Original Paper DOI: 10.5604/20831862.1003485 Biol. Sport 2012;29:211-215

RELATIVE AND ABSOLUTE RELIABILITY OF KARATE SPECIFIC AEROBIC TEST (KSAT) IN EXPERIENCED MALE ATHLETES

Accepted for publication 11.06.2012

AUTHORS: Chaabène H.[1], Hachana Y.[1-3], Attia A.[1-3], Mkaouer B.[3], Chaabouni S.[1], Chamari K.[2-3]

[1] Research Unit, Analysis and Evaluation of factors affecting the sport performance, Higher Institute of Sports and Physical Education, Ksar said, Tunisia
[2] Tunisian research laboratory ''Sport Performance Optimisation'', National Centre of Medicine and Science in Sport, Tunisia
[3] Higher Institute of Sports and Physical Education, ISSEP Ksar-Saïd, Manouba University, Tunis, Tunisia

Reprint request to:
Helmi Chaabane
Boucha, Azmour, Kélibia, Nabeul
8035 Tunis, Tunisia
E-mail: chaabanehelmi@hotmail.fr

ABSTRACT: The present investigation was conducted to evaluate the relative and absolute reliability and the minimal detectable change ($MDC_{95\%}$) of the karate specific aerobic test (KSAT) in male karate athletes. Sixteen subjects (age: 22.2±3.7 years, height: 175.7±6.7 cm, body mass: 72.7±7.2 kg and karate experience: 9.2±4.5 years) participated in this study and performed the KSAT twice (test-retest). Reliability of time to exhaustion (TE) was very good, with intraclass correlation coefficient ICC(3.1)>0.90, standard errors of measurement SEM<5% and 95% limits of agreement (LOA) -15.8 ± 74.7 s. The $MDC_{95\%}$ of the KSAT was 81.42 s, rated as "satisfactory". The dependent t test showed no significant difference between peak heart rate (HR_{peak}), peak lactate concentration (La_{peak}) and rating of perceived exertion (RPE) between test and retest ($P>0.05$). Effect size for all variables was $dz\leq0.2$, rated as trivial. Thus, when using the present test as a measure of aerobic fitness, a change greater than 9.4% (81.42 s) is necessary to be confident (at the 95% level) that the change in time to exhaustion reflects improvement and exceeds the measurement error. The KSAT is a reliable tool that can be used by practitioners for assessing aerobic fitness in experienced karate athletes.

KEY WORDS: field-specific testing, aerobic performance, absolute and relative reliability, karate

INTRODUCTION

Karate was developed as a martial art in Japan, and it has become a very popular sport over the last couple of decades in the rest of the world. Typical karate training consists of the basics, *kata* and *kumite*. *Kata* are set forms in pre-established sequences of offensive and defensive techniques and specifically movements [10]. *Kumite* is free sparring with execution of defensive and offensive techniques between two athletes, following karate rules [7,23]. However, a valid description of its physiological characteristics has been very complicated to achieve for several reasons: the difficulty of quantifying the effort in this intermittent sport in which the work rate is distributed into periods of high intensity and low intensity; and the differences between two opponents regarding anthropometric characteristics, technical–tactical level and in specific technical actions performed [10].

Officially, the total time of a *kumite* competition consists of three min for male senior athletes and two min for female senior, junior and cadet divisions [30]. Ravier and Rouillon [22] described a mean male combat time of 222 ± 73.72 s, which represents 62% of total combat time. The ratio of effort-pause was 13.65 ± 3.24 s to 8.83 ± 3.25 s, respectively. Additionally, previous studies observed that during an entire *kumite* the aerobic system is the major contributor (77.8 ± 5.8%) to the total energy produced. Regarding attacks periods, the phosphagen system (ATP-PCr) is the main mechanism employed (16 ± 4.6%) while the anaerobic lactic system used about 6.2 ± 2.4% [2].

Different authors have reported the aerobic and anaerobic alactic systems as potential predictors of karate performance [2,10,22]. However, data on aerobic demands in specific conditions of *kumite* are not largely described in investigations concerning athletes [18]. Although investigations in laboratory situations with aerobic system analysis have been done, the results are not consistent and the few available ones measured aerobic demands with a specific approach to the combat situation, which is essential for karate *kumite* athletes [5,23,24]. However, no study has investigated a specific aerobic test for karate, except that of Nunan [18], who created the karate specific aerobic test (KSAT).

The KSAT consists of repeated sequential sets involving a straight punch and roundhouse kick combination on a heavy punch/kick bag suspended from a wall mounted bracket, interspersed with recovery periods. The goal here is to maintain exercise intensity whilst making the test progressively more demanding by means of reducing the recovery periods between exercises bouts. Nevertheless, to be valid and applicable a test must be reliable [12]. Knowledge of the specific demands of karate via KSAT assessment protocol can assist coaches and even athletes in the best way to optimise training programs [12,18]. Additionally, the use of specific skills is essential in order to reflect the demands of karate; without them, little credit can be given to collected data on aerobic demands. Knowledge about the true and criterion value of the analysis and the difference between the values provide the most important indication of test quality. Another important characteristic of consistent results for this type of test is the intra- and inter-rater reliability, which concerns the reproducibility of the observed value when the measurement is repeated [9]. Since, as this paper discussed, the main attributes of specific test quality include the reproducibility of the criteria established, the aim of the present study was to determine the reliability of the KSAT.

MATERIALS AND METHODS

Participants. The sample was composed of 16 black-belt adult male athletes, who had competed at national level competitions. The participants' characteristics were as follows (mean ± SD): age 23 ± 4 years; height 175.7 ± 6.7 cm; body mass 72.7 ± 7.2 kg; karate practice time 9 ± 4 years. The study was conducted in the middle of the competitive season. This study was submitted to and approved by the Committee of Ethics in Research. All participants took part voluntarily after being informed about the risks and benefits of the procedures involved and signed an informed consent form that was previously approved by the Ethics Committee.

Experimental procedures

Subjects participated in three sessions of data collection. One practice trial was performed one week before the initiation of baseline testing to gain familiarity with the test. The two other sessions of the KSAT were evaluated by having the subjects perform the KSAT on two separate occasions (one week apart) at the same time of day to establish test-retest reliability. All tests were administered by the same researcher. Before each testing session, participants performed a 15 min standardized warm-up that included prescribed jogging and stretching, as the KSAT protocol indicates [18].

The subjects performed the two test trials wearing karate-specific clothing and footwear and they were asked to use the same set in both. They were further instructed to follow their normal diet, to consume a light meal at least three hours prior to each trial, to avoid all kinds of vigorous activity before each test session, to sleep normally and to avoid any intense physical activity in the 24 hours prior to the test. The KSAT was performed indoor in the following conditions during tests: temperature 22°C and 63% humidity. Experimenters provided vigorous verbal encouragement during the test so as to push the participant to perform maximum effort.

Heart rate was measured every 5 seconds through the KSAT (Polar S610, Kempele, Finland). Capillary blood samples were drawn within ~3 minutes after the test. Blood lactate was determined using the Lactate Pro Analyzer (Arkray, Tokyo, Japan), which has been shown to be both accurate and reliable [20]. The Lactate Pro analyser is supplied with a check strip to confirm that the analyser is operating correctly and a calibration strip that provides a non-quantitative indication of instrument accuracy.

Global RPEs were recorded immediately after the KSAT using the Borg scale (RPE, 1-10) [4]. During the test, the heavy bag (25 kg, 30 × 100 cm) was held in place to avoid unwanted movement. The test ended when the subject achieved volitional exhaustion, at which point time to exhaustion (TE), exercise level and the cycle number performed during the test were recorded.

The KSAT

This protocol test was developed and prescribed by Nunan [18]. This test was composed of sequential sets of straight punch and roundhouse kick combinations on a heavy punch/kick bag suspended from a wall mounted bracket. The exercise sequence combination included a leading straight punch (Panel A, Figure 1) followed by a rear leg roundhouse kick (Panel B, Figure 1), a rear straight punch (Panel C, Figure 1) and a leading roundhouse kick (Panel D, Figure 1), repeated twice. Each set time of movement was performed in 7 s with a maximal time of 1182 s (19 min and 42 s).

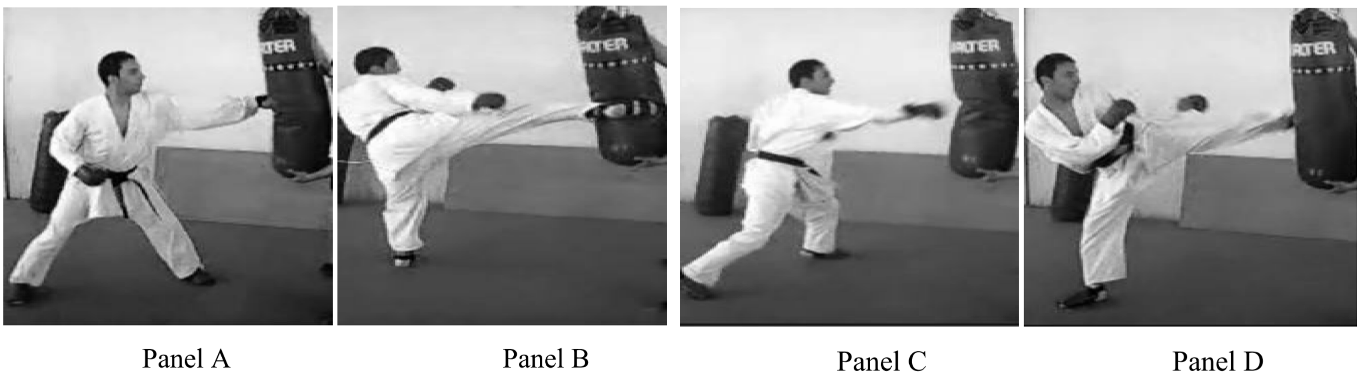

FIG. 1. TECHNIQUES USED DURING THE KARATE SPECIFIC AEROBIC TEST. PANEL A: KIZAMI-TSUKI (STRAIGHT PUNCH); PANEL B: MAWASHI-GERI (REAR LEG ROUNDHOUSE KICK); PANEL C: KYAKU-ZUKI (REAR STRAIGHT PUNCH); PANEL D: KIZA-MAWASHI-GERI (LEADING ROUNDHOUSE KICK).

The progression in intensity of the exercise during the test was based on a similar sequence of emitted audio beeps as the multistage fitness test [13]. The test was designed with two auditory signals, the first to let the participant know when to begin the bout of exercise and a second sound to indicate when they should rest (7 s later). The time to complete the exercise bout remained the same, 7 seconds, whilst the recovery time between bouts progressively decreased. Participants were encouraged to perform each strike and kick with the maximum force possible. The aim was to maintain maximal exercise intensity whilst progressively making the test more demanding by reducing the recovery in between exercise bouts. When the participant failed to complete the set of movements in the 7 s interval twice or when there was a clear decrease in the power of techniques according to the recommendations provided in the original article from Nunan [18], the time to exhaustion was recorded and represented the final test result.

Statistical analysis

Data are shown as mean ± SD. Normality was analysed using the Kolmogorov–Smirnov test. All variables presented a normal distribution. Heteroscedasticity was revealed by calculating the correlation coefficient between the absolute difference and the average of the test trials. The intraclass correlation coefficient (ICC) model 3.1 was used to examine the relative reliability of the KSAT. This model was used because this investigation had an intra-rater design with a single tester representing the only tester of interest [19]. The present study considered an ICC over 0.90 as high, between 0.80 and 0.90 as moderate and below 0.80 as low [28]. The SEM, the coefficient of variation (CV) and 95% limits of agreement (LOA) were calculated as indications of the absolute reliability [3]. The SEM was calculated using the square root of the error mean square (SEM= $\sqrt{MSE}$) [14,25]. Before reporting the relevant data in the units of measurement, heteroscedasticity was assessed using a zero order correlation between the absolute residuals and the predicted scores for each participant. The SEM also allows the calculation of the $MDC_{95\%}$. The $MDC_{95\%}$ demonstrated the 95% confidence interval (CI) of the difference in score between paired observations. In the present study the $MDC_{95\%}$ was calculated using the formula: $MDC_{95\%} = SEM \times x1.96$ [6,29]. To evaluate the usefulness of the KSAT, the smallest worthwhile change (SWC) was calculated as 0.2 × between-subject standard deviation of that test [1,8,21]. If the SEM is lower than the SWC, then the test is rated as "good". If the SEM is much greater than the SWC, then the test is rated as "marginal". If the typical error is about the same as the SWC, then the test may be useful, that is, "satisfactory" [11,26].

To investigate systematic bias, a paired Student's t-test was conducted to test the hypothesis of no difference between sample mean scores for test and retest. The effect size of the difference between test and retest (*dz*) was determined. The modified Hopkins scale was used for the interpretation of *dz*: trivial: < 0.2, small: 0.2 – 0.6, moderate: 0.6 – 1.2, and large: > 1.2 [9]. All data analyses were performed using SPSS for Windows (version 19.0, SPSS, Inc, Chicago, IL). Significance for all statistical tests was accepted at $p \leq 0.05$.

RESULTS

The heteroscedasticity coefficients for TE, HR_{peak}, La_{peak}, and RPE were r = -0.15 (p = 0.56); r = 0.32 (p = 0.29); r = -0.27 (p = 0.44) and r = -0.32 (p = 0.22) respectively, indicating no heteroscedasticity in raw data. The dependent t test, conducted to test the hypothesis of equality of means, showed no significant bias between TE (t = -1.77; p = 0.09; *dz* = 0.08), HR_{peak} (t = -1.18; p = 0.26; *dz* = 0.2 [trivial]), La_{peak} (t = 2.01; p = 0.07; *dz* = 0 [trivial]) and RPE (t = -0.41; p = 0.68; *dz* = 0.11 [trivial]) for the test and the retest (le 1).

Results of the test-retest absolute and relative reliability analyses are presented in Table 2. TE demonstrated good relative ($ICC_{(3,1)} = 0.982$) and absolute (CV = 4.57%; the mean difference

TABLE 1. PHYSIOLOGICAL AND PERCEPTUAL CHARACTERISTICS OF KSAT (MEAN ± SD) (N = 16)

Parameters	Test	Retest	P
HR peak (bpm)	197.2 ± 7.5	195.4 ± 7.5	P = 0.26
La_{peak} (mmol · l^{-1})	9 ± 2.6	8.3 ± 3	P = 0.07
RPE	7.6 ± 1	7.7 ± 1.2	P = 0.68

Note: HR_{peak}: Peak heart rate, La_{peak}: Peak blood lactate concentration, RPE: Rating of perceived exertion

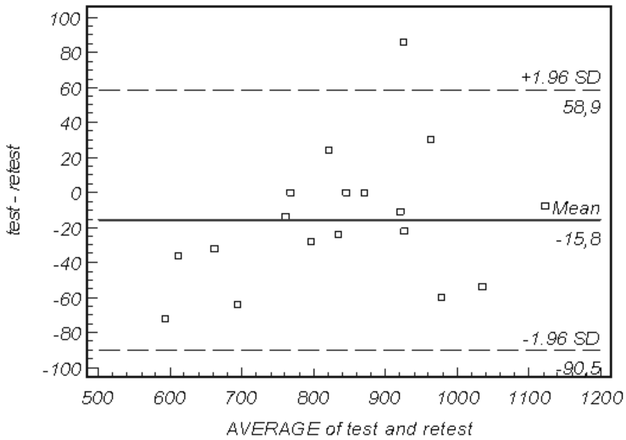

FIG. 2. BLAND AND ALTMAN PLOT FOR TIME TO EXHAUSTION TEST-RETEST SESSIONS

TABLE 2. TEST-REST RESULTS AND RELIABILITY INDICES OF THE KSAT (N = 16)

Parameters	TE Test	TE Retest	ICC	SEM(sec)	SWC(sec)	MDC95%(ses)
TE(s)	832.6±151.4	848.4±139.1	0.982	29.37	28.07	81.42

Note: s: seconds; TE: Time to exhaustion; ICC: intraclass correlation coefficient; SEM: standard error of measurement; SWC: smallest worthwhile change; $MDC_{95\%}$: minimal detectable change at 95% confidence interval.

(bias) ± the 95% limits of agreement was -15.8 ± 74.7 s (Figure 2)) reliabilities. The SEM of the TE was within 29.37; it was about the same as the SWC (28.07). The $MDC_{95\%}$ for TE was 81.42 s.

DISCUSSION

The reliability of fitness testing is a critical issue. In physical fitness testing, the use of a sport-specific test that incorporates the major key techniques of the activity targeted is crucial. In this context, despite not creating any new test, the present research was the first to attempt to further study the relative and absolute reliability and the minimal detectable change of the KSAT, which was previously presented by Nunan [18]. As mentioned above, test-retest reliability is critical in establishing the reproducibility of a test. Moreover, it is a prerequisite to establishing validity in future studies [12,16].

According to Impellizzeri and Marcora [11], reliability results are most often used to judge whether a test protocol should be widely used. The main finding of this study was that the KSAT is a reliable tool for the evaluation of specific endurance in karate athletes with a small $MDC_{95\%}$. To the best of the authors' knowledge, this is the first serious attempt to deal with the study of relative and absolute reliabilities of the KSAT and the $MDC_{95\%}$. Indeed, the first modest attempt to study the reliability of the KSAT was carried out by Nunan [18] on a reduced sample size (five subjects) and by only comparing the equality of means of KSAT test and retest using the Wilcoxon signed-rank statistical test. This statistical tool is much criticised for this kind of analysis. In this study, the approach of test-retest reliability of the KSAT was provided by identifying two complementary but distinct methods, i.e. relative and absolute reliabilities.

Relative reliability is a method that commonly relies on determining the intraclass correlation coefficient (ICC) [29]. In the present study the $ICC_{(3,1)}$ was 0.98 for TE, approaching unity. It has been reported that an ICC above 0.90 defines a high level of relative reliability [28]. In fact, this high value of TE ICC indicates that KSAT is a reliable test. According to Bland and Altman [3], ICC cannot be the only statistical measure of reliability as it is affected by the heterogeneity of the sample. Consequently, as measures of absolute reliability, SEM and CV are recommended to be used, in conjunction with the ICC [15].

If data are homoscedastic, as in the present study ($r = -0.15$; $p = 0.56$ for TE), an appropriate measure of absolute reliability is the SEM, which can be presented in absolute or relative expression. The CV is used with heteroscedastic data [3]. Thus, absolute reliability in the present study was evaluated by the SEM only. According to Nevill and Atkinson [17], any two tests would differ, because of measurement error (SEM in the present study) by no more than 5%. In fact, the absolute reliability of the TE of the KSAT was very good, with SEM 3.4%, far below the 5% limit. It is important to note that having good reliability does not mean that a variable is useful, as a number of physiological measures have high reliability but some are not sensitive measurements tools [8]. In the current study, the SWC (28.07 s) was similar to SEM (29.37 s) and the usefulness of the test was rated as "satisfactory".

The SEM allows the calculation of the minimal detectable change ($MDC_{95\%}$). The $MDC_{95\%}$ is considered as an estimation of the smallest change in score that can be detected objectively for a test, i.e. the amount by which the subject's performance needs to be changed to be sure the change is bigger than the measurement error [6]. It is important to report that $MDC_{95\%}$ differs from SWC, which is the smallest change in an outcome measure that would be considered important and so introduces an element of subjectivity [6]. $MDC_{95\%}$ provides information that less measurement error indicated more sensitivity to real change. The $MDC_{95\%}$ in this investigation indicates that 95% of the outcomes of the KSAT will demonstrate a random variation as a result of measurement error of less than 9.4% (81.42 s) for TE. No other research was found with the determination of the $MDC_{95\%}$ of the KSAT, so it is difficult to compare this result to those of the specific literature.

The $MDC_{95\%}$ is a generate statistic that provides practical relevant information in meaningful terms; nevertheless, a documented limitation of the MDC is that it is affected by the heterogeneity of the sample [27]. This limitation illustrates the necessity of having an MDC value that is specific to particular characteristics of athletes, i.e. sex, age and level of expertise. Therefore, future studies can attempt to quantify a more robust $MDC_{95\%}$ using larger samples of karate athletes, and $MDC_{95\%}$ differentiated by age, sex and expertise. Another common criterion to verify absolute reliability of a test was the Bland and Altman method [3]. In our study, there was a high level of concordance between scores of the TE test-retest verified by Bland and Altman plots (Figure 2). In these analyses, both bias and random error were found to be low, resulting in good reliability. To put these results in a practical context, if a subject from the study population performed a TE of 910 s on the first application of the KSAT, it suggests that he could perform on the second trial a score as high as 984.7 s (910+74.7), or as low as 835.3 s (910-74.7). We could consider these LOA acceptable. Due to the absence of any earlier study dealing with the reliability of the KSAT, the present research is unable to compare results with previous findings.

CONCLUSIONS

The main finding of the present study is that the karate specific aerobic test (KSAT) provides good levels of absolute and relative reliabilities and a small $MDC_{95\%}$ in a sample of experienced karate athletes. It is important for coaches administering this test to know that a change of greater than 9.4% (81.42 s) is necessary to be 95% certain that the change in performance indicates improvement and exceeds the measurement error. Further studies should also focus on validity, especially direct validity, which is a major strength of any test when established.

Acknowledgments

The present study was financially supported by the "Ministère de l'Enseignement Supérieur et de la recherche Scientifique", Tunisia.

The authors acknowledge all subjects for their enthusiasm and commitment in the completion of this research study. The authors would like to thank Ms Bianca Miarka for her help with the English revision of the manuscript.

REFERENCES

1. Atkinson G., Nevill A.M. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26:217-238.
2. Beneke R., Beyer T., Jachner C., Erasmu J., Hutler M. Energetics of karate kumite. Eur. J. Appl. Physiol. 2004;92:518-523.
3. Bland J.M., Altman D.A. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-310.
4. Borg G. Borg's Perceived Exertion and Pain Scales. Human Kinetics, Champaign, IL, 1998.
5. Doria C., Veicsteinas A., Limonta E., Maggioni M.A., Aschieri P., Eusebi F. Energitics of karate (kata and Kumite techniques) in top-level athletes. Eur. J. Appl. Physiol. 2009;107:603-610.
6. Haley S.M., Fragala-Pinkham M.A. Interpreting change scores of tests and measures used in physical therapy. Phys. Ther. 2006;86:735-743.
7. Heller J., Peric T., Olouha R., Kohtikova E., Melicha J., Novakova N. Physiological profiles of male and female taekwondo black belts. J. Appl. Sport Sci. 1998;16:243-249.
8. Hopkins W.G., Elske J., Schabort E.J., Hawley J.A. Reliability of power in physical performance tests. Sports Med. 2001;31:211-234.
9. Hopkins W.G. Measure of reliability in sports medicine and science. Sports Med. 2000;30:1-15.
10. Imamura H., Yoshimura Y., Uchida K., Nishimura S., Nakazawa A.T. Maximal oxygen uptake, body composition and strength of highly competitive and novice karate practitioners. Appl. Hum. Sci. 1998;17:215-218.
11. Impellizzeri F.M., Marcora S.M. Test validation in sport physiology: Lessons learned from clinimetrics. Int. J. Sports Physiol. Perform. 2009;4:269-277.
12. Kristina S.B., Maurice D.D., Morey J.K., Ming-Shun S.C. Test-retest reliability and minimal detectable change of the Hexagon agility test. J. Strength Cond. Res. 2009;23:2167-2171.
13. Léger L.A., Mercier D., Gadouy C., Lambert J. The multistage 20 m shuttle run test for aerobic fitness. J. Sport Sci. 1988;6:93-101.
14. Lexell J.E, Downham D.Y. How to assess the reliability of measurements in rehabilitation. Am. J. Phys. Med. Rehabil. 2005;84:719-723.
15. Looney M.A. When is the intraclass correlation coefficient misleading? Meas. Phys. Educ. Exerc. Sci. 2000;4:73-78.
16. Munro A.G., Herrington L.C. Between-session reliability of four hops tests and the agility T-Test. J. Strength Cond. Res. 2011;25:1470-1477.
17. Nevill A., Atkinson G. Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science. Br. J. Sports Med. 1997;31:314-318.
18. Nunan D. Development of a sport specific aerobic capacity test for karate – A pilot study. J. Sports Sci. Med. 2006;5:47-53.
19. Portney L.G., Watkins M.P. Foundations of Clinical Research. Applications to Practice. 3rd Ed. Pearson Prentice Hall, Upper Saddle River, NJ 2008.
20. Pyne D.B., Boston T., Martin D.T., Logan A. Evaluation of the Lactate pro blood lactate analyser. Eur. J. Appl. Physiol. 2000;82:112-116.
21. Pyne D.B. Interpreting the results of fitness testing. International Science Football Symposium. Australia 2003.
22. Ravier G., Rouillon J.D. Intensité relative et durée des différentes phases d'une compétiton de karate. Sci. Motricité 2000;41.
23. Ravier G., Dugue B., Grappe F., Rouillon J.D. Maximal accumulated oxygen deficit and blood responses of ammonia, lactate, and pH after anaerobic test: a comparison between international and national elite karate athletes. Inter. J. Sports Med. 2006;27:810-817.
24. Ravier G., Grappe F., Rouillon J.D. Application of force-velocity cycle ergometer test and vertical jump tests in the functional assessment of karate competitor. J. Sports Med. Phys. Fitness 2004;44:349-355.
25. Schreuders T.A., Roebroeck M.E., Goumans J., Van Nieuwenhuijzen J.F., Stijnen T.H., Stam H.J. Measurement error in grip and pinch force measurements in patients with hand injuries. Phys. Ther. 2003;83:806-815.
26. Spencer M., Fitzsimons M., Dawson B., Bishop D., Goodman C. Reliability of a repeated-sprint test for field-hockey. J. Sci. Med. Sport 2006;9:181-184.
27. Starford T.J., Binkley J., Solomon P., Finch E., Gill C., Moreland J. Defining the minimum level of detectable change for the Roland-Morris questionnaire. Phys. Ther. 1996;76:359-365; discussion 366-368.
28. Vincent W.J. Statistics in Kinesiology. Human Kinetics, Champaign, IL 1995.
29. Weir J.P. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J. Strength Cond. Res. 2005;19:231-240.
30. World Karate Federation (WKF). New Kata and Kumite Rules [Version 7.1]. 2012.