine learning approaches. *Conf Proc IEEE Eng Med Biol Soc*. 2017;2017:82-85.
25. Chen K, Zhu C, Cai M, et al. Integrative metabolome and transcriptome profiling reveals discordant glycolysis process between osteosarcoma and normal osteoblastic cells. *J Cancer Res Clin Oncol*. 2014;140:1715-1721.
26. Li Y, Dang TA, Shen J, et al. Identification of a plasma proteomic signature to distinguish pediatric osteosarcoma from benign osteochondroma. *Proteomics*. 2006;6:3426-3435.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Cao L-L, Chen Z, Yue Z, et al. Novel classifiers with clinical laboratory parameters for early detection of osteosarcoma. *J Clin Lab Anal*. 2020;34:e23189. https://doi.org/10.1002/jcla.23189