Critical Difference Applied to Exercise-Induced Salivary Testosterone and Cortisol Using Enzyme Linked Immunosorbent Assay (ELISA): Distinguishing Biological from Statistical Change

Lawrence D Hayes[a,b] · Nicholas Sculthorpe[a] · John D Young[c] · Julien S Baker[a] · Fergal M Grace[a]

[a] *Institute of Clinical Exercise and Health Science, University of the West of Scotland*

[b]*School of Human Sciences, London Metropolitan University*

[c]*Department of Life and Environment, University of the West of Scotland*

Corresponding Author:

Lawrence Hayes

School of Human Sciences

London Metropolitan University

166-220 Holloway Road

London N7 8DB

020 7133 4686

E-mail: L.Hayes@Londonmet.ac.uk

Abstract Due to its non-invasive, convenient, and practical nature, salivary testosterone (sal-T) and cortisol (sal-C) are frequently used in a clinical and applied setting. However, few studies report biological and analytical error and even fewer report the 'critical difference' which is the change required before a true biological difference can be claimed. It was hypothesized that (a) exercise would result in a statistically significant change in sal-C and sal-T and (b) the exercise-induced change would be within the critical difference for both salivary hormones. In study 1 we calculated the critical difference of sal-T and sal-C of eighteen healthy adult males aged 23.2 ± 3.0 years every 60 min in a seated position over a 12 h period (08:00 – 20:00 h [study 1]). As proof-of-concept, sal-C and sal-T was also obtained pre and at 5 and 60 min post a maximal exercise protocols in a separate group of seventeen healthy males (aged 20.1 ± 2.8 years [study 2]). The critical difference of sal-T calculated as 90%. For sal-C, the critical difference was 148% (Study 1). Maximal exercise was associated with a statistically significant ($p<0.05$) changes in sal-T and sal-C. However, these changes were all within the critical difference range. Results from this investigation indicate that a large magnitude of change for sal-C and sal-T is required before a biologically significant mean change can be claimed. Studies utilizing sal-T and sal-C should appreciate the critical difference of these measures and assess the biological significance of any statistical changes.

Keywords

Randomized crossover trial · Testosterone · Cortisol · Saliva · Exercise · Critical difference

Introduction

Acute elevations in plasma testosterone (T) have been associated with chronic adaptations to exercise training [7, 17]. The significance of this however, remains to be determined [16, 22, 23]. In addition, significant changes in blood cortisol (C) have also been reported following aerobic, anaerobic and resistance exercise modes [3, 11, 13, 15]. C either alone or in combination with T has been suggested to be an indicator of overtraining. In fact, threshold reductions in the T:C ratio of 30% have been suggested as a criteria to screen for overtraining [14].

Since the initial reporting of salivary hormones in marathon runners in the mid-1980's [2], measurement of salivary T (sal-T) and C (sal-C) has been widely used in the research literature with the assumption that they reflect circulating blood hormone concentrations. Saliva can provide a useful, non-invasive alternative to the collection of serum and plasma because it allows quantative measurement of T and C without the invasiveness and stress response associated with venipuncture, or the costs of specialist training and laboratories. As a result, sal-T and sal-C have also been used as a determinant of anabolic status and maladaptation in sportsmen [14]. The concentration of these hormones as indicators of overtraining have been demonstrated in football [4], cycling [8], swimming [9], and rowing [21], with authors suggesting differences between sal-T and sal-C reflect disturbance in the anabolic-catabolic balance.

Understanding the strengths and limitations of current analytical practices surrounding sal-T and sal-C measurement such as sensitivity, specificity, accuracy and standardization of procedures whilst simultaneously appreciating the level of biological variation of sal-T and sal-C allows the determination of minimum thresholds, beyond which, a true change has occurred. Concentrations of T and C meterands oscillate around an individual's homeostatic set-point and are known as the within-person 'biological variation'. When biological variation is combined with preanalytical (standardization of procedures) and analytical error (precision and bias), then the total error for any given biological measurement can be obtained.

Appreciation and application of consistencies in practices known to influence the meterand, timing of sampling, transport, handling and storage of samples allow the minimization of preanalytical error and when specific analytical variation has been identified, then serial measurements can be obtained to identify whether biological changes have occurred. In order to achieve meaningful data, it should be an *a priori* aim of both clinical practice and applied research, to control preanalytical and analytical variability, thus allowing the biological component to be the only true variable.

There are numerous potential sources of biological error in sal-T and sal-C studies including; diurnal variation, age, gender, long haul travel, diet, sitting to standing, physical activity, smoking, sexual activity, illness, stress, ethic difference, and non-traditional supplementation, that should be appreciated when attempting to minimize critical difference (CD) values. First outlined by Fraser and Fogarty [5] The 'critical difference', sometimes termed the 'clinically relevant difference', is a method of incorporating both analytical error and biological variation in order to determine biologically relevant 'least significant difference' thresholds considered to be clinically meaningful. The procedure outlined by Fraser and Fogarty [5] allows the amount of analytical error added to the biological variability to be calculated from the ratio of the analytical coefficient of variation (CV_A) to the biological (within person) coefficient of variation (CV_B) in determination of the CD.

Given the importance of establishing biologically meaningful measurement thresholds, the primary aim of this investigation (study 1) was to determine the critical difference of sal-C and sal-T using enzyme-linked immunosorbent assay (ELISA). A secondary (study 2) aim was to compare changes in sal-C and sal-T obtained before, and after exercise with the critical difference values. It was hypothesized that (a) exercise would result in a statistically significant change in sal-C and sal-T and (b) the exercise-induced change would be within the critical difference for both salivary hormones.

Materials and Methods

Study 1

Eighteen male university students age, stature, and body mass of 23.2 ± 3.0 years of age, 180.9 ± 4.3 cm in height, 84.4 ± 15.9 kg in body mass volunteered to participate in Study 1. Experimental procedures were approved by the University of the West of Scotland Research Ethics Committee. The protocol was explained and participants provided written informed consent to participate in this study. All participants were habitually physically active, and had abstained from alcohol, caffeine and exercise for 24 h preceding the investigation. Participants were permitted to drink water *ad libitum* (up to 10 min prior to salivary sample) as prescribed by manufacturer guidelines (Salimetrics, State College, PA). Exclusion criteria included poor sleep quality, recent shift work, extreme chronotype according to the Horne-Ostberg Morningness-Eveningness Questionnaire [10] or travel across multiple time zones.

Fasted participants reported to the laboratory at ~07:45 h approximately 40 min after waking. Laboratory observations were conducted in the University of the West of Scotland's Clinical Exercise Research Laboratory. The ~40m^2 room was cleared of time reference devices and blinded, ensuring no natural light entered the room. Artificial lighting and a constant temperature was maintained. Sample collection commenced at 08:00 h and was completed every 60 min until 20:00 h. Immediately after the first sampling, participants were provided with a standard breakfast (1,769 kJ, 18% protein, 9% fat, 73% carbohydrate). All meals were provided by the research team to control for any differences in dietary intake. Participants remained in the study venue until ~20:05 h and provided a saliva sample every 60 min. At 13:05 and 18:05 h participants were provided with a standardized meal (2,721 kJ, 24% protein, 22% fat, 54% carbohydrate per meal). Participants were permitted to drink water ad libitum (up to 10 min prior to salivary sample) and were instructed to rinse their mouths with water after eating. Participants were free to engage in sedentary activities of their choice (work, watching films without sexual or aggressive content, and listening to music). Physical activity and exercise were not allowed. Participants left the laboratory only to use the toilet (also light controlled).

Study 2

Seventeen male university students 20.2 ± 2.8 years, with a stature of 177.3 ± 7.3 cm, and body mass of 74.7 ± 12.4 kg participated in Study 2. Participants woke up at ~07:30 h and consumed a basic breakfast (cereal or toast) according to self-reporting. After three initial familiarization trials, participants reported to the laboratory at ~08:30 or ~16:30 h ± 5 min. Participants exacted the same preparatory guidelines, conditions and exclusion criteria as described in study 1. A randomized crossover trial design was used to assign time of day and mode of exercise for each participant and each session. Participants rested in a seated position for 5 min to allow baseline measures to be collected, including resting sal-T and sal-C. Once baseline sampling was completed, uniform and mode specific warm-ups preceded testing. Performance testing occurred at either 09:00 h or 17:00 h. Participants completed a 5 m maximal sprint and a 1 repetition maximum (RM) squat at each time of day for a total of four testing sessions. Testing sessions were separated by at least 48 h in order to allow complete recovery between maximal efforts.

Before 1RM testing, participants performed a series of submaximal warm-up sets of eight, five, two, and one repetitions with increasing load. This procedure was constant across both 1RM testing sessions. All measurements were taken by the same experienced administrator to ensure consistency. 1RM was assessed on back squat. Briefly, an Olympic standard barbell (York Barbell Company, York, PA) was placed above the posterior deltoids at the base of the neck. Participants lowered their body downward, bending at the hips, knees, and ankles until the top of the thighs were parallel to the floor then recovered to standing. Participants rested for at least 4 min between each of the 1RM trials. On average, the 1RM was determined within six trials. Failure was defined as a lift falling short of the full range of motion in two sequential attempts spaced 4 min apart. Saliva samples were collected pre, 5 min post, and 60 min post exercise.

Saliva collection and analysis

Whole salivary samples of approximately 1.8 mL were collected via expectoration into graduated 2 ml cryovials (Salimetrics, State College, PA). To prevent potential blood contamination of saliva, resulting in an overestimation of hormone concentrations, participants were advised to avoid brushing

their teeth and drinking hot fluids 2 h prior to reporting to the study venue. Salivary samples were collected and transported to a freezer immediately where they remained at -80ºC until assay (within one month). Samples were assayed in duplicate (without separation or extraction) for sal-T and sal-C using commercially available ELISA protocols according to manufacturer's guidelines (Salimetrics, State College, PA). To minimize biological error, critical difference was calculated for samples collected within seconds of each other (and compared to daily values). Participants were all Caucasian males of a similar age. Saliva was collected mid flow, without the stimulation of gum or juice and participants remained seated throughout sampling. All samples were analyzed according to manufacturer's instructions and in duplicate along with appropriate standard controls from the same company. The intra-assay and interassay variability were below 7% in all instances.

Statistical analysis

Data were analyzed using parametric statistics following mathematical confirmation of a normal distribution by repeated Shapiro–Wilks test. Between-hour differences were analysed using a one-way ANOVA with *a posteriori* Tukey test. Alpha was established at $p<0.05$, and data are reported as mean ± standard deviation (SD).

Coefficient of variation: For the purposes of determining CV_A and CV_B, CV is defined as the standard deviation (SD) divided by the mean of each measurement.

Analytical variation (CV_A)

Analytical variation (CV_A) was calculated as a percentage using pooled mean ± SD of each duplicate measure using the following equation:

$CV_A = 100 \cdot (SD \div mean)$

Biological variation (CV_B)

Biological variation (CV_B) for sal-T and sal-C was calculated using mean ± SD of 13 samples from each participant and the following equation:

$CV_B = 100 \cdot (SD \div mean)$

For example, the group mean sal-T from 08:00 – 20:00 h in study 1 was 137.3 ± 53.2 $pmol \cdot L^{-1}$. Therefore, $CV_B = 100 \cdot (53.2 \div 137.3) = 38.7\%$.

Within subject variation (CV_W)

Assay data from each subject, collected at periodic times (thirteen specimens) throughout a single day, which were used to calculate within subject biological coefficient of variation (CV_W) using the following equation:

$CV_W = CV_B(\%) - CV_A(\%)$

Critical difference

The critical difference was assessed using the equation of Fraser and Fogarty [5]:

Critical difference = $k\sqrt{CV_A^2 + CV_W^2}$

Where:

k = constant determined by the probability level (2.77 at $P<0.05$)

CV_A = coefficient of analytical variation.

CV_W = coefficient of within subject variation.

Results

Study 1

Table 1 summarizes the critical difference and its residuals (CV_A and CV_B) for sal-T, sal-C, and sal-T:C. It was observed that the highest critical difference was obtained for sal-T:C ratio as a result of combined error of sal-T and sal-C.

****INSERT TABLE 1 NEAR HERE****

Table 1 Biological variation and critical difference for sal-T, sal-C, and sal-T:C. CV_A = coefficient of analytical variation. CV_B = coefficient of biological variation. Sal-T:C = saliva testosterone/cortisol. CV_A for sal-T:C ratio was estimated using the mean of CV_A for sal-T and sal-C at that time point (*n*=18).

Study 2

****INSERT TABLE 2 NEAR HERE****

Table 2 Mean salivary hormone concentrations during the investigation (* denotes significant difference from pre exercise values).

Elevated salivary hormone concentrations were observed in the 09:00 h trials during both exercise modes ($p<0.05$, Table 2). Time of day had a significant effect on sal-C and sal-T with higher values present in the morning ($p<0.05$) however, no significant differences were observed for exercise type. There was a reduction in mean sal-C concentrations 60 min post exercise compared to pre and 5 min post values ($p<0.01$). Tukey's post hoc analysis revealed that this was only significant for 09:00 h tests.

Discussion

The main findings of the present study are that the critical difference for sal-T and sal-C, in controlled laboratory conditions, are large and that this range is extended when ratios of the two are analyzed. The data also demonstrate the potential for statistically significant changes in salivary measures in the absence of biological thresholds for meaningful changes being reached.

The results reported here are similar to those of Valero-Politi and Fuentes-Arderiu [20] who reported a critical difference of 78% for sal-T. However, the discrepancy in results (78 % vs. 90%) may be attributed to difference in methodologies; in the present investigation, we categorized CV_B as

the variation over a period of ~12 h whereas Valero-Politi and Fuentes-Arderiu [20] measured yearly variation at monthly intervals but only within a 90 min daily window. More importantly, the present investigation utilized ELISA whereas Valero-Politi and Fuentes-Arderiu [20] used radioimmunoassay (RIA), known to exhibit smaller CV_A [18].

The present study describes a CV_B of 60% for sal-C, somewhat higher than sal-T (39%). This is in line with previous investigations that report greater diurnal variation in cortisol compared to testosterone [12, 19].

The conditions of the present investigation were highly controlled (diet, light, physical activity, age, gender, ethnicity, smoking status, posture, travel, and illness), yet we still reported a high critical difference in a stepwise manner (sal-T:C > sal-C > sal-T). We reason that in an applied setting, where biological variance is known to be greater, (e.g. team sports, stressful situations, smokers, the aging, and individuals with high physical activity levels) then even greater alterations will be required to confirm a biologically meaningful change. With this in mind, previous investigations that have reported statistically significant post-exercise changes in sal-T and sal-C may have fallen victim to classic Type I error. Indeed, this is further supported by the findings from study 2, which demonstrate that an acute exercise session can generate statistically significant fluctuations in sal-C and sal-T whilst falling short of reaching the threshold of biological relevance. Sal-C decreased by 58% 60 min after the sprint protocol at 09:00 h, however, this is again within the critical difference defined in Study 1 (~147%). Similarly, sal-T increased by 19% 5 min following a sprint protocol at 17:00 h, again this was within the critical difference defined in Study 1 for sal-T (~90%). The authors propose an exercising of caution when attempting to make conclusions based solely on changes to sal-C and sal-T using enzyme immunoassay. Additionally, caution should be taken when using sal-C and sal-T concentrations in combination as a surrogate measure normally established using sera. For example, the proposed definition of overtraining was characterized by a 30% change in the systemically derived T:C ratio in serum [1]. However, some authors employ sal-T and sal-C as surrogates in the diagnosis [14]. A 30% change in serum concentrations of T and C are likely to be more meaningful than a 30% change in saliva concentrations due to differences in analytical and

biological variations [6, 20]. Our findings suggest that salivary measures exhibit a greater critical difference and are more inherently variable than serum and therefore offers limited scope for application in clinical diagnostics.

In conclusion, the findings from the present study identify the critical difference values for sal-T and sal-C to be 90% and 148% respectively in laboratory-controlled conditions. This magnitude of analytical and biological error is likely to have significant impact on their interpretive capacity. The findings further promote exercising of caution when using sal-T, sal-C, and their ratio as surrogate measures of systemic T and C concentrations both in clinical diagnostics and applied exercise settings.

References

1. Adlercreutz H, Harkonen M, Kuoppasalmi K, Naveri H, Huhtaniemi I, Tikkanen H, Remes K, Dessypris A, Karvonen J (1986) Effect of training on plasma anabolic and catabolic steroid-hormones and their response during physical exercise. Int J Sports Med 7:27-28

2. Cook NJ, Read GF, Walker RF, Harris B, Riad-Fahmy D (1986) Changes in adrenal and testicular activity monitored by salivary sampling in males throughout marathon runs. Eur J Appl Physiol 55:634-638

3. Elloumi M, Maso F, Michaux O, Robert A, Lac G (2003) Behaviour of saliva cortisol C , testosterone T and the T/C ratio during a rugby match and during the post-competition recovery days. Eur J Appl Physiol 90:23-28

4. Filaire E, Lac G, Pequignot JM. Biological, hormonal, and psychological parameters in professional soccer players throughout a competitive season. Percept Motor Skill 97:1061-1072

5. Fraser CG, Fogarty Y (1989) Interpreting laboratory results. Brit Med J (Clinical research ed) 298:1659-1660

6. Granger DA, Shirtcliff EA, Booth A, Kivlighan KT, Schwartz EB (2004) The "trouble" with salivary testosterone. Psychoneuroendocrinol 29:1229-1240

7. Hansen S, Kvorning T, Kjaer M, Sjogaard G (2001) The effect of short-term strength training on human skeletal muscle: the importance of physiologically elevated hormone levels. Scand J Med Sci Spor 11:347-354

8. Hoogeveen AR, Zonderland ML (1996) Relationships between testosterone, cortisol and performance in professional cyclists. Int J Sports Med 17:423-428

9. Hooper SL, MacKinnon LT, Gordon RD, Bachmann AW (1993) Hormonal responses of elite swimmers to overtraining. Med Sci Sports Exerc 25:741-747

10. Horne JA, Ostberg O (1976) A self assessment questionnaire to determine morningness eveningness in human circadian rhythms. Int J Chronobiol 4:97-110

11. Hough J, Corney R, Kouris A, Gleeson M (2013) Salivary cortisol and testosterone responses to high-intensity cycling before and after an 11-day intensified training period. J Sports Sci 31:1614-1623

12. Kraemer WJ, Loebel CC, Volek JS, Ratamess NA, Newton RU, Wickham RB, Gotshalk LA, Duncan ND, Mazzetti SA, Gomez AL, Rubin MR, Nindl BC, Hakkinen K (2001) The effect of heavy resistance exercise on the circadian rhythm of salivary testosterone in men. Eur J Appl Physiol 84:13-18

13. Luger A, Deuster PA, Kyle SB, Gallucci WT, Montgomery LC, Gold PW, Loriaux DL, Chrousos GP (1987) Acute hypothalamic pituitary-adrenal responses to the stress of treadmill exercise - physiological adaptations to physical-training. New Eng J Med 316:1309-1315

14. Maso F, Lac G, Filaire E, Michaux O, Robert A (2004). Salivary testosterone and cortisol in rugby players: correlation with psychological overtraining items. Br J Sports Med 38:260-263

15. McGuigan MR, Egan AD, Foster C (2004). Salivary cortisol responses and perceived exertion during high intensity and low intensity bouts of resistance exercise. J Sport Sci Med 3: 8-15

16. Mitchell CJ, Churchward-Venne TA, Bellamy L, Parise G, Baker SK, Phillips SM (2013) Muscular and systemic correlates of resistance training-induced muscle hypertrophy. PloS one 8:e78636

17. Ronnestad BR, Nygaard H, Raastad T (2011) Physiological elevation of endogenous hormones results in superior strength training adaptation. Eur J Appl Physiol 111:2249-2259

18. Shrivastav TG, Basu A, Kariya KP (2003). One Step Enzyme Linked Immunosorbent Assay for Direct Estimation of Serum Testosterone. J Immunoassay mmunochem 24:205-217

19. Teo W, McGuigan MR, Newton MJ (2011) The effects of circadian rhythmicity of salivary cortisol and testosterone on maximal isometric force, maximal dynamic force, and power output. J Strength Cond Res 25:1538-1545

20. Valero-Politi J, Fuentes-Arderiu X (1993) Within- and between-subject biological variations of follitropin, lutropin, testosterone, and sex-hormone-binding globulin in men. Clin Chem 39:1723-1725

21. Vervoorn C, Quist AM, Vermulst LJM, Erich WBM, Devries WR, Thijssen JHH (1991) The behavior of the plasma-free testosterone cortisol ratio during a season of elite rowing training. Int J Sports Med 12: 257-263

22. West DWD, Kujbida GW, Moore DR, Atherton P, Burd NA, Padzik JP, De Lisio M, Tang JE, Parise G, Rennie MJ, Baker SK, Phillips SM (2009) Resistance exercise-induced

increases in putative anabolic hormones do not enhance muscle protein synthesis or intracellular signalling in young men. J Physiol 587:5239-5247

23. West DWD, Phillips SM (2012) Associations of exercise-induced hormone profiles and gains in strength and hypertrophy in a large cohort after weight training. Eur J Appl Physiol 112:2693-2702