

论著·临床研究 doi:10.3969/j.issn.1671-8348.2019.23.018
网络首发 http://kns.cnki.net/kcms/detail/50.1097.R.20190813.1557.006.html(2019-08-14)

成年人血清细胞角质蛋白 19 可溶性片段地区性参考区间的建立与探讨

顾玉兰¹, 丘佳明², 陈丽芳³, 宛传丹^{3△}

(1. 江苏省常熟市第二人民医院肿瘤科 215500; 2. 江苏省常熟市第二人民医院病理科 215500;
3. 江苏省常熟市医学检验所 215500)

[摘要] **目的** 比较分析两种建立成年人血清细胞角质蛋白 19 可溶性片段(CYFRA21-1)参考区间方案, 探讨适用于常熟地区成年人血清 CYFRA21-1 的实验室参考区间。**方法** 调取 2017 年 6 月至 2018 年 6 月江苏省常熟市医学检验所实验室信息管理系统(LIS)中存储的所有 CYFRA21-1 检验数据, 筛选有效数据, 其中男 9 624 例, 女 6 926 例, 以年龄段分为青年组(≤ 45 岁)和老年组(> 45 岁)。采用统计学线性拟合法建立成年人血清 CYFRA21-1 间接参考区间。选取 2018 年 1—6 月送检的健康体检者($n=806$)CYFRA21-1 检测数据以同样分组法, 采用非参数方法建立各分组的成年人血清 CYFRA21-1 参考区间。**结果** 线性拟合建立的成年人血清 CYFRA21-1 间接参考区间分别为男青年组($0\sim 3.49$ ng/mL)、女青年组($0\sim 2.74$ ng/mL)、男老年组($0\sim 3.84$ ng/mL)、女老年组($0\sim 3.03$ ng/mL)。健康体检者非参数方法建立的成年人血清 CYFRA21-1 参考区间分别为男青年组($0\sim 3.35$ ng/mL)、女青年组($0\sim 3.18$ ng/mL)、男老年组($0\sim 3.52$ ng/mL)、女老年组($0\sim 3.21$ ng/mL)。两种参考区间之间差异无统计学意义($P>0.05$)。**结论** 利用已存储数据线性拟合建立的间接参考区间与健康体检者非参数法建立参考区间同等有效。初步建立了适用于常熟地区成年人血清 CYFRA21-1 的参考区间。

[关键词] 细胞角质蛋白 19 可溶性片段; 参考值; 生物标记; 肿瘤

[中图分类号] R446.1 **[文献标识码]** A **[文章编号]** 1671-8348(2019)23-4027-04

Establishment and discussion of regional reference interval of serum soluble Cytokeratin 19 fragment in adult serum

GU Yulan¹, QIU Jiaming², CHEN Lifang³, WAN Chuandan^{3△}

(1. Department of Oncology, Second People's Hospital of Changshu City, Changshu, Jiangsu 215500, China; 2. Department of Pathology, the Second People's Hospital of Changshu City, Changshu, Jiangsu 215500, China; 3. Changshu Medical Laboratory, Changshu, Jiangsu 215500, China)

[Abstract] **Objective** To compare and analyze the regional reference interval of soluble Cytokeratin 19 fragment (CYFRA21-1) in adult serum established by two methods, and to explore their application to Changshu area. **Methods** All the CYFRA21-1 test data stored in the laboratory information management LIS system of Changshu Medical Laboratory from June 2017 to June 2018 were collected and valid data were screened, including 9 624 cases of male and 6 926 cases of female. All the cases were divided into the young group (≤ 45 years) and the elderly group (> 45 years). The indirect reference interval of adult CYFRA21-1 was established by statistical linear fitting method. The health checkers were selected and the same grouping method were applied on them, the reference intervals of adults CYFRA21-1 for each group were established by non-parametric method. **Results** The indirect reference interval of adult CYFRA21-1 established by linear fitting method was $0\sim 3.49$ ng/mL for the male youth group, $0\sim 2.74$ ng/mL for the female youth group, $0\sim 3.84$ ng/mL for the male elderly group and $0\sim 3.03$ ng/mL for the female elderly group, respectively. The adult CYFRA21-1 reference interval for adults established by non-parametric methods for healthy subjects was $0\sim 3.35$ ng/mL for the male youth group, $0\sim 3.18$ ng/mL for the female youth group, $0\sim 3.52$ ng/mL for the male elderly group, and $0\sim 3.21$ ng/mL for the female elderly group, respectively, there was no significant difference between the two reference intervals ($P>0.05$). **Conclusion** The indirect reference interval established by linear fitting method of the stored data is as effective as the reference interval established by non-parametric method of health subjects. The CYFRA21-1 reference interval for adults in this area has been preliminarily established.

[Key words] soluble fragment of cytokeratin 19; reference values; biomarkers, tumor

细胞角质蛋白作为上皮细胞的骨架重要分子,广泛分布于正常组织细胞表面。在恶性上皮细胞中,因为异常活跃的蛋白酶类促进了细胞的降解,使得大量细胞角质蛋白可溶性片段释放进入血液^[1-2]。其中细胞角质蛋白 19 可溶性片段(CYFRA21-1)是一种可用于诊断及评价非小细胞肺癌的肿瘤标记物,具有较高的敏感度和特异度^[3-5]。

指标的量化检验涉及参考区间的建立,其直接影响到临床医生对疾病的判断。而肿瘤指标参考区间常受到种族、地域、年龄、性别及地区饮食习惯等众多因素的影响。因此,国际临床化学联合会(IFCC)建议应结合本地区建立适用本地区实验室的参考区间^[6]。本文以统计学线性拟合法分析基于常熟市医学检验所实验室信息管理系统(LIS)现有的 CYFRA21-1 检验数据建立间接参考区间,并与美国临床和实验室标准化委员会(CLSI)推荐的非参数法建立的参考区间进行比较,探寻更为便捷、省力且有效的参考区间验证方案,为本地区提供更为精准的 CYFRA21-1 参考区间。

1 资料与方法

1.1 一般资料 收集江苏省常熟市 LIS 中 2017 年 6 月至 2018 年 6 月所有 CYFRA21-1 的检测数据。这些标本来源于本地区 3 家大型综合医院(江苏省常熟市第一人民医院、江苏省常熟市第二人民医院和江苏省常熟市中医院)及各乡镇二级医院、卫生院等,记为临床检测组。纳入标准为经全面体检和影像学检查排除患有严重心、肺、肾、肝、肠、胃等基础疾病;无内分泌系统、免疫系统等疾病;无肿瘤病史。排除妊娠期及哺乳期的女性。将 2018 年 1—6 月各医院送检的体检健康者记为健康体检组,两组分别以不同的方法进行统计计算。

1.2 CYFRA21-1 检测 所有受检者均在早晨、空腹、安静状态下采集静脉血 5 mL 左右,室温静置 30 min 以上,2 h 内分离血清,以 $3\,000\times g$ 离心 5 min。采用美国雅培公司的 Architecti 全自动化学发光免疫分析系统及配套试剂进行检测分析。方法为化学发光微粒子免疫检测法,质控品及校准品均由美国雅培公司提供。每天质控均在控,并做好记录。

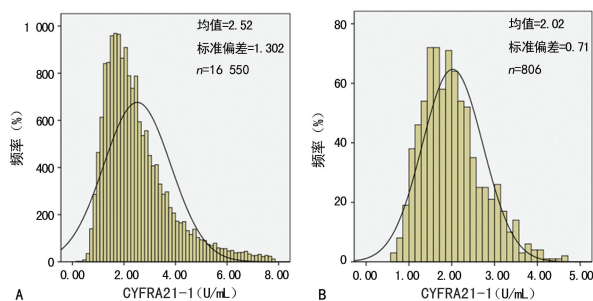
1.3 数据整理 导出的原始数据剔除同一患者重复检测数据,保存第 1 次有效数据。排除临床诊断为肿瘤相关疾病的患者检测数据。以统计学软件 SPSS20.0 四分位法将大于 $P_{75} + 3 \times (P_{75} - P_{25})$ 的数据作为离群值剔除。将最终有效数据用 SPSS20.0 软件分析其频数分布。

1.4 计算参考区间 临床检测组:时间段内经过筛选的所有有效数据将以 R 语言软件进行以频数分布表的累积频率为横坐标,检测值为纵坐标,绘制累积频率图。R 编程对任意两点间数据进行无限循环,直到识别出最佳线性。循环条件设为同时满足最小残差平方和占累积分布最大百分比的线性部分。上述

条件的两点数据间采用 Origin 软件拟合线性回归方程 $Y=bX+a$ (Y 为检测值, X 为累积频率),外插法求得间接参考区间^[7]。因为在实际临床应用中,肿瘤标记物相关指标均以检测数据高值为异常。因此本研究计算 CYFRA21-1 数据单侧 95% 位值(P_{95})作为参考区间上限,具体表示为 95% 参考区间 $[0 \sim (95b + a)]$ 。健康体检组的有效数据,利用 Kolmogorov-Smirnov 检验分析数据是否呈正态分布,正态分布的计量资料则以 $\bar{x} \pm s$ 表示。而呈偏态分布的数据采用非参数的百分位法表示,根据 CLSI 文件 C28-A3 的要求,以 $P < 0.05$ 的偏态分布数据应采用非参数排序法建立参考区间^[8],表示方法同临床检测组。

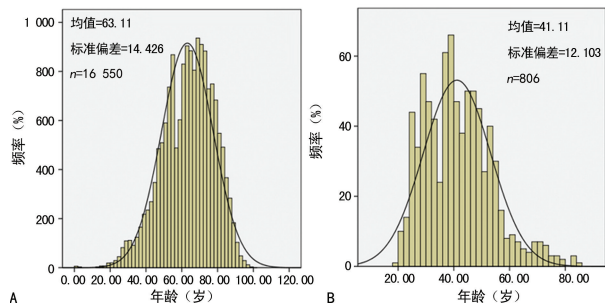
2 结果

2.1 原始数据与筛选结果 2017 年 6 月至 2018 年 6 月共检测 CYFRA21-1 28 899 例,依据四分位法统计分析检测值,将大于 $P_{75} + 3(P_{75} - P_{25})$ 的数据作为离群值剔除。再根据纳入标准剔除无效数据,最后经过筛选后得临床检测组有效数据 16 550 例,年龄 20~100 岁,平均 (63.11 ± 14.43) 岁;其中男 9 624 例,女 6 926 例。健康体检组共 818 例,经过筛选后得有效数据 806 例,年龄 19~84 岁,平均 (41.11 ± 12.10) 岁;其中男 570 例,女 236 例。将数据依次按性别、年龄进一步分为男青年组、男老年组、女青年组、女老年组进行研究,结果见表 1。年龄与 CYFRA21-1 数据频率分布直方图,见图 1、2。



A: 临床检测组; B: 健康体检组

图 1 CYFRA21-1 检测值频率分布直方图



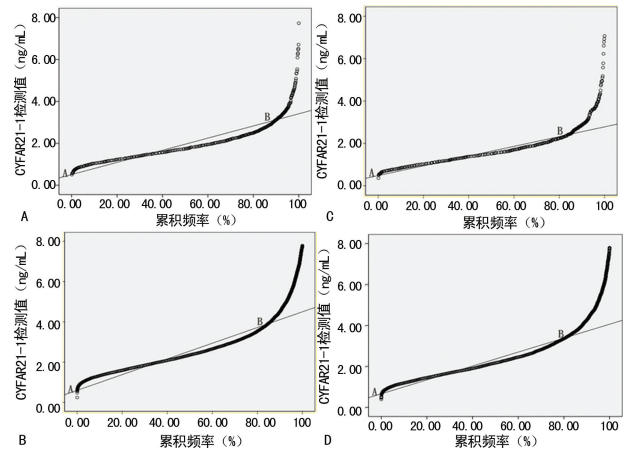
A: 临床检测组; B: 健康体检组

图 2 年龄频率分布直方图

2.2 线性拟合法计算参考区间 通过 R 语言软件利用 LIS 信息系统里存储的 CYFRA21-1 临床检测数据建立的累积频率分布直方图见图 1,通过拟合线性回归方程,求得 95% 参考区间,见表 1。

表 1 LIS 存储临床数据组血清 CYFRA21-1 间接参考区间

组别	年龄	n	年龄($\bar{x}\pm s$,岁)	拟合线性回归方程	R ²	线性比例(%)	95%参考区间(ng/mL)
男青年组	20~45岁	1 048	36.45±7.38	Y=0.029X+0.510	0.931	89.1	0~3.49
男老年组	>45岁	8 576	67.05±11.00	Y=0.039X+0.597	0.879	91.3	0~3.84
女青年组	20~45岁	842	37.23±7.15	Y=0.023X+0.482	0.962	90.6	0~2.74
女老年组	>45岁	6 084	65.74±11.64	Y=0.034X+0.661	0.908	86.5	0~3.03



A:男青年组;B:女青年组;C:男老年组;D:女老年组
图 3 CYFRA21-1 累积频率分布图(A、B 两点
间数据代表最佳线性)

2.3 非参数法计算参考区间 对 CYFRA21-1 进行

正态分布检验显示,样本各组数据呈现偏态分布($P<0.05$),参考区间的建立应取其右单侧第 95 百分位数作为参考上限。采用中位数 M 及四分位数(P_{25} , P_{75})来描述各组数据。见表 2。

表 2 健康体检组血清 CFRA21-1 参考区间

组别	年龄	n	年龄($\bar{x}\pm s$,岁)	CFRA21-1 [M(P_{25} , P_{75})]	95% 百分位数
男青年组	20~45岁	379	33.67±6.81	1.91(1.55,2.38)	3.35
男老年组	>45岁	191	55.58±9.28	2.22(1.68,2.81)	3.52
女青年组	20~45岁	154	35.60±6.30	1.57(1.26,2.06)	3.18
女老年组	>45岁	82	52.10±5.43	1.94(1.44,2.42)	3.21

2.4 两种方案参考区间比较 两种方案数据来自不同人群样本,通过 t 检验比较发现两种方案得出的各分组参考区间之间差异无统计学意义($P>0.05$)。见表 3。

表 3 两种方案参考区间比较分析

组别	年龄	临床检测组		健康体检组		t	P
		n	95%参考区间上限	n	95%参考区间上限		
男青年组	20~45岁	1 048	3.49	379	3.35	-1.228	0.220
男老年组	>45岁	8 576	3.84	191	3.52	-1.148	0.107
女青年组	20~45岁	842	2.74	154	3.18	0.170	0.865
女老年组	>45岁	6 084	3.03	82	3.21	1.650	0.091

3 讨 论

血清 CYFRA21-1 是肺癌,特别是非小细胞肺癌最为敏感的肿瘤标记物,常用于监测非小细胞肺癌的病程^[5]。其特异性较强,可用于肺癌与肺炎、肺结核、慢性阻塞性肺炎等肺部疾病的鉴别诊断。同时,血清 CYFRA21-1 针对非小细胞肺癌化疗等效果的评价与预后评估也具有十分重要的临床意义^[3,9]。血清 CYFRA21-1 参考区间对于临床上非小细胞肺癌诊断至关重要,而不同种族人群、不同地区及年龄、性别等均可影响到参考区间的设立。文献检索显示,国内仅有 3 篇论文涉及 CYFRA21-1 的参考区间建立与设定^[7-8]。3 个地区(上海、苏州、南通)所采用的方法不完全相同,且得出的结论之间存在差异。上海、苏州地区均以非参数百分位法建立参考区间,取 95%百分位数为上限^[10-11]。该方法样本较小,粗略地估算了参考区间。南通地区以体检与医院存储数据进行统计

分析,以科学统计方法评估出最适用于该地区的参考范围,数据精准可靠。常熟地区自 2010 年江苏省常熟市医学检验所成立后,将本地区所有检验资源整合一起,质量管理体系通过 ISO15189 认可,建立了完善的 LIS,保存了大量临床样本检测数据。据国外文献研究,参考区间可以基于当前数据,通过相关统计算法建立^[12-13]。

本文研究以两种常用的方法来推导参考区间,一是建立在本地区成年人人群检验数据基础之上;二是建立在健康成年人样本库,以统计学方法推算出参考区间。通过分析,两种方法的结果与厂家提供的均存在一定差异。厂家说明书的参考区间为 0~3.3 ng/mL,该参考区间不分青年组(20~45 岁)及老年组(>45 岁)。而本研究建立在 LIS 大数据分析基础上,得出的常熟地区血清 CYFRA21-1 男青年组 95%间接参考区间为 0~3.49 ng/mL;男老年组 95%间接参

考区间为 0~3.84 ng/mL。女青年组 95%间接参考区间为 0~2.74 ng/mL;女老年组 95%间接参考区间为 0~3.03 ng/mL。男性参考区间上限高于女性,且老年组要高于青年组,与原参考区间上限也存在差异。另一方案通过建立在本地区健康人群的血清 CYFAR21-1 检测数据上的分析显示,血清 CYFAR21-1 男青年组 95%参考区间上限为 3.35 ng/mL;男老年组 95%参考区间上限为 3.52 ng/mL;女青年组 95%参考区间上限为 3.18 ng/mL;女老年组 95%参考区间上限为 3.21 ng/mL。此方案与第一种方案得出的参考区间相比较差异无统计学意义($P>0.05$)。这与周边地区研究报道结果相一致^[10-11]。常熟地区成年人血清 CYFAR21-1 参考区间上限男性要高于女性,老年组高于青年组。年龄与性别是影响常熟地区成年人血清 CYFAR21-1 参考区间的重要因素,实验室在设立参考区间应区别对待,分组设定^[14]。

通过临床检测数据计算参考区间的主要不足在于临床送检数据中包含大量疾病患者数据。因此对数据进行仔细筛选十分重要,这将直接影响到结果的正确性^[15]。本文通过仔细筛选,统计分析得到的结果与以健康人群建立的参考区间,两者间差异无统计学意义($P>0.05$),说明该方法可以运用于临床检验计量项目参考区间的建立。而该方法的主要优势在于其针对的主要人群与该指标的适用人群相对应,即不以另一人群的参考值来判别本人群的异常情况。同时,该方法基于的数据虽然庞大,但不需另行设计实验来进行抽样分析,可以节省大量的人力与物力。

综上所述,通过对常熟地区成年人血清 CYFRA21-1 参考区间的精准设定,既能指导医师对病情做出准确判断,又能避免给患者造成不必要的心理恐慌与过度的医疗检查。

参考文献

[1] KUANG L I,SONG W J,QING H M,et al. CYFRA21-1 levels could be a biomarker for bladder cancer:a Meta-analysis[J]. Genet Mol Res,2015,14(2):3921-31.

[2] WANG G,ZHAO H,SHI X,et al. Serum cyfra21-1 as a biomarker in patients with nonsmall cell lung cancer[J]. J Cancer Res Ther,2014(Suppl 10):S215-217.

[3] 李玲,宋丽华,丁士超,等. 癌胚抗原和细胞角蛋白 19 片段作为晚期非小细胞肺癌疗效评价指标的临床价值[J]. 中华肿瘤杂志,2010,32(11):850-855.

[4] ZHANG Z H,HAN Y W,LIANG H,et al. Prognostic

value of serum CYFRA21-1 and CEA for non-small-cell lung cancer[J]. Cancer Med,2015,4(11):1633-1638.

[5] SIEGEL R L,MILLER K D,JEMAL A. Cancer statistics,2017[J]. CA Cancer J Clin,2017,67(1):7-30.

[6] LISTED N A. International Federation of Clinical Chemistry(IFCC),Scientific Committee,Clinical Section. Expert Panel on Theory of Reference Values (EPTRV). The theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits[J]. Clin Chim Acta,1987,165(1):111-118.

[7] 李小佩,王惠民,王建新. 应用医院存储数据建立血清 Cyfra21-1、SCCA、CEA 参考区间[J]. 临床检验杂志,2015,33(12):946-950.

[8] Insitiute CALS. Defining,establishing,and verifying reference intervalsin the clinical laboratory;approved guideline [M]. 3rd ed. Wayne: Clinecaland Laboratory Standards Institute,2008.

[9] THOMAS D S,FOURKALA E O,APOSTOLIDOU S, et al. Evaluation of serum CEA,CYFRA21-1 and CA125 for the early detection of colorectal cancer using longitudinal preclinical samples[J]. Br J Cancer,2015,113(2):268-274.

[10] 吴婷婷,邓胜明,李清茹,等. 苏南地区健康人群血清细胞角蛋白 19 片段参考区间的研究[J]. 中国血液流变学杂志,2017,27(3):323-326.

[11] 李鼎,陆云,刘兴党. 上海地区多个肿瘤标志物正常参考值的初步确定[J]. 标记免疫分析与临床,2014,21(5):593-596.

[12] SHAW J L V,COHEN A,KONFORTE D,et al. Validity of establishing pediatric reference intervals based on hospital patient data;a comparison of the modified Hoffmann approach to CALIPER reference intervals obtained in healthy children[J]. Clin Biochem,2014,47(3):166-172.

[13] POOLE S,SCHROEDER L F,SHAH N. An unsupervised learning method to identify reference intervals from a clinical database[J]. J Biomed Inform,2016,59:276-284.

[14] 岳波,刘曼娇,唐大海,等. 常规体检生物化学项目波动性参考区间的建立与验证[J]. 临床检验杂志,2018,36(6):418-421.

[15] 葛丽丽,杨俊梅,李进香. 郑州地区儿童血清神经元特异性烯醇化酶参考区间的建立[J]. 重庆医学,2016,45(31):4413-4415.