

NSCLC 骨转移风险预测模型的建立和验证

胡春晓^{1,2}, 刘亚锋^{1,2}, 苏忆欣¹, 郭健强¹, 张文婷¹, 王雪芹^{1,2},
谢军², 胡万发², 吴静^{1,3}, 邢应如^{1,2,3}, 胡东^{1,3}, 丁选胜¹

摘要 目的 构建列线图来预测非小细胞肺癌(NSCLC)患者骨转移的发生风险。方法 回顾性分析在医院确诊的NSCLC患者的临床资料,包括是否发生骨转移、年龄、性别、病理类型、吸烟状况、PS评分、TN分期、骨转移前是否有其他部位的转移、癌胚抗原(CEA)水平、甲胎蛋白(AFP)水平、血清钙浓度(Ca^{2+})、血清磷浓度(P)、碱性磷酸酶(ALP)水平,利用单因素及多因素 Logistic 回归分析建立预测模型,使用受试者工作特征曲线(ROC)、决策曲线分析法(DCA)验证模型的准确性及临床获益度,使用列线图进行模型可视化。结果 ROC 曲线下面积(AUC)显示,在建模组($n=138$)及验证组($n=92$)中,联合指标(年龄、性别、病理类型、CEA、ALP)预测的 AUC 值(建模组=0.792,验证组=0.629)均高于单一指标预测值。结论 该研究构建的预测模型效果良好,可为临床筛选 NSCLC 骨转移的高危患者提供参考。关键词 非小细胞肺癌;骨转移;危险因素;列线图

中图分类号 R 734.2

文献标志码 A 文章编号 1000-1492(2022)05-0832-05
doi:10.19405/j.cnki.issn1000-1492.2022.05.030

肺癌已成为全球发生率及病死率较高的癌症之一^[1],其中非小细胞肺癌(non-small cell lung cancer, NSCLC)比例 > 80%。随着医疗技术的发展,肺癌患者的生存期有了一定程度上延长,但五年生存率很低^[2],仍有 57% 的患者会发生远处转移,其中

骨转移发生率最高,占比 30% ~ 40%^[3]。

骨转移的早诊断、早治疗,能提高患者的生存期。目前,临床上骨转移常用的诊断方法是全身影像学检查,但在 NSCLC 的早期检出率较低,成本较高,当影像学显示肿瘤转移时,患者已处于肺癌晚期,预后较差^[4]。该研究通过筛选 NSCLC 患者发生骨转移的危险因素,构建列线图来早期预测 NSCLC 患者骨转移发生风险,为临床的早期诊治提供可靠依据。

1 材料与方法

1.1 病例资料 收集 2016 年 1 月—2020 年 12 月在安徽理工大学附属肿瘤医院(淮南东方医院集团肿瘤医院)和安徽理工大学附属医院收治的经病理学确诊的 NSCLC 患者的临床资料。排除标准:① 腺癌(LUAD)、鳞癌(LUSC)以外的病理类型。② 15 d 之内接受手术、放疗、化疗等治疗。③ 严重的骨质疏松、风湿性关节炎、外伤性骨折以及其他严重的并发症。④ 影响骨代谢的内分泌疾病(如糖尿病血糖异常、甲状腺和甲状旁腺疾病)。按照排除标准,最终纳入 230 例患者。

1.2 肿瘤标志物检测 采集两组患者空腹静脉血 3 ml,分离血清,癌胚抗原(carcinoembryonic antigen, CEA)、甲胎蛋白(alpha fetoprotein, AFP)采用化学发光法测定,CEA 参考值范围:0 ~ 10 $\mu\text{g/L}$,AFP 参考值范围:0 ~ 20 $\mu\text{g/L}$ 。

1.3 血清 Ca^{2+} 、磷(phosphorus, P)、碱性磷酸酶(alkaline phosphatase, ALP)检测 采集 2 组患者空腹静脉血 3 ml,对所有患者血清 Ca^{2+} 采用邻甲酚酞络合酮法测定,血清 P 采用磷钼酸盐法测定,血清 ALP 采用 NPP-AMP 法测定;血清 Ca^{2+} 参考范围:2.25 ~ 2.75 mmol/L,血清 P 参考范围:0.97 ~ 1.61 mmol/L,血清 ALP 参考范围:50 ~ 136 U/L。

1.4 统计学处理 采用易侗统计软件(Empower-Stats)进行统计分析,计量资料以 $\bar{x} \pm s$ 表示,采用 t 检验、方差分析比较。计数资料采用 χ^2 检验比较,以频数、构成比表示。将单因素分析筛选出的变量

2022-03-12 接收

基金项目:国家自然科学基金项目(编号:81971483);安徽省高校协同创新项目(编号:GXXT-2020-058);淮南市科技基金项目(编号:2018A367);安徽高校协同创新项目(编号:GXXT-2019-045)

作者单位:¹ 安徽理工大学医学院(安徽省职业健康安全工程实验室),淮南 232000

² 安徽理工大学附属肿瘤医院(淮南东方医院集团肿瘤医院),淮南 232000

³ 安徽工业粉尘深度净化与职业健康教育部重点实验室,淮南 232000

作者简介:胡春晓,女,硕士研究生;

丁选胜,男,博士,教授,责任作者, E-mail: xsding2013@163.com;

胡东,男,博士,教授,责任作者, E-mail: austhudong@126.com

纳入多因素 Logistic 回归分析,以确定独立危险因素。使用受试者工作特征曲线(ROC)对模型进行分析,并比较单一指标预测和联合指标预测的 ROC 曲线下面积(AUC)。“RMS”软件包绘制列线图 and 校准曲线。“pROC”软件包绘制 ROC 曲线,“RMDA”软件包通过决策曲线分析法(decision curve analysis, DCA)绘制临床获益曲线。

2 结果

2.1 建模组和验证组 NSCLC 骨转移基线情况描述 共纳入 230 例患者,按照 3 : 2 的比例随机分成建模组和验证组,其中 138 例为建模组,骨转移患者 70 例(50.7%),无骨转移患者 68 例(49.3%)。验

证组共 92 例患者,骨转移患者 47 例(51.1%),无骨转移患者 45 例(48.9%)(表 1)。

2.2 NSCLC 骨转移危险因素的单一因素及多因素分析 对建模组 138 例患者发生骨转移的危险因素进行单、多因素逻辑回归分析。在单因素分析结果中,以 $P < 0.1$ 为标准将年龄($P = 0.018$)、性别($P = 0.004$)、病理类型($P = 0.007$)、N 分期($P = 0.089$)、ALP($P = 0.015$)、CEA($P = 0.002$)纳入多因素 Logistic 回归分析,结果显示年龄($P = 0.034$)、性别($P = 0.011$)、ALP 水平($P = 0.003$)、CEA 水平($P = 0.003$)作为 NSCLC 骨转移的独立危险因子具有统计学意义,而病理类型($P = 0.053$)与骨转移密切相关但未显示统计学意义(表 2)。

表 1 建模组和验证组病例的基线数据[n(%)]

变量	建模组(n=138)	验证组(n=92)
年龄(岁, $\bar{x} \pm s$)	65.7 ± 9.9	70.5 ± 8.9
是否发生骨转移		
否	68 (49.3%)	45 (48.9%)
是	70 (50.7%)	47 (51.1%)
CEA(ng/ml)		
0~10	90 (65.2%)	57 (62.0%)
>10	48 (34.8%)	35 (38.0%)
AFP(ng/ml)		
0~20	137 (99.3%)	90 (98.9%)
>20	1 (0.7%)	1 (1.1%)
Ca ²⁺ (mmol/L)		
2.25~2.75	72 (52.2%)	29 (31.5%)
<2.25	64 (46.4%)	63 (68.5%)
>2.75	2 (1.4%)	0 (0.0%)
P(mmol/L)		
0.97~1.61	111 (80.4%)	56 (60.9%)
<0.97	24 (17.4%)	32 (34.8%)
>1.61	3 (2.2%)	4 (4.3%)
ALP(U/L)		
50~136	99 (71.7%)	34 (37.0%)
<50	7 (5.1%)	0 (0.0%)
>136	32 (23.2%)	58 (63.0%)
T分期		
T ₀ ~T ₂	53 (38.4%)	32 (34.8%)
T ₃ ~T ₄	85 (61.6%)	60 (65.2%)
N分期		
N ₀ ~N ₂	94 (68.1%)	40 (43.5%)
N ₃ 及N _x	44 (31.9%)	52 (56.5%)
PS评分(分)		
0~2	127 (92.0%)	70 (76.1%)
3~5	11 (8.0%)	22 (23.9%)
吸烟状况		
是	29 (21.0%)	2 (2.2%)
否	109 (79.0%)	90 (97.8%)
病理类型		
腺癌	98 (71.0%)	64 (69.6%)
鳞癌	40 (29.0%)	28 (30.4%)
性别		
女	45 (32.6%)	27 (29.3%)
男	93 (67.4%)	65 (70.7%)
骨转移前是否有其他部位的转移		
否	116 (84.1%)	82 (89.1%)
是	22 (15.9%)	10 (10.9%)

表 2 NSCLC 患者骨转移单因素及多因素分析[n(%)]

变量	未转移组(n=68)	转移组(n=70)	单因素分析 P值	多因素分析 P值
年龄(岁, $\bar{x} \pm s$)	67.81 ± 9.69	63.74 ± 9.82	0.018	0.034
CEA(ng/ml)			0.002	0.003
0~10	53 (77.94%)	37 (52.86%)		
>10	15 (22.06%)	33 (47.14%)		
AFP(ng/ml)			0.987	
0~20	68 (100.00%)	69 (98.57%)		
>20	0 (0.00%)	1 (1.43%)		
Ca ²⁺ (mmol/L)			0.294	
2.25~2.75	38 (55.88%)	35 (50.00%)		
<2.25	28 (41.18%)	35 (50.00%)		
>2.75	2 (2.94%)	0 (0.00%)	0.988	
P(mmol/L)			0.904	
0.97~1.61	54 (79.41%)	57 (81.43%)		
<0.97	12 (17.65%)	12 (17.14%)		
>1.61	2 (2.94%)	1 (1.43%)	0.547	
ALP(U/L)			0.170	0.291
50~136	56 (82.35%)	43 (61.43%)		
<50	2 (2.94%)	5 (7.14%)		
>136	10 (14.71%)	22 (31.43%)	0.015	0.003
T分期			0.175	
T ₀ ~T ₂	30 (44.12%)	23 (32.86%)		
T ₃ ~T ₄	38 (55.88%)	47 (67.14%)		
N分期			0.089	0.185
N ₀ ~N ₂	51 (75.00%)	43 (61.43%)		
N ₃ 及N _x	17 (25.00%)	27 (38.57%)		
PS评分(分)			0.988	
0~2	68 (100.00%)	59 (84.29%)		
3~5	0 (0.00%)	11 (15.71%)		
吸烟状况			0.904	
是	14 (20.59%)	15 (21.43%)		
否	54 (79.41%)	55 (78.57%)		
病理类型			0.007	0.053
腺癌	41 (60.29%)	57 (81.43%)		
鳞癌	27 (39.71%)	13 (18.57%)		
性别			0.004	0.011
女	14 (20.59%)	31 (44.29%)		
男	54 (79.41%)	39 (55.71%)		
骨转移前是否有其他部位的转移			0.989	
否	68 (100.00%)	48 (68.57%)		
是	0 (0.00%)	22 (31.43%)		

2.3 构建预测模型并验证 同样以 $P < 0.1$ 为标准将年龄、性别、病理类型、ALP、CEA 为自变量,以是否发生骨转移为因变量,进行 Logistic 回归,计算相应的回归系数,建立数学预测模型,NSCLC 患者发生骨转移的风险为 $P = 2.19168 + 1.26263 \times \text{CEA} + 0.76427 \times \text{ALP} - 1.07116 \times \text{病理类型} - 1.12207 \times \text{性别} - 0.04354 \times \text{年龄}$ 。对建模组绘制 ROC,CEA、年龄、ALP、病理类型、性别的 AUC 值分别为 0.625、0.607、0.604、0.394、0.382,其中 CEA 的 AUC 值最高,性别和病理类型的预测效能不佳。而联合指标(年龄、性别、病理类型、CEA、ALP)预测的 $\text{AUC} = 0.792$ 要高于单一指标预测,说明联合指标预测要比单一指标预测有更高的准确性(图 1A)。年龄、CEA、ALP 的特异度分别为 72.1%、77.9%、82.4%,灵敏度分别为 52.9%、47.1%、38.6%,而联合指标的特异度和灵敏度分别为 73.5%、77.1%(表 3),其综合表现优于单一指标。基于同一个模型,将验证组的 92 例 NSCLC 患者进行 ROC 分析,联合指标预测的 AUC 值为 0.629,特异度为 46.7%,灵敏度为 83.0%(表 3),ROC 曲线显示联合指标预测要比单一指标预测效能更佳(图 1B)。

2.4 预测模型的可视化 将多因素分析筛选出来的因素即年龄、性别、病理类型、ALP、CEA 纳入预测模型,使用“RMS”软件包建立列线图。分为每个危险因素对应的分值,总分为全部危险因素对应的分值相加得到的总分数,每一个总分对应下方的骨转移风险,见图 2。

2.5 绘制列线图的校准曲线及 DCA 绘制列线图的校准图,红线代表该模型的标准曲线,黑线代表该模型的校准曲线,校准图显示该模型的校准曲线趋势与标准曲线贴合程度尚可,说明该模型用于预测 NSCLC 患者骨转移风险具有可靠性(图 3)。通常应用 DCA 评价 NSCLC 骨转移风险预测模型的临床适用性,横线表示所有样本都是阴性,净获益为 0;

斜线表示所有样本都是阳性,净获益是个斜率为负值的反斜线。决策曲线距离 2 条极端线越远,说明列线图临床应用价值越高,DCA 的结果表明,列线图模型可在很大的风险阈值概率范围内提高对骨转移风险的预测能力。见图 4。

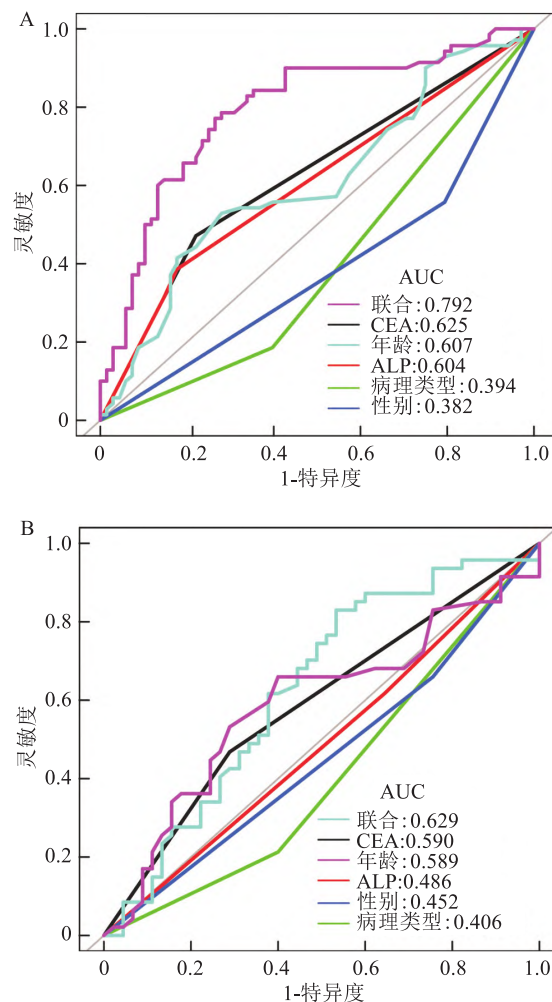


图 1 建模组及验证组的 ROC 曲线分析

3 讨论

骨相关事件(SREs)常在肺癌患者骨转移后发

表 3 建模组和验证组各因素及联合指标预测的 AUC

检测指标	建模组				验证组			
	AUC	95% CI	特异度	灵敏度	AUC	95% CI	特异度	灵敏度
联合	0.792	0.71 ~ 0.87	0.735	0.771	0.629	0.51 ~ 0.75	0.467	0.830
CEA	0.625	0.55 ~ 0.70	0.779	0.471	0.590	0.49 ~ 0.69	0.711	0.468
ALP	0.604	0.53 ~ 0.68	0.824	0.386	0.486	0.39 ~ 0.59	0.000	1.000
病理类型	0.394	0.32 ~ 0.47	0.000	1.000	0.406	0.31 ~ 0.50	0.000	1.000
性别	0.382	0.31 ~ 0.46	0.000	1.000	0.452	0.36 ~ 0.55	0.000	1.000
年龄	0.607	0.51 ~ 0.70	0.721	0.529	0.589	0.47 ~ 0.71	0.600	0.660

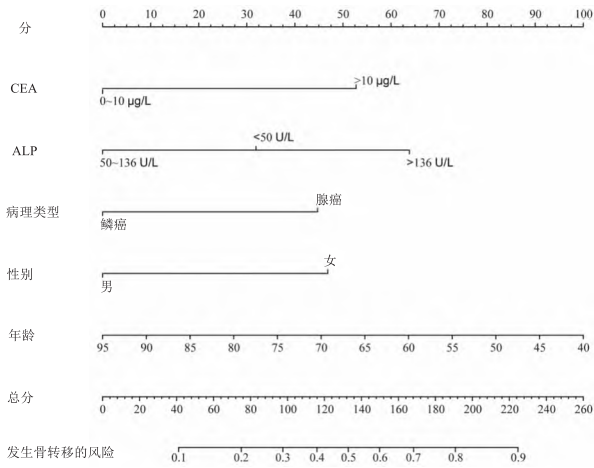


图2 NSCLC患者骨转移列线图预测模型

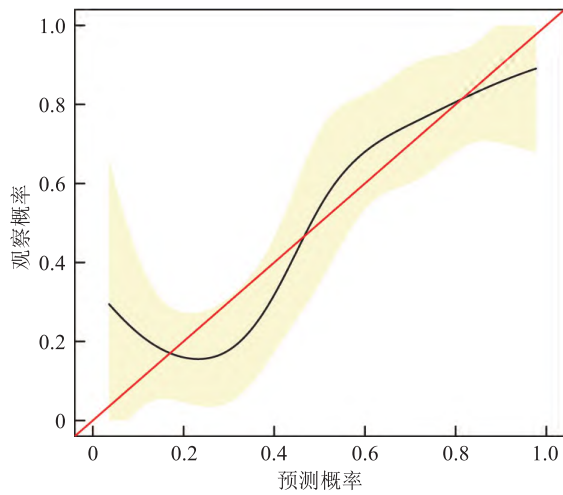


图3 列线图的校准曲线

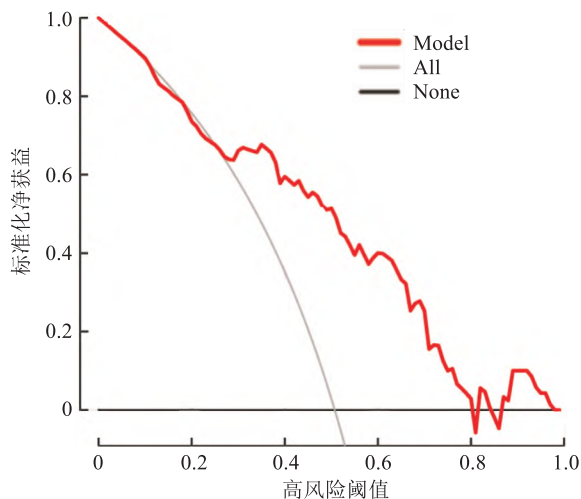


图4 NSCLC骨转移预测模型的DCA分析

生,包括疼痛、病理性骨折、脊髓压迫、高钙血症等。其中,疼痛是肺癌发生骨转移后最为明显的临床特征^[5]。但是,由于骨转移发病隐匿,骨相关事件大

多在晚期才会出现,往往导致医生不能及时发现骨转移发生,使患者错过诊治的最佳时机。CT、MRI、骨X线全身显像等常规影像学技术是目前临床上常用来诊断骨转移的方法,但也有一定局限性,例如患者往往在出现明显的临床特征后才做影像学检查,可能已处于肺癌晚期,此时已不能取得最佳的治疗效果。常规的影像学检查既不能在早期及时发现NSCLC患者骨转移的发生,又会给患者带来一定的经济压力,这就推动人们寻找更加简便适用的方法来预测NSCLC患者发生骨转移的早期风险,其对于延长患者生存期、改善临床治疗效果具有十分重要的意义。本研究共纳入了14个可能与NSCLC骨转移发生相关的因素,其中CEA ($P = 0.003$)、ALP ($P = 0.003$)、病理类型 ($P = 0.053$)、年龄 ($P = 0.034$)、性别 ($P = 0.011$)可作为NSCLC患者骨转移发生风险的预测因素。韩春明等^[6]研究表明,肺癌骨转移与病理类型具有相关性,腺癌骨转移发生率高,与本研究结果一致。有研究^[7]表明,骨转移的患者比未转移的患者更年轻;本研究同样表明年龄小的NSCLC患者更容易发生骨转移。在临床实践中,病理类型 ($P = 0.053$)有意义且该指标接近于0.05的统计学标准,所以这个指标可作为预测因素纳入预测模型。

Ayan et al^[8]研究NSCLC患者和SCLC患者血清CEA水平,结果显示,发生骨转移的NSCLC患者血清的CEA水平明显高于未发生骨转移患者。与其相符,在本研究中,骨转移组CEA水平显著高于未转移组。ALP是反映成骨细胞活性的重要标志物,不仅由成骨细胞分泌,还可通过肝脏、小肠和胎盘等各种器官产生。有文献^[9]报道,肺癌骨转移多为溶骨性骨破坏,在发生骨转移的乳腺癌及肺癌患者中,血清骨形成及骨吸收标志物均可见上升表达,骨转移的发生造成骨受损会使成骨细胞分泌大量ALP来修复受损骨骼,因此,肺癌骨转移的患者ALP水平显著增高,这与本研究结果一致。该研究显示,建模组中,骨转移和未转移组的NSCLC患者性别具有显著差异,女性发生NSCLC骨转移的风险更高,这与冯小双等^[10]的研究结果近似,骨转移在肺癌和食管癌患者转移部位中的占比男性低于女性 ($P < 0.01$)。有研究^[11]报道,血清Ca²⁺、ALP、细胞角蛋白19片段抗原21-1(CYFRA21-1)联合预测肺癌骨转移的AUC值为0.765,而在本研究中,联合指标预测的AUC值为0.792,提示后者预测效能更高。

列线图能够预测疾病风险或生存结局,整合了

各种个体危险因素,量化了每个变量的影响,从而可视化不同患者的结果^[12]。该模型可预测 10% ~ 90% 的骨转移风险。假设一个 NSCLC 男性患者测得 CEA 水平为 12.3 ng/ml, ALP 水平为 142 U/L, 病理类型为腺癌, 年龄 68 岁, 那该患者的总分为 53 (CEA > 10 ng/ml) + 64 (ALP > 136 U/L) + 45 (腺癌) + 49 (年龄 = 68) + 0 (男性) = 211, 说明该患者此刻骨转移的风险 = 86%, 发生骨转移的风险较大, 建议尽快行骨转移相关检测及治疗。

参考文献

- [1] Jun T, Zheng F S, Ren K M, et al. Long non-coding RNA UCA1 regulates the proliferation, migration and invasion of human lung cancer cells by modulating the expression of microRNA-143 [J]. *Eur Rev Med Pharmacol Sci*, 2018, 22(23):8343-52.
- [2] 秦新新, 邱 竞, 吕雪莲, 等. miR-17 通过调节脂肪酸代谢抑制非小细胞肺癌 A549 细胞的生长和转移 [J]. *安徽医科大学学报*, 2020, 55(7):1057-63.
- [3] 郭庆志, 吴梅娜, 安彤同, 等. 322 例非小细胞肺癌骨转移临床特点及治疗的回顾性分析 [J]. *中国肺癌杂志*, 2014, 17(9):656-62.
- [4] 史大宝, 高小卓, 张 瑞, 等. 非小细胞肺癌骨转移瘤的 MRI 影像及免疫组化特征 [J]. *现代肿瘤医学*, 2019, 27(17):3039-43.
- [5] Kosteva J, Langer C. The changing landscape of the medical management of skeletal metastases in nonsmall cell lung cancer [J]. *Curr Opin Oncol*, 2008, 20(2):155-61.
- [6] 韩春明. 肺癌骨转移的相关因素研究 [J]. *中国医药科学*, 2012, 2(10):16-17.
- [7] Zhou Y, Chen W Z, Peng A F, et al. Neuron-specific enolase, histopathological types, and age as risk factors for bone metastases in lung cancer [J]. *Tumour Biol*, 2017, 39(7):1010428317714194.
- [8] Ayan A K, Erdemci B, Orsal E, et al. Is there any correlation between levels of serum osteopontin, CEA, and FDG uptake in lung cancer patients with bone metastasis? [J]. *Rev Esp Med Nucl Imagen Mol*, 2016, 35(2):102-6.
- [9] 韩丽敏, 魏丽荣, 杜玉珍. 骨代谢标志物 tP1NP 和 β -CTx 及 BAP 在肺癌骨转移中的临床应用 [J]. *中华检验医学杂志*, 2017, 40(11):860-4.
- [10] 冯小双, 周昌明, 莫 森, 等. 以大型单中心医院登记为基础的 2 万例癌症转移患者的转移特征及生存分析 [J]. *中国癌症杂志*, 2021, 31(1):11-9.
- [11] 赵车冬, 徐 倩, 马 婧, 等. 血清钙离子、ALP 及 CYFRA21-1 对肺癌早期骨转移的预测价值 [J]. *分子诊断与治疗杂志*, 2020, 12(10):1323-7.
- [12] 黄 凯, 张长乐, 吴文涌, 等. 结肠癌列线图预后风险模型构建: 基于 SEER 数据库的回顾性研究 [J]. *安徽医科大学学报*, 2021, 56(2):299-30.

Establishment and validation of risk prediction model for bone metastasis of NSCLC

Hu Chunxiao^{1,2}, Liu Yafeng^{1,2}, Su Yixin¹, Guo Jianqiang¹, Zhang Wenting¹, Wang Xueqin^{1,2},
Xie Jun², Hu Wanfa², Wu Jing^{1,3}, Xing Yingru^{1,2,3}, Hu Dong^{1,3}, Ding Xuansheng¹

¹School of Medicine, Anhui University of Science & Technology (Anhui Province Engineering Laboratory of Occupational Health and Safety), Huainan 232000; ²Cancer Hospital Affiliated to Anhui University of Science and Technology (Huainan Oriental Hospital Group Cancer Hospital), Huainan 232000; ³Key Laboratory of Anhui Industrial Dust Deep Purification and Occupational Health Ministry of Education, Huainan 232000

Abstract Objective To construct nomogram to predict the risk of bone metastasis in patients with non-small cell lung cancer (NSCLC). **Methods** The clinical data of NSCLC patients diagnosed in the hospital were retrospectively analyzed, including the occurrence of bone metastasis, age, gender, pathological type, smoking status, PS score, TN stage, metastasis of other sites before bone metastasis, carcinoembryonic antigen (CEA) level, alpha fetoprotein (AFP) level, serum calcium (Ca^{2+}), serum phosphorus (P), alkaline phosphatase (ALP) level, which were determined by univariate and multivariate logistic regression analysis. Receiver operating characteristic curve (ROC) and decision curve analysis were used, DCA was used to verify the accuracy and clinical benefit of the model, and nomogram was used to visualize the model. **Results** Area under the ROC curve (AUC) showed that in the modeling group ($n = 138$) and the validation group ($n = 92$), the AUC value predicted by combined indicators (age, gender, pathological type, CEA, ALP) (modeling group = 0.792, validation group = 0.629) was higher than that predicted by single indicator. **Conclusion** The prediction model constructed in this study has good effect and can provide reference for clinical screening of high-risk patients with bone metastasis of NSCLC.

Key words NSCLC; bone metastasis; risk factors; nomogram