

• 论 著 •

新疆地区糖链抗原 72-4 生物参考区间的初步建立与探讨

韩松涛, 赵晓群, 黄艳春[△]

(新疆医科大学附属肿瘤医院检验科, 新疆乌鲁木齐 830011)

摘要:目的 使用间接法建立该地区实验室的糖链抗原 72-4(CA72-4)参考区间,验证间接法可靠性。方法 通过新疆医科大学附属肿瘤医院及新疆医科大学附属第四医院实验室信息系统收集 2014 年 1 月至 2019 年 3 月的 39 957 例体检人群的 CA72-4 检测结果。对数据进行“Box-Cox”正态转换后,使用 Turkey method 对数据进行剔除和入组。筛选入组的数据经过男、女性别组间差异检验和年龄相关性分析后,利用非参数法建立 CA72-4 参考区间。结果 不同性别 CA72-4 测量值的差异有统计学意义($Z = -11.425, P < 0.05$);CA72-4 的生物参考区间分别为男 < 8.65 U/mL,女 < 8.77 U/mL。结论 使用间接法建立检验项目的参考区间是一种比较简便快捷的方法,临床实验室可依据不同地区、不同人群建立符合该实验室的参考区间。

关键词:糖链抗原 72-4; 间接法; 参考区间

DOI:10.3969/j.issn.1673-4130.2020.03.012

中图法分类号:R-331

文章编号:1673-4130(2020)03-0301-04

文献标识码:A

Preliminary establishment and discussion of the biological reference interval of sugar chain antigen 72-4 in Xinjiang

HAN Songtao, ZHAO Xiaojun, HUANG Yanchun[△]

(Department of Clinical Laboratory, Cancer Hospital Affiliated to Xinjiang Medical University, Urumqi, Xinjiang 830011, China)

Abstract: Objective To establish a reference interval for glycan antigen 72-4 (CA72-4) in laboratories in this area by indirect method to verify the reliability of indirect method. **Methods** The CA72-4 test results of 39 957 physical examination patients from January 2014 to March 2019 were collected through the laboratory information system of Cancer Hospital Affiliated to Xinjiang Medical University and the Fourth Hospital Affiliated to Xinjiang Medical University. After performing "Box-Cox" normal transformation on the data, the data is removed and grouped using the Turkey method. After screening the data, the non-parametric method was used to establish the CA72-4 reference interval after the difference test and age correlation analysis between the male and female groups. **Results** The difference in CA72-4 measurements between different genders was statistically significant ($Z = -11.425, P < 0.05$); the biological reference intervals for CA72-4 were males < 8.65 U/mL and females < 8.77 U/mL. **Conclusion** The indirect method to establish the reference interval of the test items is a relatively simple and fast method. Clinical laboratories can establish reference intervals that fit themselves according to different regions and different populations.

Key words: sugar chain antigen 72-4; indirect method; reference interval

生物参考区间通常指参考值分布的 95% 置信区间^[1],是解释检验结果、分析检验信息的基本依据和尺度,同时可靠的参考区间对指导临床患者诊治提供坚实依据。国际临床化学和实验室医学联盟推荐的参考区间(RIs)建立方法对临床实验室过于复杂,错误使用率高^[2]。国际临床化学和实验室医学联盟和美国临床实验室标准化协会在 2010 年共同修订了名

为《临床实验室中参考区间的定义、建立和验证》的指南第 3 版(EP28-A3c)。在指南中,建立参考区间的方法得到了补充,使用不少于 120 例参考个体的非参数法,并指出参考区间的建立是繁琐、耗时且昂贵的,因此允许临床实验室使用已建立的参考区间^[3]。糖链抗原 72-4(CA72-4)指标广泛用于胃肠道肿瘤和卵巢癌的诊断、治疗和预后监测^[4-8]。新疆医科大学附属

作者简介:韩松涛,男,研究生在读,主要从事肿瘤免疫方面的研究。△ 通信作者,E-mail:huangyanchun0619@sohu.com。

本文引用格式:韩松涛,赵晓群,黄艳春.新疆地区糖链抗原 72-4 生物参考区间的初步建立与探讨[J].国际检验医学杂志,2020,41(3):301-

肿瘤医院和新疆医科大学附属第四医院 CA72-4 目前均采用试剂说明书中推荐的 6.9 U/mL 作为参考区间的上限,但在体检中发现有些体检者无明显不适症状,却表现出 CA72-4 异常高于试剂说明书中的参考区间的上限,初步猜测可能跟选择的健康人群及地域有关。因此,本研究选取一家肿瘤专科医院和一家综合性医院的体检人群,初步建立 CA72-4 的参考区间,并评估其研究价值。

1 资料与方法

1.1 一般资料 数据来自新疆医科大学附属肿瘤医院及新疆医科大学第四附属医院的实验室信息管理系统,调取既往体检健康者基本信息及 CA72-4 的检测值,共 39 957 例。本研究仅对 >18 岁的成年人的指标进行分析,纳入标准:(1)检测信息完整准确,既往体检健康者;(2)无胃肠、妇科疾病、药物治疗(如非甾体抗炎药、糖皮质激素等)等基础疾病史。排除标准:(1)有慢性疾病、代谢性疾病、营养性疾病的患者;(2)既往有相关治疗病史。纳入 39 889 例,其中男性 17 484 例,女性 22 405 例。本研究经医院伦理委员会批准。

1.2 仪器和试剂 全自动电化学发光免疫分析仪购自瑞士 Roche 公司,型号 Cobase601;CA72-4 检测试剂盒和校准品购自瑞士 Roche 公司。采用经典的 Westgard 多规则质控作为室内质量控制(1_{2s} 作为警告规则,1_{3s}、2_{2s}、R_{4s}、4_{1s}、10_x 作为失控规则),新疆医科大学附属肿瘤医院实验室累积的变异系数(CV)为 5%,附属第四医院实验室累积的 CV 为 7%。

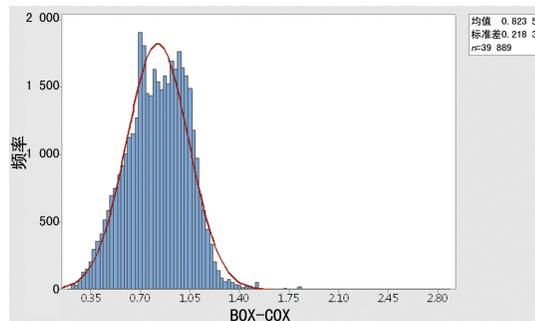
1.3 方法 使用偏度-峰度值检验(Skewness-Kurtosis)对数据进行正态性分析,当偏度值和峰度值分别小于对应偏度和峰度标准差的 1.96 倍即可判定为呈正态分布。非正态分布数据通过 Box-Cox 变换转变为近似正态分布^[9],并对变换后数据使用偏度-峰度值检验分析正态性,并绘制频数分布直方图。离群值剔除采用四分位间距法^[10],剔除均值±1.5 倍四分位间距之外的数值。运用非参数秩和检验、Spearman 相关性分析计算该项目不同性别、年龄之间是否有统计学差异。采用非参数法进行 RIs 的建立。运用非参数统计法计算 97.5% 的值作为上限。

1.4 统计学处理 运用 Office Excel 2010 及 SPSS19.0 软件分析,通过非参数法,并依据参考变化值(RCV) = 2^{1/2} × Z × [实验室累积变异系数(CVA)² + 生物变异度(CVI)²]^{1/2} 计算本研所得到的结果差异是否具有统计学意义,比较 CA72-4 检测项目的上限水平,建立符合本实验室的 CA72-4 的 RIs,并与当前所运用的 RIs 进行比较。

2 结果

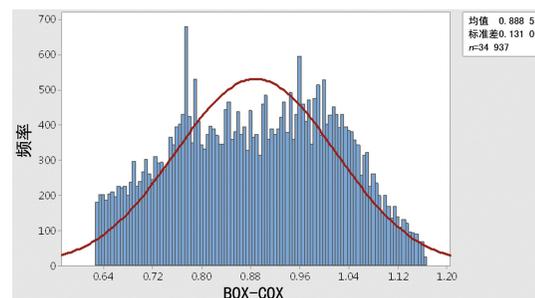
2.1 数据分布 经分析整理,纳入 39 889 例,剔除离群值后,最终纳入的数据为 34 937 例,其中男性

15 544 例,18~<30 岁 1 626 例,30~<50 岁 7 561 例,50~<70 岁 5 247 例,≥70 岁 1 110 例;女性 19 393 例,18~<30 岁 682 例,30~<50 岁 8 430 例,50~<70 岁 8 956 例,≥70 岁 1 325 例,经 Box-Cox 变换后,绘制频数分布直方图观察数据的分布情况,见图 1、2。



注:横坐标表示 CA72-4 测量值(U/mL);纵坐标表示频率(例)。

图 1 39 889 例 CA72-4 检测数据分布直方图



注:横坐标为 CA72-4 测量值(U/mL),纵坐标为频率(例)。

图 2 最终纳入 34 937 例 CA72-4 值 Box-Cox 变换后分布直方图

2.2 CA72-4 检测结果和参考区间建立 运用非参数秩和检验,得出 Z = -11.425, P < 0.05,不同性别间差异有统计学意义,可依据性别建立 RIs。运用 Spearman 相关性分析,CA72-4 检测结果与年龄差异无统计学意义(P > 0.05),不能根据年龄建立 RIs。采用非参数排序法,分别对已剔除过离群值的男、女数据进行升序排列,取 97.5% 的数值作为参考区间的上限。见表 1。

表 1 不同性别之间 CA72-4 测量值比较和参考区间的建立

性别	n	秩均值	秩和	Z	P	RIs (U/mL)
男性	15 544	18 157.56	282 241 093.50	-11.425	0.000	<8.65
女性	19 393	16 917.10	328 073 359.50			<8.77

2.3 运用 RCV 公式比较 CA72-4 的生物变异度为 16.06%^[11],新疆医科大学附属肿瘤医院实验室累积的 CV 为 5%,附属医院第四实验室累积的 CV 为 7%,平均 CV 为 6%。Z 为 95% 的置信区间,在本研究中取 1.96 数值,经计算 RCV = 2^{1/2} × 1.96 × [0.160 62 + 0.062]^{1/2} = 47.52%;比较 RCV_{max} 及 RCV

的大小。得出男性 $RCV_{\max} = 25.36$, 女性 $RCV_{\max} = 27.10$, 不同性别所建立的 RIs 差异无统计学意义。

2.4 运用 EP28-A3c 文件验证参考区间 运用本研究建立的 RIs 发现, 在 34 937 例数据中, 异常例数 1 367 (占比 3.91%), 低于厂家提供的 RIs 刷选出的异常例数 2 060 (占比 5.90%); 按照 EP28-A3c 文件根据不同性别检测 20 例健康体检者的 CA72-4, 并进行验证。经过验证, 本研究所建立的参考区间适用于本实验室, 见表 2。

表 2 CA72-4 参考区间验证和新 RIs 正常者、异常者的频数变化

组别	n	厂家提供(U/mL)	本研究 RIs	RIs 验证
男性	20	0~6.9	<8.65	<8.13
女性	20	0~6.9	<8.77	<8.72

3 讨 论

直接法建立参考区间需要筛选健康人群纳入参考个体, 这一过程极为昂贵和耗时, 并且在当今的临床工作中难以实现, 尤其是当参考区间需要根据年龄和性别差异分别建立或者需要针对一些特殊人群进行研究时, 入组参考个体的选取会变得异常困难。因此许多实验室改为引用试剂厂商或者其他实验室的参考区间, 但是直接套用他人建立的参考区间可能会存在服务人群和检测系统的差异, 会对临床工作和患者造成困扰^[2]。因此, 迫切需要一种快捷及简单方便的方法建立本地区或者适合本实验室的参考区间。国内外有许多研究基于大量样本统计分析计算出检测项目的 RIs, 如 Robust 法、Bhattacharya 法、传统的非参数法等^[12]。随着实验室信息管理系统的发展, 调取大量实验室数据已并非难事, 也为运用间接法建立实验室 RIs 奠定了很好的基础。同时制定严格的纳入、排除标准可为数据的可靠性提供重要的保证。在本研究中, 回顾性分析了近 5 年体检人群的 CA72-4 结果, 在制定纳入、排除标准时, 由于使用相对较长时间范围内的数据, 对于分析前的因素的控制并不是很明确, 从而导致在建立参考区间受到影响, 因此在制定纳入、排除标准的同时, 观察体检人群其他相应指标(生化、免疫、血常规等)或者是翻阅既往病例和进行电话随访并制定患者调查问卷, 以排除潜在对该指标检测有影响的人群; 其次加大数据量可以优化所建立的参考区间, 并尽可能将检测结果不确定的人群控制在 5% 以下, 以此来控制分析前因素对该指标的影响。

建立生物参考区间过程中, 数据的剔除是关键步骤。本研究采用剔除四分位间距法^[10], 剔除均值 ± 1.5 倍四分位距之外的数值。运用此剔除方法要求数据呈正态或对称分布, 如原数据呈偏态分布, 需将数

据转换成正态或对称分布形式。本研究数据呈偏态分布, 故选择“Box-Cox”转换方式对数据进行正态性转换^[9]。然而在间接法建立实验室生物参考区间的文献当中, 数据剔除法各有不同^[13-14], 每种方法各有优缺点, 在选择数据剔除时可根据数据量和数据分布情况, 选用不同剔除方法以到达剔除效果。许多离群值剔除方法都是以数据正态分布为前提^[15], 所剔除的离群值均按照剔除公式严格计算。本研究最终纳入的数据为 34 937 例, 共剔除 4 952 例, 剔除比例为 12.4%。出现这种情况, 这与离群值剔除既按照剔除公式计算, 又循环使用本研究中的方法直到剔除所有离群值为止有关; 同时亦不排除本研究中数据基数大导致剔除比例较高。如果离群值剔除不正确会导致参考区间反常缩小或者扩大, 因此只有严格按照剔除公式进行反复循环剔除离群值才能保证所建参考区间的正确性。

确定参考区间的上、下限是最重要的一环。通常确定参考区间的上、下限有参数法、非参数法及罗伯斯特法等^[15]。参数法是建立参考区间的最经典的方法, 但是一般要求数据必须呈正态分布。而大多数的数据并非正态分布, 这也是限制了参数法在确立参考区间的应用。罗伯斯特法是一种介于参数法与非参数法之间的方法^[16], 具有类似参数估计的过程并且对样本量没有太高的要求, 对于数据分布形态的要求不如其他的方法严格。正是这种特性使其在非正态分布的数据中同样表现出色^[3]。从严格意义上来说, 罗伯斯特法是一种小样本估计参考区间的方法, 被广泛应用于样本量少于 120 例的情况中^[15]。EP28-A3c 指南明确指出非参数法是建立生物参考区间的推荐方法之一。由于参数法的计量水准明显高于非参数法, 因此非参数法对样本数量的要求相对较高。根据相关研究者建议, 使用非参数法进行参考区间研究, 最小的可靠观测量为 120 例。无论获得多大量的参考值, 数据仍然应该根据非参数方法分析并根据适当的百分位数报告^[3]。对大部分分析物而言, 应该选取数据分布的 2.5% 和 97.5% 百分位值作为上、下参考限的取值^[3]。

现中国大多数医院所使用的参考区间均是依据欧美健康人群所建立的, 且根据卫健委最新颁布的有关肿瘤标志物的参考区间行业标准(WS/T645. 2-2018)中并没有涉及 CA72-4。本研究通过采用非参数法建立参考区间, 得出 CA72-4 参考区间男性 < 8.65 U/mL, 女性 < 8.77 U/mL, 与试剂厂商所提供的参考区间有一定的差别, 这种差别有可能是因为地区或者人群的选择不同, 也有可能是在本研究中仅采用了非参数法这一种方法进行参考区间的建立, 不排除此种方法存在局限性。因此在建立参考区间的过

程中可进行两种以上方法的比较,以得出更加符合本实验室或者本地区健康人群的参考区间。

4 结 论

本研究通过非参数法初步得到新疆地区 CA72-4 的参考区间,具有一定参考价值。同时随着医院的实验室信息管理系统逐步完善,使用间接法建立实验室指标的 RIs 比直接法更加简单明了,相应技术要求较低。对于各级医院实验室可依据不同地区、不同人群建立符合本地区或本实验室的 RIs,为临床或患者提供更好参考依据。

参考文献

[1] 中国合格评定国家认可委员会. 医学实验室质量和能力认可准则(ISO15189:2007):CNAS-CLO2[S]. 北京:中国标准出版社,2008.

[2] GEFFRÉ A, FRIEDRICHS K, HARR K, et al. Reference values; a review [J]. *Vet Clin Pathol*, 2010, 38(3): 288-298.

[3] Clinical and Laboratory Standards Institute. Defining, establishing, and verifying reference intervals in the clinical laboratory: approved guideline-third edition; C28-A3[S]. Wayne, PA, USA: CLSI, 2010.

[4] SUN Z, ZHANG N. Clinical evaluation of CEA, CA19-9, CA72-4 and CA125 in gastric cancer patients with neoadjuvant chemotherapy[J]. *World J Surg Oncol*, 2014, 12(1): 397.

[5] CHEN X Z, ZHANG W K, YANG K, et al. Correlation between serum CA724 and gastric cancer; multiple analyses based on Chinese population[J]. *Mol Biol Rep*, 2012, 39(9): 9031-9039.

[6] GRANATO T, MIDULLA C, LONGO F, et al. Role of HE4, CA72. 4, and CA125 in monitoring ovarian cancer [J]. *Tumour Biol*, 2012, 33(5): 1335-1339.

[7] LENHARD M S, NEHRING S, NAGEL D, et al. Predictive value of CA 125 and CA72-4 in ovarian borderline tumors[J]. *Clin Chem Lab Med*, 2009, 47(5): 537-542.

[8] EMOTO S, ISHIGAMI H, YAMASHITA H, et al. Clinical significance of CA125 and CA72-4 in gastric cancer with peritoneal dissemination[J]. *Gastric Cancer*, 2012, 15(2): 154-161.

[9] YANG Y, CHRISTENSEN O F, SORENSEN D. Analysis of a genetically structured variance heterogeneity model using the Box-Cox transformation[J]. *Genet Res*, 2011, 93(1): 33-46.

[10] HORN P S, FENG L, LI Y, et al. Effect of outliers and nonhealthy individuals on reference interval estimation [J]. *Clin Chem*, 2001, 47(12): 2137-2145.

[11] QI Z, ZHANG L, CHEN Y, et al. Biological variations of seven tumor markers[J]. *Clin Chim Acta*, 2015, 450: 233-236.

[12] THYGESEN K, ALPERT J S, JAFFE A S, et al. Third universal definition of myocardial infarction[J]. *Circulation*, 2012, 126(16): 2020-2035.

[13] 郭玮, 宋斌斌, 沈隽霏, 等. 基于实验室数据信息库建立丙氨酸转氨酶生物参考区间[J]. *中华医学杂志*, 2015, 95(32): 2599-2602.

[14] ARAL H, USTA M, CILINGIRTURK A M, et al. Verifying reference intervals for coagulation tests by using stored data[J]. *Scand J Clin Lab Invest*, 2011, 71(8): 647-652.

[15] 沈隽霏, 宋斌斌, 潘柏申. 间接法建立生物参考区间[J]. *检验医学*, 2015, 30(4): 391-396.

[16] HORN P S, PESCE A J, COPELAND B E. A robust approach to reference interval estimation and evaluation[J]. *Clin Chem*, 1998, 44(3): 622.

(收稿日期: 2019-06-27 修回日期: 2019-10-15)

(上接第 300 页)

[18] 陈琛, 苏华, 蒋丽娜, 等. 血清 IL-4、NT-proBNP 及 PCT 检测在脓毒症患者预后的应用[J]. *检验医学与临床*, 2017, 14(6): 762-764.

[19] 黄朋, 方恋, 陈波, 等. 重症肺炎患者血清降钙素原、C-反应蛋白、D 二聚体及炎症因子水平变化及其临床意义[J]. *安徽医药*, 2018, 22(3): 478-482.

[20] 陈华, 周懿忆, 戴屹东. NT-proBNP 的临床应用进展[J]. *标记免疫分析与临床*, 2015, 22(1): 69-72.

[21] 张翊, 颜海青, 张茜. 血浆氨基末端 B 型脑利钠肽前体对小儿重症肺炎心功能状态的评估价值[J]. *国际儿科学杂志*, 2016, 43(8): 654-656.

[22] DUARTE J C, CASTRO A T E, SILVA R, et al. Prognostic value of plasma D-dimer level in adults with com-

munity-acquired pneumonia; A prospective study[J]. *Rev Port Pneumol*, 2015, 21(4): 218-219.

[23] 张辉, 包玉娟, 左祥荣, 等. 炎症与右心衰竭的研究进展[J]. *中华危重病急救医学*, 2018, 30(8): 815-818.

[24] 宋立成, 孟激光, 韩志海. 血栓调节蛋白在凝血功能调节及炎症反应中作用的研究进展[J]. *解放军医学院学报*, 2017, 38(10): 984-986.

[25] 郭冰冰, 路明, 武怡, 等. 重症肺炎支原体肺炎患儿血清 Toll 样受体 4/髓样分化因子 88 及 D-二聚体水平检测的意义[J]. *中华实用儿科临床杂志*, 2017, 32(22): 1709-1712.

(收稿日期: 2019-06-25 修回日期: 2019-10-16)