

论著·临床研究 doi:10.3969/j.issn.1671-8348.2017.08.010

Logistic 回归和 PLS-DA 模型评价肿瘤标志物对结肠癌的诊断价值*

曾龙飞¹, 田刚², 姚健^{1,3△}

(1. 西南医科大学附属医院肝胆外科, 四川泸州 646000; 2. 西南医科大学附属医院检验科, 四川泸州 646000; 3. 四川省泸州市人民医院肝胆外科 646000)

[摘要] **目的** 探讨血清肿瘤 M2 型丙酮酸激酶(TuM2-PK)、癌胚抗原(CEA)、糖类抗原 CA199、CA724、CA125 和 CA242 对结肠癌的诊断价值。**方法** 采用酶联免疫吸附法(ELISA)和电化学发光法测定 231 例结肠癌、105 例结肠良性病和 166 例健康对照血清 TuM2-PK、CEA、CA199、CA724、CA125 和 CA242 水平。采用 ROC 曲线和 Logistic-ROC 曲线比较单一和组合标志物的诊断效能,以最优组合标志物建立结肠癌 Logistic 回归和偏最小二乘判别分析(PLS-DA)诊断模型。**结果** 结肠癌组血清 TuM2-PK、CEA、CA199、CA125 和 CA242 水平显著高于结肠良性病组和健康对照组($P < 0.05$)。血清 CA199 诊断结肠癌的 ROC 曲线下面积(AUC)最高[0.79 (95%CI:0.75~0.83)],在临界值为 69.5 U/L 时灵敏度和特异性分别为 64.1%和 89.7%。Logistic-ROC 曲线中各组合标志物较单一指标的 AUC 高,其中(CEA+CA199+TuM2-PK)的 AUC 较为理想[0.89(95%CI:0.86~0.92)]。血清(CEA+CA199+TuM2-PK)的 Logistic 回归和 PLS-DA 模型对结肠癌的诊断准确率分别为 82.7%和 77.5%,对非结肠癌的诊断准确率分别为 93.7%和 95.6%。**结论** 联合检测血清 CEA、CA199 和 TuM2-PK 是较为理想的肿瘤标志物组合,有助于结肠癌的早期诊断。

[关键词] 结肠肿瘤;肿瘤 M2 型丙酮酸激酶;ROC 曲线;Logistic 回归分析**[中图分类号]** R735.3**[文献标识码]** A**[文章编号]** 1671-8348(2017)08-1038-04

Diagnostic of serum tumor markers for colon cancer assessed with logistic regression analysis and PLS-DA model*

Zeng Longfei¹, Tian Gang², Yao Jian^{1,3△}

(1. Department of Hepatobiliary Surgery, the Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan 646000, China; 2. Department of Clinical Laboratory, the Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan 646000, China; 3. Department of Hepatobiliary Surgery, the People's Hospital of Luzhou City, Luzhou, Sichuan 646000, China)

[Abstract] **Objective** To investigate the diagnostic value of serum tumors M2 pyruvate kinase (TuM2-PK), carcino-embryonic antigen (CEA), CA199, CA724, CA125 and CA242 in colon cancer. **Methods** Serum levels of TuM2-PK, CEA, CA199, CA724, CA125 and CA242 in 231 patients with colon cancer, 105 patients with colon benign lesion diseases and 166 normal controls were measured by the ELISA and electrochemiluminescence assays. The operation working characteristic curve (ROC) and the logistic regression-ROC were used to evaluate the diagnostic value of tumor markers for colon cancer individually and in combination. The models of logistic regression analysis and partial least-squares discriminant analysis (PLS-DA) were established to diagnose patients with colon cancer based on the optimal panel of serum tumor markers. **Results** Concentrations of serum TuM2-PK, CEA, CA199, CA125 and CA724 in the colon cancer group are higher than those in the colon benign diseases group and the normal controls ($P < 0.05$). The area under the operation working characteristic curve (AUC) of CA199 is the highest 0.79 [95% confidence interval (95%CI), 0.75-0.83] at the cutoff value of 69.5 U/L with 64.1% of sensitivity and 89.7% of specificity. The AUC of combined serum tumor markers based on logistic regression analysis is higher than those in individuals, of which serum (CEA+CA199+TuM2-PK) is the optimal [AUC=0.89, 95% confidence interval (95%CI), 0.86-0.92]. The diagnostic accuracy of logistic analysis and PLS-DA model for colon cancer is 82.7% and 77.5%, and for colon cancer is 93.7% and 95.6%, respectively. **Conclusion** The combination of serum tumor markers CEA, CA199 and TuM2-PK is more suitable as a diagnostic model for the screening of colon cancer.

[Key words] colon cancer; tumor M2 pyruvate kinase; receiver operating characteristic curve; logistic regression analysis

结肠癌早期症状不典型,约 60% 的患者确诊时已到中晚期^[1]。目前尚无准确诊断结肠癌的单一肿瘤标志物,而癌胚抗原(carcino-embryonic antigen, CEA)、糖类抗原(carbohydrate antigen, CA)CA199、CA724 等诊断结肠癌的灵敏度也不理想^[2-3]。肿瘤 M2 型丙酮酸激酶(tumors M2 pyruvate kinase,

TuM2-PK)是近年来发现与结肠癌有关的新型标志物,其诊断价值尚未被完全阐明^[4]。研究表明,联合检测血清 TuM2-PK、CEA 和 CA199 等能提高结肠癌诊断的灵敏度和阳性率^[4-5]。尽管如此,并联实验通过叠加各指标的阳性率提高灵敏度的同时降低了特异性,增加了检测成本^[6]。串联实验提高了诊断的

表 1 血清 CEA、CA199、CA724、CA125、CA242 和 TuM2-PK 水平($\bar{x}\pm s$)

组别	n	CEA(ng/mL)	CA199(U/L)	CA724(U/L)	CA125(U/L)	CA242(U/L)	TuM2-PK(U/mL)
健康对照组	166	4.0±2.02	35.9±18.8	6.6±5.89	22.5±10.1	16.2±11.0	13.3±6.48
良性病组	105	13.2±10.3 ^a	54.3±30.3 ^a	12.7±7.46	26.1±12.7 ^a	28.1±23.6 ^a	24.4±14.0 ^a
结肠癌组	231	19.9±17.9 ^{bc}	96.0±53.4 ^{bc}	14.9±10.4 ^{bc}	32.8±17.8 ^c	34.6±28.9 ^{bc}	36.2±23.8 ^{bc}
F		75.1	137.9	51.2	31.6	58.2	91.5
P		0.000	0.000	0.000	0.000	0.000	0.000

^a: P<0.05, 与结肠良性病组比较; ^b: P<0.05, 与健康对照组比较; ^c: P<0.05, 结肠癌与健康对照组比较。

表 2 血清 CEA、CA199、CA724、CA125、CA242 和 TuM2-PK 对结肠癌的诊断价值比较

标志物	灵敏度[n(%)]	特异性[n(%)]	AUC(95%CI)	临界值
CEA(ng/mL)	152(65.8)	202(74.5)	0.70(0.66~0.74)	8.37
CA199(U/L)	148(64.1)	243(89.7)	0.79(0.75~0.83)	69.50
CA724(U/L)	187(61.0)	150(85.4)	0.72(0.66~0.77)	5.08
CA125(U/L)	93(40.3)	230(84.9)	0.65(0.61~0.71)	36.90
CA242(U/L)	106(45.9)	222(81.9)	0.68(0.63~0.73)	33.10
TuM2-PK(U/mL)	144(62.3)	221(81.5)	0.73(0.68~0.78)	25.20

特异性但牺牲了灵敏度。因此,寻找最佳的肿瘤标志物组合并评价其联合诊断效能显得尤为重要。受试者工作特征(ROC)曲线是一种兼顾灵敏度和特异性的诊断试验评价工具,而基于 Logistic 回归的 ROC 曲线可用于联合指标的诊断效能评价^[6]。因此,本研究以结肠癌患者为研究对象,采用 ROC 曲线和 Logistic 回归-ROC 曲线探讨血清 TuM2-PK、CEA、CA199、CA724、CA125 和 CA242 单项和经 Logistic 回归后组合标志物对结肠癌的诊断效能。进一步建立最优肿瘤标志物的二分类 Logistic 回归分析和偏最小二乘判别分析(PLS-DA)线性模型,验证其对结肠癌的鉴别诊断价值。

1 资料与方法

1.1 一般资料 选取 2013 年 5 月至 2016 年 1 月就诊于西南医科大学附属第一医院胃肠外科的结肠癌患者 231 例(结肠癌组),其中男 139 例,女 92 例,平均年龄(64.4±10.9)岁。结肠良性病患者 105 例(良性病组),男 76 例,女 29 例,平均年龄(52.5±19.7)岁。其中溃疡性结肠炎 37 例,结肠息肉 30 例,急、慢性结肠炎 28 例,肠易激综合征 10 例。健康对照组为本院体检者,共 166 例,其中男 96 例,女 70 例,平均年龄(49.1±26.6)岁,各项体检指标均正常,无心、肺、肾脏疾病等。

1.2 方法

1.2.1 仪器与试剂 血清 CEA、CA199、CA724、CA125 和 CA242 测定均采用深圳新产业(Maglumi 2000 Plus)电化学发光仪及配套试剂,质控均符合要求。血清 TuM2-PK 的测定采用 ELISA,采用苏州卡尔文公司提供的试剂盒并严格按试剂说明书操作,批内和批间变异系数(CV)均控制在 10%以内。

1.2.2 测定方法 采用电化学发光法测定血清中 CEA、CA199、CA724、CA125 和 CA242 水平。采用酶联免疫吸附试验(ELISA)检测血清 TuM2-PK 水平。各指标的参考范围为:CEA 0~6.0 ng/mL、CA199 0~37 U/L、CA724 0.2~6.0 U/L、CA125 0~35 U/L、CA242 0~15 U/L 和 TuM2-PK 0~15 U/mL。

1.3 统计学处理 ROC 曲线分析和 Logistic 回归分析均采用 SPSS 17.0 软件完成。PLS-DA 模型的建立采用 SIMCA

13.0 软件完成。计量资料以 $\bar{x}\pm s$ 表示,组内比较采用单因素方差分析,组间比较采用 LSD 检验, P<0.05 表示差异有统计学意义。

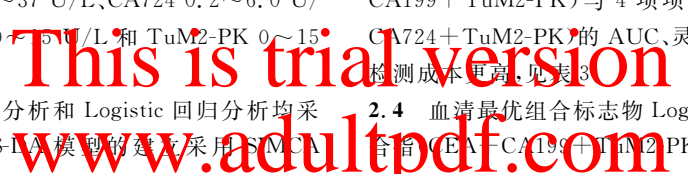
2 结果

2.1 血清 CEA、CA199、CA724、CA125、CA242 和 TuM2-PK 水平比较 结肠癌组血清中 CEA、CA199、CA125、CA242 和 TuM2-PK 水平均显著高于良性病组和健康对照组,差异有统计学意义(P<0.05)。良性病组血清 CEA、CA199、CA724、CA242 和 TuM2-PK 也显著高于对照组,差异有统计学意义(P<0.05),见表 1。

2.2 血清 CEA、CA199、CA724、CA125、CA242 和 TuM2-PK 对结肠癌的诊断价值比较 6 种血清肿瘤标志物中 CEA、CA199、CA724 和 TuM2-PK 的诊断效能较高(AUC≥0.70)。其中 CA199 具有最高的诊断效能(AUC=0.79),在临界值为 69.5 U/L 时,具有 64.1%的灵敏度和 89.7%的特异性,见表 2。

2.3 基于 Logistic 回归-ROC 曲线的血清组合标志物对结肠癌的诊断价值比较 以诊断效能较高(AUC≥0.70)的血清 CEA、CA199、CA724 和 TuM2-PK 建立 Logistic 回归-ROC 曲线模型,评估组合标志物的诊断效能(图 1)。两项或多项组合标志物的诊断效能均优于单一指标。两项联合指标中(CA199+TuM2-PK)的诊断效能最佳,AUC、灵敏度和特异性分别为 0.87、77.5%和 90.0%。3 项联合指标中,(CEA+CA199+TuM2-PK)的诊断效能最高,AUC、灵敏度和特异性分别为 0.89、82.3%和 91.5%。4 项联合指标(CEA+CA199+CA724+TuM2-PK)的 AUC、灵敏度和特异性分别为 0.90、84.4%和 88.9%。(CEA+CA199+TuM2-PK)的 AUC 和灵敏度均优于(CA199+TuM2-PK)。3 项组合标志物(CEA+CA199+TuM2-PK)与 4 项组合标志物(CEA+CA199+CA724+TuM2-PK)的 AUC、灵敏度和特异性相当,但后者的检测成本更高,见表 3。

2.4 血清最优组合标志物 Logistic 诊断模型分析 以 3 项联合指标(CEA+CA199+TuM2-PK)建立 Logistic 诊断模型,该模



型对结肠癌和非结肠癌(良性病组和对照组)诊断的准确率分别为 82.7%和 93.7%,见图 2。

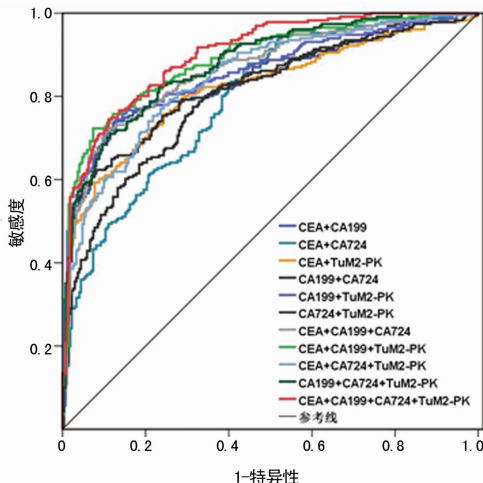


图 1 基于血清组合肿瘤标志物的 Logistic 回归-ROC 曲线

表 3 血清组合标志物的 Logistic 回归-ROC 曲线对结肠癌的灵敏度、特异性和 AUC 比较

组合标志物	灵敏度 (%)	特异性 (%)	AUC (95%CI)
CEA+CA199	71.4	88.6	0.84(0.81~0.89)
CEA+CA724	61.5	79.0	0.80(0.76~0.83)
CEA+ TuM2-PK	68.4	82.3	0.82(0.78~0.86)
CA199+CA724	65.4	87.6	0.83(0.79~0.87)
CA199+TuM2-PK	77.5	90.0	0.86(0.82~0.88)
CA724+TuM2-PK	64.1	80.8	0.80(0.76~0.84)
CEA+CA199+CA724	71.4	89.7	0.87(0.83~0.90)
CEA+CA199+TuM2-PK	82.3	91.5	0.89(0.86~0.92)
CEA+CA724+TuM2-PK	68.0	82.3	0.85(0.81~0.88)
CA199+CA724+TuM2-PK	71.9	87.1	0.88(0.85~0.91)
CEA+CA199+CA724+TuM2-PK	84.4	88.9	0.90(0.88~0.93)

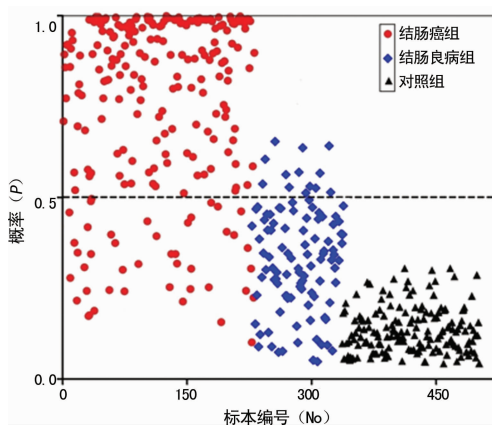
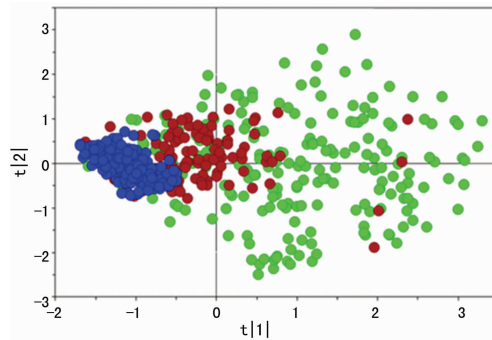


图 2 基于血清 CEA、CA199 和 TuM2-PK 的 Logistic 回归诊断模型

2.5 血清最优组合标志物 PLS-DA 诊断模型的建立 以 3 项联合指(CEA+CA199+TuM2-PK)建立 PLS-DA 诊断模型,

该模型对结肠癌和非结肠癌(良性病组和健康对照组)诊断的准确率分别为 77.5%(179/231)和 95.6%(259/271),见图 3。



蓝色:健康对照组;红色:结肠良性病组;绿色:结肠癌组。
图 3 基于血清 CEA、CA199 和 TuM2-PK 的 PLS-DA 诊断模型

3 讨论

CEA、CA199、CA724、CA125 和 CA242 是肿瘤常规筛查项目,其中 CA199 是结肠癌最常用的筛查指标,在消化道肿瘤中其血清水平显著升高^[2-4]。古丽等^[7]研究表明,结肠癌组血清 CA199 水平显著高于对照组,诊断结肠癌的 AUC、灵敏度和特异性分别为 0.76、53.1%和 89.9%。本研究发现,6 种肿瘤标志物中 CA199 的 AUC 最高,诊断效能与特异性与古丽等^[7]报道基本一致。

丙酮酸激酶(PK)是参与糖酵解的关键酶,主要有 L-PK、R-PK、M1-PK 和 M2-PK 4 种亚型。M2-PK 有二聚体和四聚体两种形式,其中活性较低的二聚体主要参与碳水化合物的合成,活性较高的四聚体主要参与糖无氧酵解的能量代谢^[8]。肿瘤细胞发生时常伴有 M2-PK 四聚体水平的减少和 M2-PK 二聚体含量的增加,故 M2-PK 又称为 TuM2-PK。由于肿瘤细胞中 TuM2-PK 与磷酸烯醇式丙酮酸的亲和力远低于四聚体丙酮酸激酶,故糖酵解过程受阻,中间产物大量堆积、肿瘤细胞快速增殖^[9]。Fatel-Cantillo 等^[10]证实,TuM2-PK 在结肠癌细胞增殖和转移时释放入血,导致血清 TuM2-PK 水平显著升高。本研究发现,结肠癌组血清 TuM2-PK 水平显著高于良性病组,可能与结肠癌患者肿瘤细胞的糖酵解代谢异常有关。结肠癌组血清 TuM2-PK 水平显著高于良性病组与对照组,提示增高的血清 TuM2-PK 水平有助于结肠癌的辅助诊断。此外,单项标志物中 TuM2-PK 诊断结肠癌的灵敏度与 CEA 和 CA199 相当,高于 Schneider 等^[11]报道,可能与人群和诊断阈值不同有关。此外,本研究结果提示,联合检测 CEA 和 TuM2-PK 可显著提高结肠癌诊断的灵敏度,与 Kim 等^[12]结论一致。

目前,但鲜有 TuM2-PK 与 CEA、CA199、CA724、CA125 和 CA242 不同组合方式的 Logistic 回归-ROC 曲线对结肠癌的诊断效能比较。Molodianovitch 等^[13]发现,基于 Logistic 回归的 ROC 曲线可以更好地评价模型的拟合优度、兼顾了灵敏度和特异性。卢灿荣等^[14]基于 Logistic 回归的 ROC 曲线表明,联合检测血清 CEA、CA199 和 C 反应蛋白有助于提高结肠癌的诊断效能。Fung 等^[15]采用 Logistic 回归联合检测 TuM2-PK、Dickkopf-3 和胰岛素样生长因子结合蛋白(IGFBP2)诊断结肠癌 AUC 达到 0.87~0.91,灵敏度达到 73%。本研究对比分析发现,Logistic 回归-ROC 曲线模型中,血清(CEA+CA199+TuM2-PK)与(CEA+CA199+CA724+TuM2-PK)组合标志物的 AUC 相当,均能较好地提高结肠癌诊断的灵敏度但后者的检测成本较高。进一步以组合标志物

(CEA+CA199+TuM2-PK)建立 Logistic 回归和 PLS-DA 诊断模型,两类模型对结肠癌和非结肠癌的诊断准确率也均较为理想。

综上所述,血清(CEA+CA199+TuM2-PK)是鉴别诊断结肠癌较为理想的肿瘤标志物组合,有助于结肠癌的早期诊断。尽管如此,本研究尚未全面考察 TuM2-PK 各类常规肿瘤标志物联合检测对结肠的诊断价值,标本量也较为有限。因此,寻找更加理想的肿瘤标志物组合,评价其诊断结肠癌的效能、分析组合标志物与病理金标准结果的一致性还需大样本深入研究。

参考文献

- [1] 吴在德. 外科学[M]. 6 版. 北京:人民卫生出版社,2004:510-514.
- [2] 王健,叶林加,李万海,等. 联合检测血清 CEA、CA19-9 与 CA72-4 对老年结肠癌诊断的价值[J]. 中国老年学杂志,2014,10(19):5452-5453.
- [3] 荆卫娟,任传路,丁磊,等. 血清 TuM2-PK、TSGF 与 CEA、CA19-9、CA242 联合检测在结肠癌诊断中的临床价值[J]. 标记免疫分析与临床,2015,22(11):1121-1124.
- [4] 顾群浩,张晓东,蔡照弟,等. 肿瘤型 M2-PK、CEA 与 CA199 检测在结直肠癌诊断中应用的临床观察[J]. 现代生物医学进展,2012,12(31):6083-6085,6075.
- [5] 李冬梅,赵秀媛,朱传金,等. M2-PK 与 CEA、CA19-9、CA72-4 对消化系统肿瘤诊断的比较分析[J]. 临床消化病杂志,2007,19(1):53-55.
- [6] 刘靳波,余晓林,夏乾峰,等. ROC 曲线及 Logistic 回归评价肿瘤标志物在胃结肠肿瘤的诊断价值[J]. 第四军医大学学报,2008,29(14):1298-1300.
- [7] 古丽. 结肠癌患者 CK19 的表达及意义[J]. 重庆医学,2015,44(29):4120-4122.
- [8] 周莹艳. 肿瘤 M2 型丙酮酸激酶在肺癌中的研究进展[J]. 临床肺科杂志,2010,15(8):1139-1141.
- [9] Christofk HR, Vander Heiden MG, Harris MH, et al. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth[J]. Nature, 2008, 452(7184):230-233.
- [10] Fatela-Cantillo D, Fernandez-Suarez A, Marin Moreno MA, et al. Prognostic value of plasmatic tumor M2 pyruvate kinase and carcinoembryonic antigen in the survival of colorectal cancer patients[J]. Tumor Biology, 2012, 33(3,SD):825-832.
- [11] Schneider J, Schulze G. Comparison of tumor M2-pyruvate kinase (tumor M2-PK), carcinoembryonic antigen (CEA), carbohydrate antigens CA 19-9 and CA 72-4 in the diagnosis of gastrointestinal cancer [J]. Anticancer Res, 2004, 23(6D):5089-5093.
- [12] Kim CW, Kim JI, Park SH, et al. Usefulness of plasma tumor M2-pyruvate kinase in the diagnosis of gastrointestinal cancer[J]. Korean J Gastroenterol, 2003, 42(5):387-393.
- [13] Molodianovitch K, Faraggi D, Reiser B. Comparing the areas under two correlated ROC curves; parametric and non-parametric approaches[J]. Biom J, 2006, 48(5):745-757.
- [14] 卢灿荣,张士武,张勇,等. 联合检测癌胚抗原、糖链抗原 19-9 和 C 反应蛋白对结肠癌的诊断价值[J]. 标记免疫分析与临床,2012,19(1):3-6.
- [15] Fung KY, Tabor B, Buckley MJ, et al. Blood-based protein biomarker panel for the detection of colorectal cancer[J]. PLoS One, 2015, 10(3):e0120425.

(收稿日期:2016-09-30 修回日期:2016-11-28)

(上接第 1037 页)

- [2] Ibrahim N, Januzzi JL. The potential role of natriuretic peptides and other biomarkers in heart failure diagnosis, prognosis and management[J]. Expert Rev Cardiovasc Ther, 2015, 13(9):1017-1030.
- [3] Wayne Causey M, Singh N. Clinical implications of B-type natriuretic peptide and N-terminal pro-B-type natriuretic peptide in the care of the vascular surgery patient[J]. Semin Vasc Surg, 2014, 27(3/4):143-147.
- [4] Don-Wauchope AC, McKelvie RS. Evidence based application of BNP/NT-proBNP testing in heart failure[J]. Clin Biochem, 2015, 48(4/5):236-246.
- [5] Sun RR, Lu L, Liu M, et al. Biomarkers and heart disease [J]. Eur Rev Med Pharmacol Sci, 2014, 18(19):2927-2935.
- [6] Santaguida PL, Don-Wauchope AC, Oremus M, et al. BNP and NT-proBNP as prognostic markers in persons with acute decompensated heart failure: a systematic review[J]. Heart Fail Rev, 2014, 19(4):534-540.
- [7] Oremus M, Don-Wauchope A, McKelvie R, et al. BNP and NT-proBNP as prognostic markers in persons with chronic stable heart failure[J]. Heart Fail Rev, 2014, 19(4):471-505.
- [8] Hill SA, Booth RA, Santaguida PL, et al. Use of BNP and NT-proBNP for the diagnosis of heart failure in the emergency department; a systematic review of the evidence[J]. Heart Fail Rev, 2014, 19(4):421-438.
- [9] Oremus M, McKelvie R, Don-Wauchope A, et al. A systematic review of BNP and NT-proBNP in the management of heart failure; overview and methods [J]. Heart Fail Rev, 2014, 19(4):413-419.
- [10] Savarese G, Musella F, D'Amore C, et al. Changes of natriuretic peptides predict hospital admissions in patients with chronic heart failure: a meta-analysis [J]. JACC Heart Fail, 2014, 2(2):148-158.

(收稿日期:2016-08-04 修回日期:2016-10-04)