

# 胃癌外周血淋巴细胞数关联因素的横断面研究

汪圣毅<sup>1</sup>, 周浩<sup>2</sup>, 刘虎<sup>3</sup>

**摘要** 目的 识别胃癌患者外周血淋巴细胞数(PBLC)的关联因素。方法 横断面设计,收集行胃癌手术的患者资料,用单因素分析、多元线性回归、变量重要性评价,分析术前PBLC变化的关联因素。局部加权回归和稳健线性模型进一步验证。结果 术前PBLC  $< 1.1 \times 10^9/L$  (A组)138例(20.72%),PBLC  $\geq 1.1 \times 10^9/L$  (B组)528例(79.28%)。相对于B组,A组的年龄较大[(64.61  $\pm$  10.42)岁 vs (62.18  $\pm$  10.41)岁,  $P < 0.05$ ],中性粒细胞较低[(3.21  $\pm$  1.41)  $\times 10^9/L$  vs (3.59  $\pm$  1.31)  $\times 10^9/L$ ,  $P < 0.01$ ]。淋巴细胞减少与较高的胃癌分期有关,  $P < 0.01$ 。多元线性回归分析显示模型残差随机分布,年龄( $\beta = -0.01$ ,  $t = -3.70$ ,  $P < 0.01$ )、肿瘤分期[ $\beta$  (II vs I) = -0.16,  $t = -2.79$ ,  $P < 0.01$ ;  $\beta$  (III vs I) = -0.18,  $t = -3.87$ ,  $P < 0.01$ ;  $\beta$  (IV vs I) = -0.21,  $t = -2.16$ ,  $P < 0.05$ ]是淋巴细胞减少的关联因素,中性粒细胞增加与PBLC升高有关( $\beta = 0.05$ ,  $t = 3.61$ ,  $P < 0.01$ )。连续自变量的相对重要性分析显示,年龄、中性粒细胞、癌胚抗原(CEA)的LMG指标分别为55.55%、44.14%、0.31%。局部加权回归和稳健线性模型显示,年龄是PBLC的负向关联因素。结论 胃癌PBLC与中性粒细胞正向关联。

**关键词** 胃肿瘤;外周血淋巴细胞数;多元线性回归;横断面研究;局部加权回归;稳健线性模型;TNM分期

**中图分类号** R 735.2; R 730.3

**文献标志码** A 文章编号 1000-1492(2023)01-0151-05  
doi:10.19405/j.cnki.issn1000-1492.2023.01.025

淋巴细胞在维持机体免疫功能方面发挥重要作用<sup>[1]</sup>。外周血淋巴细胞数(peripheral blood lymphocyte count, PBLC)增加多见于血液系统肿瘤,减少多见于流感、艾滋病、新型冠状病毒感染、肿瘤<sup>[2]</sup>患

者。胃癌细胞侵犯肌层及其以上的患者,PBLC存在降低现象<sup>[3]</sup>,老年胃癌患者下降明显<sup>[4]</sup>。胃癌患者PBLC减少与年龄、病理特征、肿瘤标志物<sup>[5]</sup>存在关联关系。但既往研究多局限于简单对比,未排除混杂因素,内部真实性受限,迫切需要采用校正混杂因素的分析方法,深入揭示胃癌患者PBLC的影响因素。由此采用横断面设计,收集术后病理证实为胃癌的登记数据,用单因素分析、多元线性回归分析、变量重要性评价、局部加权回归和稳健线性模型方法,探索胃癌患者术前PBLC的影响因素,为相关研究提供参考。

## 1 资料与方法

**1.1 病例资料** 以2015—2020年行胃癌手术693例患者的横断面数据为研究对象,收集术前的年龄、性别、PBLC、中性粒细胞、癌胚抗原(carcinoembryonic antigen, CEA)、术后的病理分期(TNM分期)、分化等。入选标准:①术后病理为胃癌;②术前未行放疗或中医中药治疗。排除标准:①无病理诊断;②响应变量和自变量存在缺失值;③含异常值的个案。

**1.2 变量及哑变量设置** 术前PBLC、中性粒细胞、CEA抽取空腹静脉血检验。PBLC按照正常值范围的下限分为 $< 1.1 \times 10^9/L$ 组(A)、 $\geq 1.1 \times 10^9/L$ 组(B)。病理按美国癌症联合会(American Joint Committee on Cancer, AJCC)的标准,分为pTNM的I、II、III、IV期,分化:G<sub>x</sub>为未评估,G<sub>1</sub>为高分化,G<sub>2</sub>为中分化,G<sub>3</sub>为分化不良,G<sub>x</sub>、G<sub>1</sub>合并,低分化和未分化纳入G<sub>3</sub>。哑变量设置:①变量因子化:factor、as.factor函数将变量转为因子型。②变量排序数字化:ifelse函数设女=0,男=1,未评估和高分化、中分化、低分化和未分化腺癌分别为0、1、2,I、II、III、IV期分别为0、1、2、3。③lm函数依据排序自动设置哑变量,contrasts或model.matrix函数显示对比编码:①性别:男1、女0;②分化:未评估和高分化腺癌=0,中分化=1或低分化和未分化=1时,其余均为0;③分期:I期=0,II、III、IV期分别为1时,其余为0。

2022-08-24 接收

基金项目:国家自然科学基金(编号:81472750)

作者单位:<sup>1</sup>安徽医科大学第一附属医院普外科,合肥 230022

<sup>2</sup>中国人民解放军联勤保障部队第901医院外科,合肥 230031

<sup>3</sup>中国科学技术大学附属第一医院西区(安徽省肿瘤医院)肿瘤内科,合肥 230031

作者简介:汪圣毅,男,主任医师;

刘虎,男,副教授,主任医师,硕士生导师,责任作者,E-mail: drliuhu@yahoo.com

**1.3 统计学处理** R 3.6.2 行数据分析, ①缺失值: ! complete. case 函数识别, with 函数定位, 原始数据集分割为有和无缺失值的亚集。②异常值 (outlier): PBLC 为响应变量, 多变量模型检测异常值, 代入 cooks. distance 函数计算 Cook 距离, 其均数 4 倍以上为真实异常值 (true outlier) 并排除。正态分布用 K-S (kolmogorov-smirnov) 检验, 正态、非正态变量分别用  $\bar{x} \pm s$ 、 $M(P_{25}, P_{75})$  表示, 组间比较用  $t$  检验、Mann-Whitney U 检验。分类变量 [  $n(\%)$  ] 用  $\chi^2$  检验。多元线性回归 (multiple linear regression, MLR) 用 lm 函数, step 函数筛选自变量, 残差诊断图及 car 包 vif 函数计算方差膨胀因子 (variance inflation factor, VIF), relaimpo 包评估变量重要性<sup>[6]</sup>, ggplot2 的 qplot 函数拟合局部加权回归 (locally weighted regression, loess) 和稳健线性模型 (robust linear model, rlm)。  $P < 0.05$  为差异有统计学意义。

**2 结果**

**2.1 一般特征** 693 例中删除 CEA 缺失的 3 例、有异常值和较强影响点的 24 例, 666 例纳入, 男 504 例, 女 162 例, 年龄 23 ~ 85 (62.68 ± 10.45) 岁。I、II、III、IV 期分别为 155、130、352、29 例, 未评估和高分化、中分化、低分化和未分化腺癌分别为 78、307、281 例。与 B 组比较, A 组患者年龄较大、中性粒细胞较少、I 期少而 III 期较多, 见表 1。

**2.2 淋巴细胞数关联因素的 MLR 结果** PBLC 为响应变量 Y, 服从正态分布 ( $P = 0.15$ )。年龄、中性

粒细胞、CEA、性别、肿瘤分化和分期为自变量, step 函数优化模型的方程是  $Y = 1.921 - 0.007(\text{age}) + 0.051(\text{neutrophil}) - 0.162(\text{TNM\_II}) - 0.182(\text{TNM\_III}) - 0.214(\text{TNM\_IV})$ , 最小  $AIC = -950.81$ ,  $F = 8.96$ ,  $R^2 = 0.6358$ ,  $P < 0.01$ 。年龄每增加 1 岁, PBLC 平均减少  $7 \times 10^6/L$ , 中性粒细胞每增加 1 个单位 ( $1 \times 10^9/L$ ), 胃癌患者的 PBLC 平均增加  $5.1 \times 10^7/L$ 。调整其他变量, 与 TNM 的 I 期比较, II 期患者的 PBLC 平均减少  $1.62 \times 10^8/L$ , III 期患者的 PBLC 减少  $1.82 \times 10^8/L$ , IV 期患者的 PBLC 减少  $2.14 \times 10^8/L$ , 见表 2。

**2.3 优化模型的评价结果** step 函数筛选的年龄、中性粒细胞、TNM 分期作为自变量, 构建包含分类自变量的多元线性回归的优化模型, 对模型进行残差诊断的可视化分析, 模型残差分布基本随机, 见图 1。年龄、中性粒细胞、TNM 分期的 VIF 值分别为 1.006、1.014、1.018。

**2.4 自变量重要性及关联关系的考察** 连续自变量年龄、中性粒细胞数、CEA 纳入线性回归模型, relaimpo 包分析各变量的相对重要性统计值, 林德曼、玛琳达和金 (Lindeman, Merenda and Gold, LMG) 氏法, 最后纳入计算变量贡献 (variables contribution when included last, Last) 法、平方标准系数 (squared standardized coefficient, Betasq) 法、普拉特 (Pratt) 法、高尼兹 (Genizi) 法、CAR 评分 (CAR score, Car) 法的值分别为: 年龄 = 55.55%、55.67%、55.67%、55.56%、55.54%、55.55%, 中性粒细胞 = 44.14%、

表 1 胃癌患者的基本特征 [  $M(P_{25}, P_{75}), n(\%), \bar{x} \pm s$  ]

变量	总体 (666 例)	淋巴细胞 ( $\times 10^9/L$ )		统计值	P 值
		< 1.1 (A 组) (138 例)	$\geq 1.1$ (B 组) (528 例)		
年龄 (岁)	62.68 ± 10.45	64.61 ± 10.42	62.18 ± 10.41	-2.44	0.02
中性粒细胞 ( $\times 10^9/L$ )	3.51 ± 1.34	3.21 ± 1.41	3.59 ± 1.31	2.85	0
淋巴细胞 ( $\times 10^9/L$ )	1.54 ± 0.50	0.90 ± 0.15	1.71 ± 0.42	37.13	0
CEA (ng/ml)	2.34 (1.28, 4.78)	2.55 (1.40, 4.70)	2.28 (1.28, 4.80)	35 087.00	0.50
性别				0.04	0.84
女	162 (24.32)	35 (25.36)	127 (24.05)		
男	504 (75.68)	103 (74.64)	401 (75.95)		
肿瘤分化				3.50	0.17
Gx + G1	78 (11.71)	10 (7.25)	68 (12.88)		
G2	307 (46.10)	65 (47.10)	242 (45.83)		
G3	281 (42.19)	63 (45.65)	218 (41.29)		
TNM 分期				12.90	0.01
I	155 (23.27)	17 (12.32)	138 (26.14)		
II	130 (19.52)	27 (19.57)	103 (19.51)		
III	352 (52.85)	88 (63.77)	264 (50.00)		
IV	29 (4.35)	6 (4.35)	23 (4.36)		

统计值包括  $t$  值、 $\chi^2$  值、 $W$  值

表 2 胃癌患者淋巴细胞数关联因素的多元线性回归分析

截距及自变量	$\beta$ (95%可信区间)	标准误	t 值	P 值
截距	1.921(1.667, 2.175)	0.129	14.858	<0.01
年龄	-0.007(-0.010, -0.003)	0.002	-3.698	<0.01
中性粒细胞	0.051(0.023, 0.079)	0.014	3.607	<0.01
TNM_II	-0.162(-0.276, -0.048)	0.058	-2.789	<0.01
TNM_III	-0.182(-0.275, -0.090)	0.047	-3.865	<0.01
TNM_IV	-0.214(-0.408, -0.019)	0.099	-2.159	<0.05

TNM\_II:TNM II期(为1的哑变量) vs TNM I期;TNM\_III:TNM III期 vs TNM I期;TNM\_IV:TNM IV期 vs TNM I期

44.07%、44.06%、44.14%、44.14%、44.14%，CEA = 0.31%、0.27%、0.27%、0.30%、0.32%、0.31%， $R^2$  贡献最大的自变量为年龄，见图 2。拟合考察分析中，qplot 函数绘制淋巴细胞数与年龄的关系，lo-ess 拟合时 (span = 0.2)，曲线波动，但总体上呈负关联 (残差标准误 = 0.4967)，见图 3，rlm 拟合时，负性关联关系依然存在，见图 4 (残差标准误 = 0.5082)。

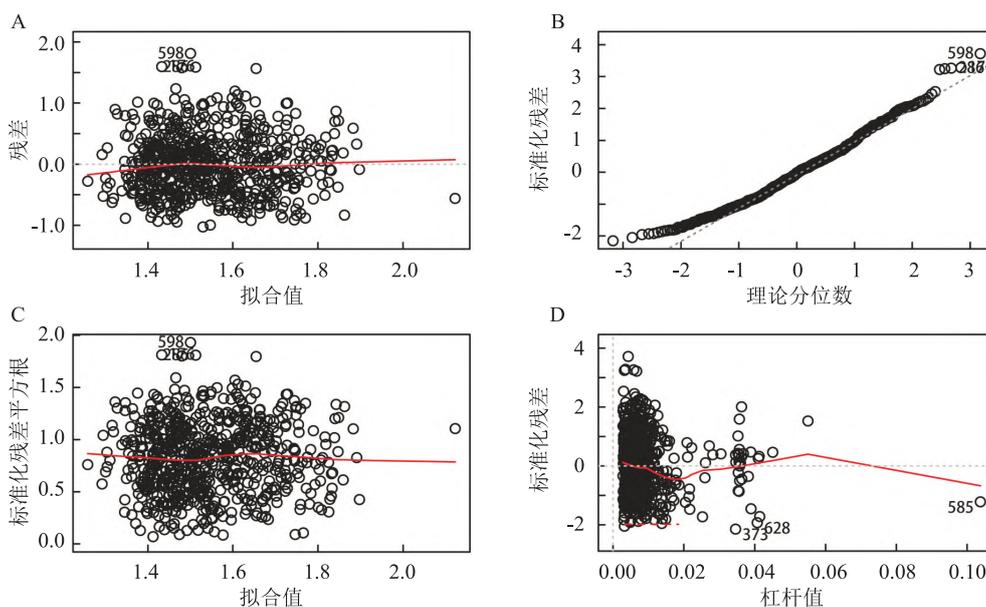


图 1 模型残差分析的诊断图

A:残差对拟合值;B:正态 QQ 图;C:标准化残差平方根对拟合值;D:标准化残差对杠杆值

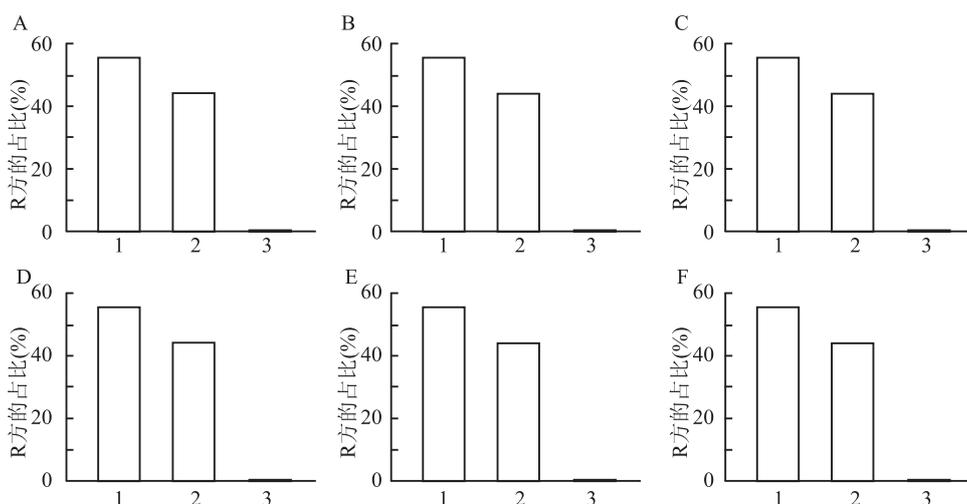


图 2 线性模型相对重要性指标

A:LMG 方法;B:Last 方法;C:Betassq 方法;D:Pratt 方法;E:Genizi 方法;F:Car 方法;1:年龄(岁);2:中性粒细胞( $\times 10^9/L$ );3:癌胚抗原(ng/ml)

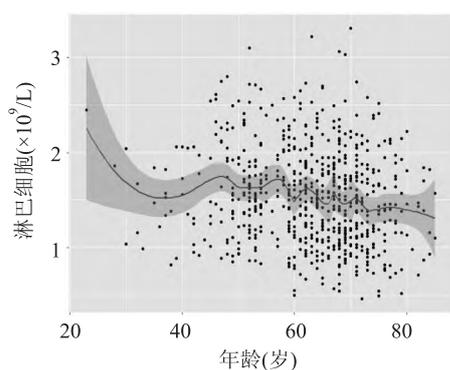


图3 Loess 平滑时年龄与淋巴细胞数的关系

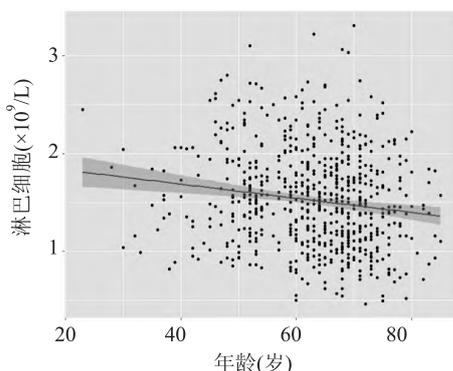


图4 稳健线性模型拟合时年龄与淋巴细胞数的关系

### 3 讨论

胃癌浸润深度超过黏膜下层(>T1期)的患者,PBLC 减低<sup>[3]</sup>,老年胃癌患者的 PBLC 与生存时间密切相关,其变化具有重要的临床意义<sup>[7]</sup>,但其影响因素尚不明确。本研究显示,随着年龄的增长、胃癌 TNM 分期的增加,PBLC 减少,中性粒细胞与 PBLC 存在正向关联。

本研究显示,年龄每增加 1 岁,胃癌患者每毫升外周血中的淋巴细胞数减少 7 000 个。Tatara et al<sup>[7]</sup>显示,≥75 岁的胃癌,47% 的患者 PBLC 低于 1 462/μl。Saito et al<sup>[4]</sup>显示,≥70 岁胃癌患者的术前 PBLC 显著低于 <70 岁的患者。Park et al<sup>[8]</sup>以 I~III 期胃癌为研究对象,发现术前 PBLC 减少(<1 000/μl)的患者达 24.8%,最低者为 640/μl,年龄增加是 PBLC 减少的独立影响因素,与本研究的结果一致。loess 方法局部拟合,分析全部数据范围中两个变量间的关系,结果更真实可靠<sup>[9]</sup>,rlm 可规避离群值的影响,拟合线与真实线更接近,结果更可靠<sup>[10]</sup>。本研究首次用 loess 和 rlm 拟合方法,显示胃癌患者 PBLC 与年龄的负向关联关系,表明二者的关系较为稳定。胃癌患者 PBLC 减少与年龄增加有关的机制尚不明确。淋巴细胞功能和数量变化与衰老存在共同的调节基因<sup>[11]</sup>,因此,衰老有关的通路

及其分子改变,可能会抑制淋巴细胞活化、促进凋亡或焦亡使其数量减少,是有待研究的领域。

胃癌患者外周血淋巴细胞亚群与分期有关,分期较晚时,CD3<sup>+</sup>、CD3<sup>+</sup>CD4<sup>+</sup>、CD4<sup>+</sup>/CD8<sup>+</sup> 淋巴细胞减少<sup>[12]</sup>,诱导免疫耐受和抑制的调节性 T 细胞增加<sup>[13]</sup>,但淋巴细胞总数与 TNM 分期的关系尚未见文献报告。本研究显示胃癌患者外周血淋巴细胞总数与 TNM 分期负相关,与 I 期比较,II、III、IV 期胃癌患者的淋巴细胞减少显著。因此,晚期胃癌患者可能存在抑制淋巴细胞增殖的机制。Fu et al<sup>[14]</sup>对胃癌的免疫细胞进行 RNA 测序发现,分期越晚的患者,血液 CD8<sup>+</sup> T 细胞中,转录因子 IRF8 的表达越低,而 IRF8 可调控 Zbtb32,促进 NK 细胞的增殖<sup>[15]</sup>。IRF8、Zbtb32 是否与晚期胃癌淋巴细胞数减少有关,是未来的研究方向之一。

本研究显示 PBLC 与中性粒细胞存在的正向关联关系,与 Saito et al<sup>[3]</sup>的结果一致,另外,PBLC 较高组(≥1.1×10<sup>9</sup>/L)的 CEA 低于 PBLC 较低组(1.1×10<sup>9</sup>/L)的 CEA,但单因素和多元回归分析,CEA 对淋巴细胞的影响不显著。直接用回归系数比较变量重要性可能导致错误结论,一般采用相对重要性指标,relaimpo 包提供多种不同的指标,评估线性模型中变量的重要性,可增加研究的可靠性,是此类研究的优势。本研究 relaimpo 不同算法的结果均显示,年龄对 PBLC 的影响较为重要,表明年龄对 PBLC 的作用较为稳定。

MLR 是基于普通最小二乘法的统计分析方法,存在一定的局限性,未来可行条件宽泛的分位数回归、贝叶斯回归、神经网络模型等分析,深入了解胃癌 PBLC 的影响因素及其变化规律,为临床诊断、治疗和预后分析提供参考。

### 参考文献

- [1] Wang F, Nie J, Wang H, et al. Characteristics of peripheral lymphocyte subset alteration in covid-19 pneumonia [J]. J Infect Dis, 2020, 221(11):1762-9.
- [2] Xu H, Lin M, Hu Y, et al. Lymphopenia during definitive chemoradiotherapy in esophageal squamous cell carcinoma: Association with dosimetric parameters and patient outcomes [J]. Oncologist, 2021, 26(3): e425-34.
- [3] Saito H, Shimizu S, Kono Y, et al. Score of the preoperative absolute number of lymphocytes, monocytes, and neutrophils as a prognostic indicator for patients with gastric cancer [J]. Surg Today, 2019, 49(10): 850-8.
- [4] Saito H, Kono Y, Murakami Y, et al. Prognostic significance of pre- and postoperative lymphocyte counts in patients with gastric

- cancer [J]. Dig Surg, 2019, 36(2): 137–43.
- [5] 刘红芬, 王征, 王莉, 等. 肿瘤标记物对胃癌诊断和预后判断的价值以及与免疫炎症反应指标的相关性 [J]. 河北医科大学学报, 2019, 40(7): 855–8.
- [6] Srinivasa R D, Kannipamula S J. Assessing the relative importance of predictors in linear regression [J]. EJCMC, 2020, 7(4): 970–6.
- [7] Tataru T, Suzuki S, Kanaji S, et al. Lymphopenia predicts poor prognosis in older gastric cancer patients after curative gastrectomy [J]. Geriatr Gerontol Int, 2019, 19(12): 1215–9.
- [8] Park S J, Lee J, Kim H, et al. Association between absolute lymphocyte count and overall mortality in patients with surgically resected gastric cancer [J]. Korean J Intern Med, 2021, 36(3): 679–88.
- [9] Hoffman H, Bunch K M, Mikhailova T, et al. Transition from proximal to distal radial access for diagnostic cerebral angiography: Learning curve analysis [J]. World Neurosurg, 2021, 152: e484–91.
- [10] Yang C X, Schon E, Obeidat M, et al. Occurrence of accelerated epigenetic aging and methylation disruptions in human immunodeficiency virus infection before antiretroviral therapy [J]. J Infect Dis, 2021, 223(10): 1681–9.
- [11] Walkiewicz D, Wicik Z, Puzianowska-Kuznicka M. Gonadotropin-releasing hormone receptor pathway affects the function of human EBV-transformed B lymphocytes in an age-independent way [J]. Exp Gerontol, 2021, 152:111471.
- [12] 宋谦, 党海珍, 李露嘉, 等. 胃癌患者不同分期外周血 CD3<sup>+</sup>、CD4<sup>+</sup> 和 CD8<sup>+</sup> 的表达结果比较 [J]. 实用医药杂志, 2012, 29(7): 587–9.
- [13] 陆威, 李永翔, 张尚鑫, 等. 胃癌患者外周血及癌组织中 CD4<sup>+</sup> CD25<sup>+</sup> 调节性 T 细胞、转录因子 Foxp3 的表达及临床意义 [J]. 安徽医科大学学报, 2013, 48(7): 790–3.
- [14] Fu K, Hui B, Wang Q, et al. Single-cell RNA sequencing of immune cells in gastric cancer patients [J]. Aging (Albany NY), 2020, 12(3): 2747–63.
- [15] Adams N M, Lau C M, Fan X, et al. Transcription factor IRF8 orchestrates the adaptive natural killer cell response [J]. Immunity, 2018, 48(6):1172–82.

## Associated factors of peripheral blood lymphocyte counts in gastric cancer: a cross-sectional study

Wang Shengyi<sup>1</sup>, Zhou Hao<sup>2</sup>, Liu Hu<sup>3</sup>

(<sup>1</sup>Dept of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei 230022;

<sup>2</sup>Dept of Surgery, 901th Hospital of the Chinese People's Liberation Army Joint Logistics

Support Force, Hefei 230031; <sup>3</sup>Dept of Oncology, West District of The First Affiliated Hospital of University of Science and Technology of China, Hefei 230031)

**Abstract Objective** To identify the associated factors of peripheral blood lymphocyte counts (PBLC) in gastric cancer patients. **Methods** In this cross-sectional study, data were collected from patients underwent gastric cancer surgery. Univariate analysis, multiple linear regression and variable importance metrics were used to assess factors associating PBLC in pre-operative patients. The findings were further investigated by fitting locally weighted regression (loess) and robust linear model (rlm). **Results** Patients with pre-operative PBLC  $< 1.1 \times 10^9/L$  (group A) were 138 (20.72%), whereas patients with PBLC  $\geq 1.1 \times 10^9/L$  (group B) were 528 (79.28%). Compared to group B, patients in group A were older [(64.61  $\pm$  10.42 years) vs (62.18  $\pm$  10.41 years),  $P < 0.05$ ] and showed lower neutrophils [(3.21  $\pm$  1.41)  $\times 10^9/L$  vs (3.59  $\pm$  1.31)  $\times 10^9/L$ ,  $P < 0.01$ ]. Lymphopenia was associated with later stage of gastric cancer,  $P < 0.01$ . A multiple linear regression model with randomly distributed residuals found that age ( $\beta = -0.01$ ,  $t = -3.70$ ,  $P < 0.01$ ) and tumor stage [ $\beta$  (II vs I) =  $-0.16$ ,  $t = -2.79$ ,  $P < 0.01$ ;  $\beta$  (III vs I) =  $-0.18$ ,  $t = -3.86$ ,  $P < 0.01$ ;  $\beta$  (IV vs I) =  $-0.21$ ,  $t = -2.16$ ,  $P < 0.05$ ] were to be statistically significant factors for lymphopenia. Neutrophil increase ( $\beta = 0.05$ ,  $t = 3.61$ ,  $P < 0.01$ ) was associated with a higher PBLC. The relative importance analysis for continuous independent variables showed that relative importance metric LMGs in variable age, neutrophil and CEA were 55.55%, 44.14% and 0.31% respectively. Loess and rlm models showed that age was the negatively influencing factor of PBLC. **Conclusion** PBLC in gastric cancer patients were positively associated with neutrophils, and negatively associated with age and tumor stages.

**Key words** stomach neoplasms; peripheral blood lymphocyte count; multiple linear regression; cross-sectional study; locally weighted regression; robust linear model; TNM staging