

• 运动人体科学 •

不同方法定量评价跑步经济性隔日测量的可靠性

王志明

(广州体育学院 田径教研室, 广东 广州 510500)

摘 要: 为了探索 RE 隔日测量结果的可靠性, 采用配对 *t* 检验、Pearson 相关系数、组内相关系数(intra-class correlation coefficients, ICC)以及 Bland-Altman 法进行验证。选择 12 名定向越野和中长跑项目学生运动员作为研究对象, 其中女生 5 名, 年龄(20.50±1.00)岁, 身高(163.01±4.27) cm、体重(52.96±3.28) kg; 男生 7 名, 年龄(20.25±0.71)岁, 身高(173.64±4.64) cm, 体重(66.57±4.60) kg。隔日测量 RE, 测试时间和测量顺序均一致, 室内气温 20 °C, 湿度 45%, 结果发现, RE 隔日测量结果配对 *t* 检验差异没有显著性, 相关性呈高度相关; 组内相关系数 ICC 结果较高, Bland-Altman 图形中 2 次 RE 测量结果差值的分布服从正态分布, 所有的差值数据均在 95% 的一致性界限以内。结果表明, RE 隔日测量可靠性比较高, 但是, 影响 RE 的因素非常复杂, 在实际测量的过程中, 要始终保证受试者在前后测量身体处于非疲劳状态, 测试前两天保证受试者进行 30 min 以上的熟悉跑台跑步, 前后两次测试顺序和测试时间尽量保持一致。

关 键 词: 运动生理学; 跑步经济性; 重测信度

中图分类号: G804.7 文献标志码: A 文章编号: 1006-7116(2014)04-0125-05

Evaluating the reliability of every other day running economy measurement quantitatively by using different methods

WANG Zhi-ming

(Department of Track and Field, Guangzhou Sport University, Guangzhou 510500, China)

Abstract: In order to probe into the reliability of the results of every other day running economy (RE) measurement, the author verified the results by using paired *t*-test, Pearson correlation coefficient, intra-class correlation coefficient (ICC) and Bland-Altman method. The author selected 12 student orienteers and middle and long distance runners as his research subjects, including 5 female students at an age of 20.50±1.00, with a height of 163.01±4.27cm and a weight of 52.96±3.28kg, 7 male students at an age of 20.25±0.71, with a height of 173.64±4.64cm and a weight of 66.57±4.60kg, measured RE at a room temperature of 20°C and a relative humidity of 45% and under the condition of consistent measurement time and order every other day, and revealed the following findings: the results of every other day RE measurement verified by using paired *t*-test had no significant difference, showing a high correlation; the results verified by using ICC were relatively high, the distribution of the differences between the 2-time RE measurement results shown in the Bland-Altman diagram presented a normal distribution, all the difference data were within the 95% consistency limit. The said findings indicate the followings: the reliability of every other day RE measurement is high; although it was found by the author that the reliability of every other day RE measurement is high, RE affecting factors are very complicated; during actual measurement, the researcher should always make sure that the testees are in a non fatigue condition physically during the two successive times of measurement, let the testees have an over 30-minute treadmill familiarizing run two days before measurement, and keep the order and time of the two successive times of measurement as consistent as possible.

Key words: sports physiology; running economy; retest reliability

收稿日期: 2015-05-19

基金项目: 国家自然科学基金青年科学基金项目(11002036); 广东省高等学校优秀青年教师培养计划资助项目(Yq2013105)。

作者简介: 王志明(1959-), 男, 副教授, 硕士, 研究方向: 体育教育训练学。通讯作者: 任占兵。E-mail: gztywzm@21cn.com

所谓跑步经济性(Running Economy, RE)是指在次极限负荷的特定速度下跑步,摄氧量达到稳定状态时每单位体重的摄氧量^[1-4]。采用摄氧量单位表达 RE 在不考虑底物代谢对能量消耗的影响时^[5],RE 的测试强度要尽可能使受试者处于有氧运动。RE 是人体功能内稳态对运动应激的一种反应^[6],研究认为,内稳态是由生物系统的各种调节机制调控而维持的一种动态平衡,是生物系统从进化适应中获得的维持整个生物系统生存的基本条件^[7]。通过 RE 评价运动员的有氧代谢能力已经比较常见^[8-9]。所谓 RE 测量可靠性,是指在相同的测量条件下,对同一批受试者使用相同的测量方法,重复测量结果的一致性程度,本研究所关注的 RE 测量可靠性是隔日测量 RE 结果的一致性。前人在定量评价 RE 测量结果一致性方面,所采用的方法也不尽相同。例如, Morgan 等^[10]采用组内相关系数(Intraclass Correlation Coefficient, ICC)和配对样本 *t* 检验的方法研究了优秀长跑运动员连续两天 RE 测量结果的稳定性,认为连续 2 天测试 RE 结果的可靠性比较高($r=0.95$),平均变异系数为 1.32%; Morgan 等^[11]选择让受试者周一到周五连续测试 5 次,通过均值的多重比较方法进一步研究了优秀长跑运动员在不同负荷下 RE 在 1 周内随时间的变化,认为 1 周的时间因素并不是对 RE 造成干扰的有效指标。江崇民等^[12]通过平板运动跑台和场地两种不同的测试方法,对我国成年男性走、跑过程中的气体代谢和能量消耗进行比较,运用了配对 *t* 检验、直线相关分析、ICC 和 Bland-Altman 图形法分析表明两种测试方法有非常显著的一致性和相关性。魏登云^[13]认为,可靠性检验的方法与测验类型有关,频数等于 1 和频数大于 1 的两类测验,其可靠性检验的方法有本质区别,认为对于“测量——再测量”方法,可靠性可以两次重复测验之间的样本相关系数 *r* 估计。袁尽州等^[14]认为,同一测量方法的信度用不同估算方法计算,结果差异很大。本研究零假设为 RE 隔日测量存在差异,采用配对 *t* 检验、Pearson 相关系数、组内相关系数(ICC)以及 Bland-Altman 法定量评价 RE 隔日测量可靠性。

1 研究对象与方法

1.1 研究对象

来自某体育学院定向越野队和中长跑队的 12 名受试者,其中女生 5 名,年龄(20.50±1.00)岁,身高(163.01±4.27)cm,体重(52.96±3.28)kg;男生 7 名,年龄(20.25±0.71)岁,身高(173.64±4.64)cm,体重(66.57±4.60)kg。为了保证 RE 测试结果不受训练疲劳的影响,在正式测试前 1 d,所有受试者停止高强度训练,但保持正常

的饮食和一般日常活动。测试前,详细说明该实验的目的、注意事项以及可能出现的问题,受试者填写知情同意书,志愿按照实验要求配合工作人员进行测试。

1.2 RE 测试方法

RE 测试仪器主要采用 MAX-II 运动心肺功能测试系统、POLAR 心率遥测系统、秒表等。RE 开始测试的时间分别是上午的 09:00 和隔日上午 09:00,室内气温 20℃,相对湿度 45%。受试者穿着自己最舒适的运动服装。在 RE 的正式测试前 3 d,每名受试者均进行了跑台跑步适应,跑台跑步适应分 2 个阶段:第 1 阶段,跑台行走熟悉阶段,受试者在跑台以 4 km/h 的速度行走 10 min,进行 3 次;第 2 阶段,跑台跑步熟悉阶段,受试者在跑台以 10 km/h 的速度跑 10 min,两个阶段累计 30 min 以上。跑台跑步适应阶段结束后隔日,开始测试 RE。根据 RE 的定义,RE 主要是指在次极限负荷的特定速度下跑步摄氧量达到稳定状态时每单位体重的摄氧量^[1-4],本研究让受试者在 10 km/h 的速度下跑 5 min,取跑台跑步最后 2 min 的摄氧量的平均值作为 RE 值。由于可靠性主要指相同条件下同一试验对相同人群重复试验获得相同结果的稳定程度,因此,本研究 2 次测试的对象、测试仪器以及 RE 的测试方法、受试者的测试顺序等都保持一致。

采用 SPSS18.0 对各项测量指标的平均数和标准差进行描述性统计,采用配对样本 *t* 检验、Pearson 相关系数、组内相关系数(ICC)以及 Bland-Altman 法分析 RE 隔日测试结果的可靠性。

2 研究结果及分析

2.1 RE 隔日测量配对样本 *t* 检验

表 1 是隔日测量 RE 的结果,从表 1 可见,受试者初次测量的 RE 结果为(33.54±3.57) mL·kg⁻¹·min⁻¹,略低于隔日 RE 测量的结果(34.06±3.18) mL·kg⁻¹·min⁻¹,初次测量的 *V*(O₂)绝对值(2 080.00±392.52) mL·min⁻¹,也低于隔日所测量的 *V*(O₂)绝对值(2 116.46±398.33) mL·min⁻¹,但进一步对初次和隔日 RE 测量的相对值和绝对值进行配对样本 *t* 检验发现,当显著性水平 α 为 0.05 时,初次和隔日 RE 测量结果差异并不具有显著性(见表 2)。

2.2 RE 隔日测量相关性

通常情况下,当相关系数 $|r|$ 大于 0.8 的时候,认为两个变量的线性相关性很强。两次 RE 测量结果的相关性为 0.82,两次绝对值测量结果的相关系数为 0.95,均大于 0.8,说明两次 RE 测量结果呈高度相关。从初次与隔日 RE 测量结果相关性散点图(见图 1)也可以说明两次 RE 测试结果呈高度相关。

表1 隔日测量 RE 代谢指标统计结果

测量时间	n/人	RE(mL·kg ⁻¹ ·min ⁻¹)			V(O ₂)(mL·min ⁻¹)		
		平均值	标准差	标准误	平均值	标准差	标准误
初次	12	33.54	3.57	1.03	2 080.00	392.52	113.31
隔天	12	34.06	3.18	0.92	2 116.46	398.33	114.99

表2 隔日测量 RE 及 V(O₂) 差异的配对 t 检验结果

配对	平均	标准差	标准误	95%置信区间		t 值	自由度	Sig.(2-tailed)
				下限	上限			
初次&隔日 RE	-0.52	2.08	0.60	-1.84	0.80	-0.86	11.00	0.41
初次&隔日 V(O ₂)	-36.46	124.56	35.96	-115.60	42.68	-1.01	11.00	0.33

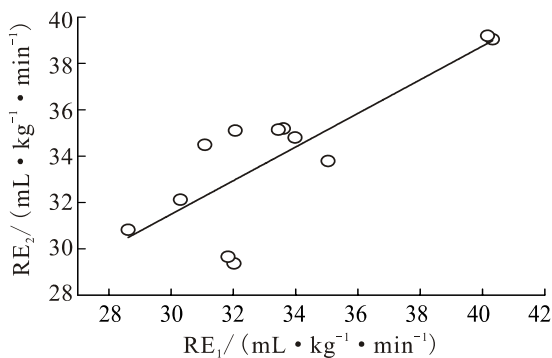


图1 RE 两次测量结果的相关性散点图

2.3 RE 隔日测量组内相关系数

组内相关系数(ICC)最先由 Bartko^[15]用于测量和评价信度大小,RE 和 V(O₂)的 ICC 分别为 0.895(95%置信区间为 0.636~0.969)、0.975(95%置信区间为 0.912~0.993), ICC 值介于 0~1 之间,0 表示不可信,1 表示完全可信。0.90~0.99 表示具有极好信度、0.80~0.89 表示具有良好信度,0.70~0.79 表示具有中等信度,小于 0.69 表示信度较差。本研究 RE 和 V(O₂)的 ICC 值均大于 0.89,说明这两个指标的重测信度良好。

2.4 RE 隔日测量的 Bland-Altman 图形结果

Bland-Altman 图形分析也可以用来定量评价测量的一致性^[6]。显示 RE 两次测量绝对误差为 0.5 mL·kg⁻¹·min⁻¹,一致性界限为 3.6~4.6 mL·kg⁻¹·min⁻¹。RE 两次测量平均差异值分布显示没有系统性误差(见图 2)。

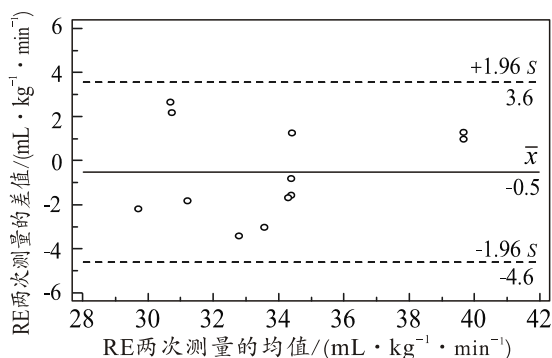


图2 RE 两次测量均值的系统性误差

V(O₂)两次测量绝对误差为-36.5 mL·min⁻¹,一致性界限为 207.7~280.6 mL·min⁻¹,V(O₂)两次测量平均差异值分布显示没有系统性误差(见图 3)。

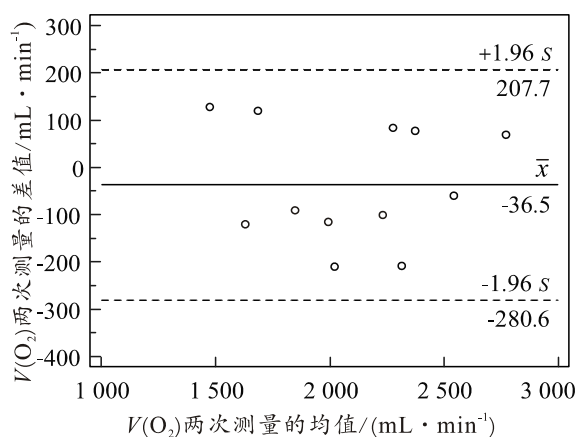


图3 RE 两次测量均值的系统性误差

3 讨论

研究发现,影响 RE 的因素是一个比较复杂的系统^[17-18],受试者的疲劳状态、饮食、服装以及室温和湿度等都有可能影响 RE 测量的信度^[4]。对样本均数的 t 检验又称配对 t 检验,适用于配对设计的计量资料均数的比较,其比较的目的是检验两样本均数所代表的未知总体均数是否有差异^[19]。当显著性水平 α 为 0.05 时,本研究初次和隔日 RE 测量结果差异并不具有显著性。两次 RE 测量结果的配对 t 检验主要检验的是 RE 测量的系统误差差别是否显著,并不能对随机误差的差异进行测量。而根据统计学知识^[20],随机误差对 RE 均数差异的比较影响不大,因此,配对 t 检验对随机误差的敏感度较低,如果本研究中假设前后测量 RE 的方法无差别,而个体 RE 测量结果的差异显著时,采用配对 t 检验来验证 RE 测量方法的一致性不准确。综上可见,RE 测量结果的差异不显著,说明两次 RE 反应的总体均数可能相同,但不能说明两次 RE 测量方法的一致性较好。

相关性分析的零假设为两次 RE 测量结果不呈线

性相关, 本研究结果可见, 两次 RE 测量结果的相关系数 r 为 0.895, 说明拒绝零假设, 两次 RE 测量结果呈高度线性相关。然而, 相关分析只是检验数据同步变化的方向与紧密程度, 即一个变量(首次测试结果 RE)随着另一个变量(隔日测试结果 RE)呈现出线性变化规律。然而, 较高的相关性并不代表着两次 RE 测量结果的一致性。因为 Bland 等^[16]认为, 相关性分析是测量两个变量之间的相关性强度, 并不是两个变量之间的一致性程度; 测量的刻度比例对相关性没有影响, 但是影响测量的一致性; 相关系数受到样本数量的分布范围影响, 而一致性不受样本数量分布范围影响, 当测量样本数量的分布范围很小, 即被测样本数量个体间的变异与测量仪器间变异相当时, 很难观察到两测量样本数据的线性关系; 测量的显著性对相关性也有影响, 而与一致性没有什么关系; 较差的测量一致性可能会产生出较高的相关性^[21]。

研究表明, 在配对 t 检验、相关性系数以及组内相关系数 3 种方法中, 组内相关系数是评价一致性最为理想的指标, 因为组内相关系数对测量的系统误差和随机误差均敏感^[20, 22]。但是, 当把所有观测值范围变小, 相当于只对某一局限范围的个体进行测量时, 组内相关系数有时也会做出错误的判断。本研究中, RE 和 $V(O_2)$ 的 ICC 值均大于 0.89, 说明隔日测量跑步经济性的可靠性较好。

Bland-Altman 图形法^[23]是定量分析与定性分析的有机结合。该方法的基本思想是, 利用原始数据的均值与差值, 分别以均值为横轴, 以差值为纵轴做散点图, 计算差值的均数以及差值的 95% 分布范围(一致性界限), 认为应该有 95% 的差值位于该一致性界限以内。一般认为, 图形中的点位于一致性界限范围内的要占到所有点的 95%, 同时还要考虑不超出专业上可接受的临界值范围。满足这两点一般即可认为两种方法的一致性较好, 可以互换。李镒冲等^[20]认为, Bland-Altman 法在评价一致性的时候既考虑了随机误差同时也考虑了系统误差对一致性的影响, 同时可结合专业意义进行判断, 具有独特的优势。配对 t 检验与简单相关分析具有明显的片面性, 不能同时兼顾随机误差与系统误差, 用它们评价一致性所得的结论可能是误导的。在本研究中, 图 2 和图 3, 两次 RE 测量结果差值的分布服从正态分布, 所有的差值数据均位于 $d-1.96s$ 和 $d+1.96s$ 之间, 这个区间被称之为 95% 的一致性界限, 由于 RE 差值均位于 95% 的一致性界限以内, 因此, RE 隔日测量具有较好的一致性。

本研究结论否定零假设, 即认为 RE 隔日测量具有较好的可靠性, 通过不同方法的分析讨论, 本研究

发现, 采用配对 t 检验、Pearson 相关系数、组内相关系数 ICC 以及 Bland-Altman 图形法定量评价 RE 隔日测量可靠性方面, 并没有发现哪一种方法否定 RE 隔日测量结果的一致性, RE 隔日测量结果配对 t 检验差异不显著, 相关性呈高度相关, 组内相关系数 ICC 结果较高, Bland-Altman 图形中两次 RE 测量结果差值的分布服从正态分布, 所有的差值数据均在 95% 的一致性界限以内。

刘承宜等^[7]认为, 在运动训练中存在大量的内稳态现象。RE 充分反映了人体有效的利用内稳态抵抗外界环境干扰的能力, RE 也是机体对外界或内部各种刺激所产生的非特异性应答反应的总和, RE 是个体内稳态生理学因素和心理学因素的整体反映, 当外界自然环境处于稳态(外稳态), 跑步速度由 0 转为 10 km/h 时, 人的机体应激反应随着时间的变化也必然由一种稳态转化为另一种稳态(内稳态)。RE 同时反映了组织细胞在特定速度下所能消耗或利用氧气的的能力, 如果 RE 的测试负荷与运动负荷一样, 底物代谢对 RE 的干扰会很大, 测试的 RE 如果没有考虑底物代谢的影响, 则 RE 并不能完全反应运动经济性。本研究选择 $10 \text{ km} \cdot \text{h}^{-1}$ 的速度保证了学生耐力运动员在此速度下处于有氧运动的状态, 通过不同方法定量评价 RE 隔日测量可靠性发现, 人体对跑步速度的应激反应比较稳定, 人体内稳态会随自然环境的外稳态而发生变化, 并且进一步形成新的内稳态。

通过讨论本研究认为, 在实际定量评价 RE 测量可靠性方面, 应该重点以组内相关系数和 Bland-Altman 图形法的结果为重点参考。虽然本研究发现隔日测量 RE 的可靠性比较高, 但是, 影响 RE 的因素非常复杂, 在实际测量的过程中, 要始终保证受试者在前后测量身体处于非疲劳状态, 测试前两天保证受试者进行 30 min 以上的熟悉跑台跑步阶段, 前后两次测试顺序和测试时间尽量保持一致等。

参考文献:

- [1] Conley D L, Krahenbuhl G S. Running economy and distance running performance of highly trained athletes[J]. Med Sci Sports Exerc, 1980, 12(5): 357-360.
- [2] Anderson T. Biomechanics and running economy[J]. Sports Med, 1996, 22(2): 76-89.
- [3] Conley D L, K G B L. Following Steve Scott: physiological changes accompanying training[J]. Phys Sports Med, 1984, 12: 103-106.
- [4] Morgan D W, Craib M. Physiological aspects of running economy[J]. Med Sci Sports Exerc, 1992, 24(4):

456-461.

- [5] 郜卫峰. 利用能量单位评价跑步经济性的有效性研究[J]. 体育科学, 2012, 32(5): 49-57.
- [6] 刘承宜, 任占兵. 生理经济学[J]. 华南师范大学学报: 自然科学版, 2011(2): 1-6.
- [7] 刘承宜, 袁建琴, 付德荣, 等. 以赛带练的内稳态研究[J]. 体育学刊, 2008, 15(5): 81-84.
- [8] Paavolainen L, Hakkinen K, Hamalainen I, et al. Explosive-strength training improves 5-km running time by improving running economy and muscle power[J]. J Appl Physiol, 1999, 86(5): 1527-1533.
- [9] Lake M J, Cavanagh P R. Six weeks of training does not change running mechanics or improve running economy[J]. Med Sci Sports Exerc, 1996, 28(7): 860-869.
- [10] Morgan D W, Martin P E, Krahenbuhl G S, et al. Variability in running economy and mechanics among trained male runners[J]. Med Sci Sports Exerc, 1991, 23(3): 378-383.
- [11] Morgan D W, Craib M W, Krahenbuhl G S, et al. Daily variability in exercise ventilation[J]. Respir Physiol, 1994, 96(2-3): 345-352.
- [12] 江崇民, 邱淑敏, 王欢, 等. 平板运动跑台和场地环境测试走、跑运动能量消耗的比较研究[J]. 体育科学, 2011, 31(7): 30-36.
- [13] 魏登云. 体育测量可靠性检验的基本思想与方法[J]. 体育科学, 2001, 21(3): 85-88.
- [14] 袁尽州, 朱长跃, 刘芦萍. 几种测量信度估价方法的比较研究[J]. 西安体育学院学报, 2003, 20(4): 64-66.
- [15] Bartko J J. The intraclass correlation coefficient as a measure of reliability[J]. Psychol Rep, 1966, 19(1): 3-11.
- [16] Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement[J]. Lancet, 1986, 1(8476): 307-310.
- [17] Morgan D W, Martin P E, Krahenbuhl G S. Factors affecting running economy[J]. Sports Med, 1989, 7(5): 310-330.
- [18] Saunders P U, Pyne D B, Telford R D, et al. Factors affecting running economy in trained distance runners[J]. Sports Med, 2004, 34(7): 465-485.
- [19] 颜虹, 徐勇勇, 赵耐青, 等. 医学统计学[M]. 北京: 人民卫生出版社, 2005: 122.
- [20] 李镒冲, 李晓松. 两种测量方法定量测量结果的一致性评价[J]. 现代预防医学, 2007(17): 3263-3266.
- [21] Altman D G. Estimation of gestational age at birth-comparison of two methods[J]. Arch Dis Child, 1979, 54(3): 242-243.
- [22] 潘晓平, 倪宗瓚. 组内相关系数在信度评价中的应用[J]. 华西医科大学学报, 1999(1): 62-63.
- [23] 萨建, 刘桂芬. 定量测量结果的一致性评价及Bland-Altman法的应用[J]. 中国卫生统计, 2011(4): 409-411.

