

# Reliability and criterion-related validity of the 20-yard shuttle test in competitive junior tennis players

Anna Eriksson<sup>1</sup>  
Fredrik R Johansson<sup>2</sup>  
Maria Bäck<sup>3-5</sup>

<sup>1</sup>Rehab City Östermalm, Primary Health Care, <sup>2</sup>Department of Environmental Medicine, Musculoskeletal and Sports Injury Epidemiology Center, Karolinska Institutet, Stockholm, <sup>3</sup>Department of Occupational Therapy and Physiotherapy, Sahlgrenska University Hospital, <sup>4</sup>Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, <sup>5</sup>Department of Medical and Health Sciences, Division of Physiotherapy, Linköping University, Linköping, Sweden

**Purpose:** This study adds to the previous work in the field of sport-specific fitness testing by evaluating a tennis-specific agility test called “the 20-yard shuttle test”. The aim of the study was to evaluate the test–retest reliability, the inter-rater reliability, and the criterion-related validity of the 20-yard shuttle test on competitive junior tennis players.

**Participants and methods:** Totally, 34 Swedish tennis players (13 girls), mean age 14±1.6 years, participated in the study. To examine test–retest reliability, the subjects performed the 20-yard shuttle test three times on the same day and then the same procedure was repeated after 3 days. To test the inter-rater reliability, the time was measured with a stopwatch simultaneously by two different raters. The time recorded manually was compared to the gold standard of digital timing to evaluate the criterion-related validity.

**Results:** Excellent test–retest reliability was found both within the same day (intraclass correlation coefficient [ICC] 0.95) and between days (ICC 0.91). Furthermore, the results showed excellent inter-rater reliability (ICC 0.99) and criterion-related validity on both test occasions (ICC 0.99).

**Conclusion:** We have provided introductory support for the 20-yard shuttle test as a reliable and valid test for use in competitive junior tennis players. The ease of administration makes this test a practical alternative to evaluate physical fitness in order to optimally train the athletes.

**Keywords:** agility, physical fitness, physiotherapy, performance

## Introduction

Tennis is a complex sport since it involves technical and tactical skills. Further, it is also a physically demanding sport.<sup>1</sup> This complexity requires tennis athletes to have fast reaction times and the ability to perform explosive agility movements. The movement pattern is characterized by quick starts and stops, accelerations, decelerations, and multidirectional movements.<sup>2</sup> A study in junior tennis players that compared short straight sprints with agility, including multidirectional running, showed that these are specific qualities that are most of the time unrelated.<sup>3</sup> Due to these findings, it is important to train and test tennis players in the specific movement patterns and running distances that are encountered during match play, typically high-intensity work for approximately 4–10 seconds.<sup>4</sup>

Repetitive multidirectional movement patterns can lead to lower extremity injury.<sup>5</sup> Also, acute injuries are common problems in pivoting sports that include cut movements, sudden accelerations, stops and turns, since these can place substantial demands on ankles, knees, and hips.<sup>6</sup> Furthermore, it is hypothesized that a nonfunctional movement pattern can predispose for injuries.<sup>7</sup> To identify areas of reduced fitness, it is important to conduct a regular physical fitness testing.

Correspondence: Maria Bäck  
Department of Physiotherapy,  
Sahlgrenska University Hospital, Vita  
stråket 13, 413 45 Gothenburg, Sweden  
Tel +46 31 342 11 95  
Fax +46 31 825 604  
Email maria.m.back@vgregion.se

The usefulness of a test depends on its reliability or the extent to which a test is consistent and free from error.<sup>8</sup> The test also needs to be specific to the demands of the sport that the subject is practicing.<sup>9</sup> To the best of our knowledge, there are a restricted number of reliable functional tests for determining physical fitness, especially concerning speed and agility, that exist in the literature. Moreover, these parameters are most often tested in a laboratory setting.<sup>10</sup> Compared to functional tests, laboratory measurements are less accessible and also often expensive. Furthermore, these tests are aimed at measuring only one specific parameter, eg, muscle strength. How do the results of these tests represent for physical performance is still not clear.<sup>10</sup> Therefore, there is a need for more reliable and valid sports-related functional tests, aiming to test different aspects of physical fitness and athletic performance. Such tests are often inexpensive and easy to perform.<sup>10</sup>

A physical fitness test that is frequently used by the US Tennis Association to evaluate agility in competitive tennis players is “the spider run test”.<sup>11</sup> This test is easy to administer and the movement patterns simulate a lot to the actual movements observed during tennis play. However, the average time to perform the test is approximately 15–17 seconds for junior males and females, which is longer than an average point in tennis.<sup>4</sup> The Swedish Tennis Federation has a test battery including flexibility and fitness testing for competitive tennis players aged 12–20 years. According to the results of these tests, training can be individualized, supplemented, or adjusted to optimize performance. The test used to evaluate agility is called “the 20-yard shuttle test”.<sup>12</sup> It originates from American football and involves acceleration, deceleration, and multidirectional short distance speed. The time to complete the test is approximately 5–6 seconds. Accordingly, the test resembles the movements in tennis and also reproduces the time frame for the majority of points in match play.

Although the 20-yard shuttle test is frequently used, studies evaluating the reliability of the test are scarce. To our knowledge, only two studies have examined the test–retest reliability using the 20-yard shuttle test, and the tests are performed on athletes mostly involved in different team sports, and test–retest reliability is performed within the same day.<sup>13,14</sup> The reliability has not been tested in tennis players. Furthermore, the 20-yard shuttle test is in practice evaluated with a manual stopwatch.<sup>12</sup> Studies investigating the criterion-related validity, ie, the manual recording of time compared to the gold standard of digital timing are lacking for this test. Considering the significance of physical fitness testing in

tennis players to evaluate and optimize performance and to reduce the risk of injuries, a reliable and valid tennis-specific test is needed. Therefore, the aim of our study was to evaluate the reliability and criterion-related validity of the 20-yard shuttle test in Swedish competitive junior tennis players.

## Methods

### Participants

A sample of 39 competitive junior tennis players (mean height 164±10.6 cm, mean mass 52.5±10.5 kg, mean age 14±1.6 years) volunteered to participate in the study. All members in two different tennis clubs in Stockholm, Sweden, that met the inclusion criteria were asked to participate in the study by their tennis coach. Inclusion criteria were as follows: boys and girls, aged from 12 to 20 years, regularly competing (a minimum of five tournaments per year), and enrolled in ≥2 tennis training sessions per week. Exclusion criteria were injuries in the lower extremity by the time for the test that preclude maximal performance in terms of speed. Change of shoes from one session to the other was also an exclusion criterion.

From a total of 44 persons who fulfilled the inclusion criteria, 39 took part in the study. The main reason for exclusion was ongoing tournament play. Thirty-four persons participated in both sessions. Descriptive characteristics of the study population are presented in Table 1. Most subjects stated that their physical training included training on speed, strength, conditioning, and power. Fourteen subjects were practicing other sports in addition to tennis. The most common sports were soccer, hockey, handball, and golf.

The research protocol was approved by the Regional Ethical Review Board at the University of Gothenburg. A written informed consent was provided prior to the participation in the study. For subjects below 15 years, an informed consent was also provided to their guardians.

### Procedures

A pilot study including five subjects was completed prior to the study to ensure the procedure of the test and for the raters to become familiarized with the stopwatch. There was no occasion for retest in the pilot study.

The 20-yard shuttle test was performed indoors on a tennis court. The surface was hard court, since it has been frequently used in various tournaments all over the world. The first test occasion started off by weighing and measuring all the subjects. They were instructed to wear shorts and T-shirt and to take off their shoes. They also answered a questionnaire about their training frequency, number of

**Table 1** Descriptive characteristics of the study population (N=34)

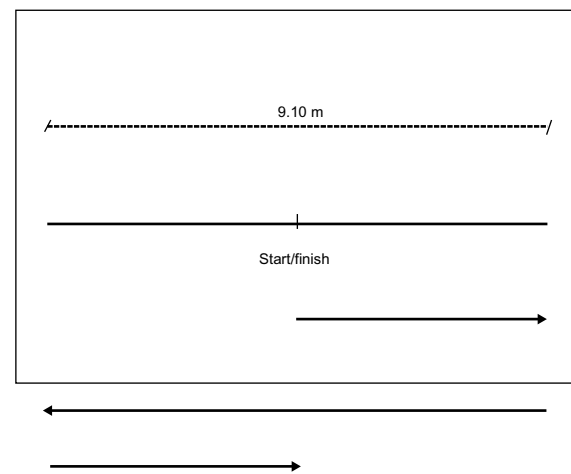
Age (years)	14±1.6
Minimum–maximum	12–17
Sex, n girls/boys (%)	13/21 (38.2/61.8)
Height (cm)	164.0±10.6
Mass (kg)	52.5±10.5
Years of playing tennis	7.2±2.4
Hours of tennis training/week	4.8±2.3
Hours of physical training/week	2.0±1.4
Number of tennis tournaments/year	12.3±7.2
Hours of alternative training/week	1.3±2.2
Subjects participating in other sports, n (%)	14 (41.2)
Current injury in upper body, n (%)	3 (8.8)
History of leg/foot injury in last 6 months (%)	6 (17.2)
History of infection in last month (%)	14 (41.2)

**Notes:** Continuous data are presented with mean ± standard deviation. Categorical data are presented as absolute numbers and proportions.

**Abbreviation:** n, number.

tournaments they participated in 1 year, and injury history. After this, a standardized warm-up of 10 minutes was performed including jogging, lateral displacements, sprints, and dynamic stretching. The test was performed at the same place and approximately at the same time of the day for both test sessions to avoid the effects of diurnal variation. Participants were asked to refrain from strenuous exercise 24 hours prior to the test and not to consume food, caffeine, or nicotine for 3 hours before the testing session. Participants received thorough standardized instructions on how to practice the test and were given instructions to perform the test as fast as they could. No verbal encouragement was used during the performance.

The test was set up in the following manner: Three marker cones were placed along a line 4.55 m apart. The players were instructed to straddle a marked tape (48 cm) behind the middle line, which served as the start/finish line (where the photoelectric barriers were placed), and put one hand down in a three-point stance. On hearing the command “ready, steady, go”, the subject started, and the raters started the stopwatches as soon as they crossed the start line. The subject turned and ran as fast as possible 4.55 m to the right side and touched one foot behind the line. The subject then ran 9.1 m to the left and touched one foot behind the other line and finally finished by running back through the finish line. When the subject crossed the line, the stopwatches, both manually and digitally, were stopped. The test is illustrated in Figure 1. The duration of each trial was recorded to the nearest 100th of a second. The photoelectric cell timer was automatically activated as the subject crossed the first cell and stopped when the subject crossed the last cell.

**Figure 1** A schematic diagram of the 20-yard shuttle test.

**Note:** The arrows indicates the distances that the subject is running at the given start command.

## Reliability and validity analyses

### Test–retest reliability

The test was performed three times, with 5 minutes of rest in between trials, according to the test procedure of the Swedish Tennis Federation. The same procedure was then repeated after 3 days. The same person (a physiotherapist) executed all the tests for both the test sessions. The test leader did not have access to the results obtained from the previous test session.

### Inter-rater reliability

During the first test session, a tennis coach was also present to manually keep track of the time along with the physiotherapist. The time required to complete the test was measured simultaneously with a stopwatch by rater 1 and rater 2. The raters started the stopwatches when the subject crossed the start line (and the photoelectric cells) and stopped as the subject crossed the finish line. The two raters were standing on opposite sides of the start line (where the photocells were placed) facing each other. The raters were blinded to the results of one another.

### Criterion-related validity

The time recorded manually was compared to the gold standard of digital timing in both test sessions. Time was recorded by the photoelectric cells equipment “IVAR” (Ivar Krause, Tallin, Estland).

## Statistical analyses

Data were analyzed by the Statistical Package for the Social Sciences (SPSS 20.0, Chicago, IL, USA). All the study variables

were normally distributed. Descriptive measures for continuous data were calculated with mean  $\pm$  one standard deviation (SD). Categorical variables were described as absolute numbers and proportions. The average intraclass correlation coefficient (ICC 3,3) with a 95% confidence interval (CI) was used to determine analyses within sessions, including test–retest reliability, inter-rater reliability, and criterion-related validity (concurrent validity). To calculate ICC (3,1) for between-session analyses, the best value from each session was used. The ICC varies from 0 to 1, where 1 is considered perfectly reliable. For this study, an ICC greater than 0.75 was considered excellent, from 0.4 to 0.75 was considered fair to good, and less than 0.4 was considered poor.<sup>15</sup> A complementary standard error of measurement (SEM) and SEM% were presented in relation to the ICC. Bland–Altman plots were performed to visualize the difference against the mean of best manual and digital test–retest values between sessions.<sup>16</sup> An analysis of variance (ANOVA) with repeated measures was performed to test the presence of systematic trends in measurements. The following design was used: SESSION (1,2)  $\times$  TYPE (digital, manual)  $\times$  TRIAL (1,2,3). Moreover, inter-rater effects were investigated with another ANOVA with repeated measures using the following design: RATER (1,2)  $\times$  TRIAL (1,2,3). All tests were two-sided and considered significant if  $P < 0.05$ .

A sample size calculation for the differences in seconds between test and retest was performed before the start of the study. The power was 0.80 and  $\alpha$ -value was 0.05. A medium effect size of 0.5 with a mean difference of 0.2 seconds

between test and retest resulted in an SD of difference 0.4 seconds, which generated a sample size of 34 subjects.

## Results

Descriptive data from each test session are listed in Table 2, including mean times in seconds, minimum and maximum values, and SD from each trial. In addition, the best manual and digital times from each test session are presented.

### Test–retest reliability

The results indicated excellent same-day test–retest analyses for manual tests in session 1 (ICC 0.95, 95% CI 0.91–0.97) and session 2 (ICC 0.96, 95% CI 0.92–0.98). Furthermore, the within-session test–retest analyses for digital times showed excellent results in session 1 (ICC 0.95, 95% CI 0.92–0.98) and session 2 (ICC 0.96, 95% CI 0.94–0.98). Moreover, the results showed excellent reliability for the between-sessions test–retest reliability for both the best manual scores (ICC 0.95, 95% CI 0.90–0.97) and the best digital scores (ICC 0.91, 95% CI 0.83–0.96). For more detailed results, see Table 3.

In addition, Bland–Altman plots showed that the mean difference between the best manual (Figure 2) and the best digital test–retest scores between sessions (Figure 3) was close to zero.

Results from the ANOVA with repeated measures showed statistically significant main effects for TYPE ( $P < 0.001$ ) and TRIAL ( $P = 0.001$ ), and there was a significant interaction

**Table 2** Descriptive data from each test session (N=34)

20-yard shuttle test	Mean time (seconds)	Minimum time (seconds)	Maximum time (seconds)	SD (seconds)
Sess 1, TL 1, T1	5.23	4.60	5.88	0.34
Sess 1, TL 1, T2	5.17	4.47	6.00	0.35
Sess 1, TL 1, T3	5.13	4.53	5.87	0.33
Sess 1, best manual, TL1	5.07	4.47	5.84	0.34
Sess 1, TL 2, T1	5.28	4.60	5.85	0.36
Sess 1, TL 2, T2	5.22	4.56	5.97	0.33
Sess 1, TL 2, T3	5.19	4.68	5.85	0.34
Sess 1, dig, T1	5.30	4.67	5.89	0.34
Sess 1, dig, T2	5.23	4.56	6.00	0.30
Sess 1, dig, T3	5.21	4.59	5.87	0.32
Sess 1, dig, best	5.15	4.56	5.87	0.33
Sess 2, TL 1, T1	5.21	4.60	5.84	0.34
Sess 2, TL 1, T2	5.17	4.53	5.84	0.33
Sess 2, TL 1, T3	5.17	4.53	5.90	0.34
Sess 2, best manual, TL1	5.10	4.53	5.75	0.32
Sess 2, dig, T1	5.24	4.67	5.77	0.32
Sess 2, dig, T2	5.19	4.58	5.91	0.33
Sess 2, dig, T3	5.20	4.56	5.91	0.33
Sess 2, dig, best	5.13	4.56	5.71	0.32

**Abbreviations:** N, number; SD, standard deviation; TL, test leader; T, trial; Sess, session; dig, digital time.

**Table 3** Test–retest manual and digital measurements: within-session and between-session reliabilities (N=34)

Type of measurement	Number of measurements	ICC	95% confidence interval		SEM	SEM %
			ICC lower	ICC upper		
Session 1: manual	3	0.95	0.91	0.97	0.22	1.43
Session 2: manual	3	0.96	0.92	0.98	0.20	1.30
Best manual over sessions	2	0.96	0.91	0.98	0.14	1.33
Session 1: digital	3	0.95	0.92	0.98	0.21	1.31
Session 2: digital	3	0.96	0.94	0.98	0.18	1.16
Best digital over sessions	2	0.96	0.91	0.98	0.14	1.32

**Notes:** Best manual over session is the best manual time from test leader 1 for sessions 1 and 2. Best digital over sessions is the best digital time recorded from sessions 1 and 2.  
**Abbreviations:** ICC, intraclass correlation coefficient; SEM, standard error of measurement.

effect between SESSION and TYPE ( $P<0.001$ ). The mean score for digital time was higher compared to manual time ( $P<0.001$ ). Furthermore, post hoc comparisons showed a significant lower mean time for trial 2 vs 1 ( $P=0.007$ ) and trial 3 vs 1 ( $P=0.002$ ) for session 1 (Figure 4). The interaction effect between SESSION and TYPE showed a significantly ( $P<0.001$ ) larger difference between digital and manual time for test session 1 compared to test session 2.

### Inter-rater reliability

The results showed excellent inter-rater reliability for best values between rater 1 and rater 2 (ICC 0.99, 95% CI 0.98–1.00, SEM 0.06).

Results from the ANOVA with repeated measures showed statistically significant main effects for RATER ( $P<0.001$ )

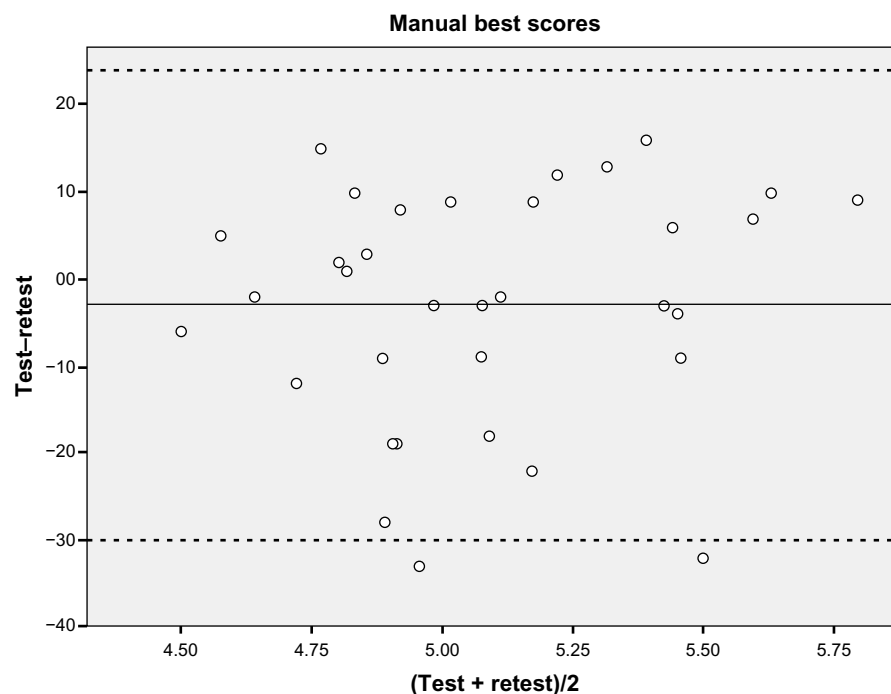
and TRIAL ( $P=0.013$ ). The mean score for rater 1 was lower compared to rater 2.

### Criterion-related validity

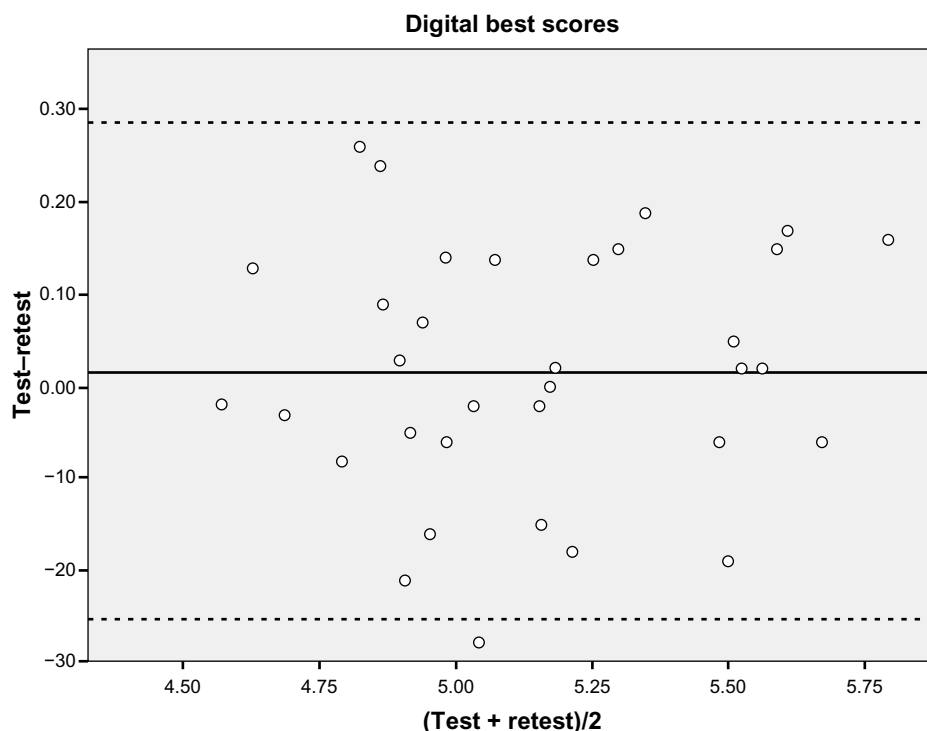
The results demonstrated an excellent criterion-related validity between manual and digital (gold standard) measurements using best individual values among trials from session 1 (ICC 0.99, 95% CI 0.98–1.00) and session 2 (ICC 0.99, 95% CI 0.99–1.00).

### Discussion

Our results support the reliability and the criterion-related validity of the 20-yard shuttle test conducted among competitive junior tennis players. To be relevant to a sport, a fitness test must mimic the demands of that particular sport. Hence,



**Figure 2** Bland–Altman plot showing the difference against the mean of the best manual test–retest values between sessions (n=34), with mean and limits of agreement, including two standard deviations.



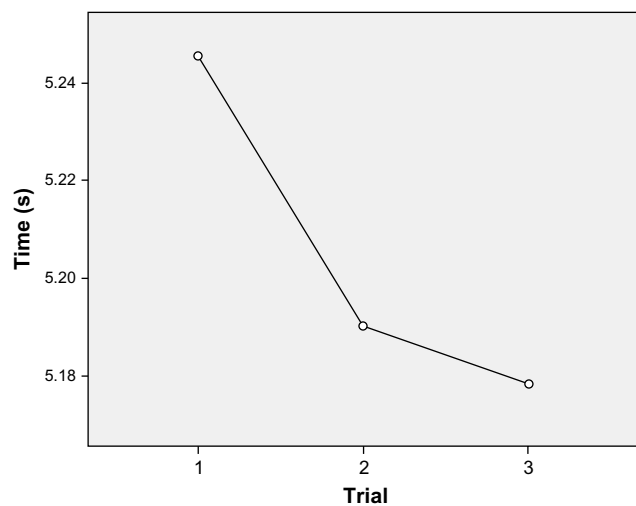
**Figure 3** Bland–Altman plot showing the difference against the mean of best digital test–retest values between sessions ( $n=34$ ), with mean and limits of agreement, including two standard deviations.

the fitness components that contribute to improvements in performance of that sport should be tested and evaluated.<sup>9</sup> These results add to the body of knowledge regarding the usefulness of the 20-yard shuttle test as a test tool in clinical and research practice for junior tennis players.

Test–retest reliability is important in establishing the reproducibility of a test.<sup>8</sup> Moreover, the reliability of a test is essential when being utilized to detect the improvements in physical abilities.<sup>17</sup> Our results showed that the 20-yard

shuttle test is highly reliable both when conducted within the same day and when repeated after 3 days at two different occasions. The test–retest reliability in this study was slightly better than that found in the other two studies for the 20-yard shuttle test on high school-aged<sup>14</sup> and college-aged boys and girls<sup>13</sup> involved in different sports. Our highly reliable results could be possibly due to the fact that the subjects in our study were exclusively tennis players who are used to this type of movement patterns in their sport. The subjects in the other two studies were all physically active but in different sports, some of them were involved in gymnastics and some in dance where this type of movement does not occur.<sup>13,14</sup> This theory is strengthened by the results of yet another study that evaluated the reliability of other agility tests and also obtained somewhat higher reliability parameters than those obtained by Stewart et al<sup>14</sup> and Sekulic et al<sup>13</sup> In that study, all the subjects were soccer players who are also used to agility movements in their sport.<sup>18</sup>

In the present study, a significant difference in time was seen between tests 1 and 2 and between tests 1 and 3 on test day 1. The subjects got progressively better. This was not seen on test day 2. Since most of the subjects had never performed the test prior to this study, it is likely to attribute this difference as learning effects. In a study by Sporis et al,<sup>19</sup> reliability was examined for six different soccer-specific agility tests,



**Figure 4** Means of measurement of time in trials 1, 2, and 3 ( $n=34$ ).



and the results of the first trial in all the agility tests were the weakest. They recommend at least one maximal practice trial before the actual test. The same results were concluded by another study when interpreting the descriptive statistics data of explosive power tests obtained from students.<sup>20</sup> Based on these results, at least one maximal test trial should precede the testing to reduce certain motor learning effects. In our study, there was no test trial prior to the actual test, but three tests were performed by each individual and only their best time was analyzed.

Our results showed excellent inter-rater reliability for best values between rater 1 and rater 2. The raters in our study did participate in a pilot study to get familiarized with the testing protocol and the stopwatch. A study by Vicente-Rodriguez et al<sup>21</sup> evaluated the inter-rater reliability of manual timing between trained and untrained raters, for the 4×10 m shuttle test and 30 m running speed tests, and the results showed a significant difference between raters with the trained rater measuring better times. When compared to digital timing (photoelectric cells), greater reliability (smaller systematic error) was observed between the trained rater and the digital timing. These results suggest that raters should be trained and be familiar with how to handle the stopwatch so as to minimize systematic error and to ensure accurate measurements.

Accurate timing in sprinting activities is of interest to athletes, coaches, and scientists. Although the ideal option always would be to use photoelectric cells to record the timing for different field tests, the most commonly used measurement tool is a manual stopwatch because it is easier to administer and a cheaper alternative.<sup>21</sup> There is limited research on the validity of handheld stopwatches compared with digital timing in speed and agility testing. The results in our study showed good criterion-related validity, which indicates that physiotherapists and tennis coaches can acceptably measure the 20-yard shuttle test on tennis players using a manual stopwatch. This result is in accordance with the study by Vicente-Rodriguez et al<sup>21</sup> which also found considerably small differences between manual timing by a trained rater compared with electronic timing when assessing the speed and agility of adolescents. On the other hand, another study by Mayhew et al<sup>22</sup> showed larger variations when studying the difference between manual and electronic timing of the 40-yard dash in college football players. The results showed that manual timing was significantly faster than electronic timing, although the raters were trained. The method used for manual timing was different in this study though. The electronic timing was started when the subject lifted their hand from a switched

mat, which is likely to cause a certain reaction time for the raters.<sup>22</sup>

The 20-yard shuttle test can be used to gain information about a tennis player in order to optimize performance and to reduce injuries. Tennis players need to have enough strength to be able to decelerate the movements of the body with control in order to quickly change direction.<sup>23</sup> The faster the player is moving, the bigger load the player will be exposed to. In order to accomplish this, the player must have vast eccentric strength.<sup>23</sup> Eccentric strength is also crucial for athletes from an injury prevention standpoint, since a lot of injuries occur during deceleration.<sup>24</sup> Furthermore, this test could be used by physiotherapists to detect weaknesses in different physiological parameters, such as muscle strength and balance, in addition to monitoring the development of performance. Dynamic balance, or the ability to keep the center of gravity over the base of support while the body is moving, is an important skill.<sup>23</sup> The 20-yard shuttle test is a test that is performed at full speed and is therefore a good complement to other tests that are performed in a controlled setting.

The results need to be considered in relation to the study's limitations. The author of this study was involved in the timing procedure which may be a potential source of bias. The results can also be discussed in terms of generalization. The subjects varied in ages, years of tennis played, and playing capacity. Our experience from the testing is that there was a larger variation in time between test and retest for the players who were performing weaker and scored poor test results (higher times). Also, we believe that motivation is a crucial factor for maximal performance. It is likely to believe that elite players are more motivated than players of lower level. Therefore, it would be interesting for future research to investigate whether the test-retest reliability varies between groups of elite players vs groups of amateur players.

## Conclusion

In conclusion, we have provided introductory support for the 20-yard shuttle test as a reliable and valid test for use in competitive junior tennis players. There is a need for further research to evaluate the usefulness and impact of this test among tennis players, in terms of optimizing performance and reducing injuries.

## Acknowledgments

The authors would like to thank Bengt Jansson for statistical advice. This work was supported by the Memorial Foundation at the Swedish Association of Physiotherapists.

## Disclosure

The author reports no conflicts of interest in this work.

## References

1. Reid M, Schneiker K. Strength and conditioning in tennis: current research and practice. *J Sci Med Sport*. 2008;11(3):248–256.
2. Llana-Belloch S, Brizuela G, Perez-Soriano P, Garcia-Belenguer AC, Crespo M. Supination control increases performance in sideward cutting movements in tennis. *Sports Biomech*. 2013;12(1):38–47.
3. Leone M, Comtois A, Tremblay F, Leger L. Specificity of running speed and agility in competitive junior tennis players. *Med Sci Tennis*. 2006;1:10–11.
4. Fernandez J, Mendez-Villanueva A, Pluim BM. Intensity of tennis match play. *Br J Sports Med*. 2006;40(5):387–391.
5. Ellenbecker TS, Roetert EP, Sueyoshi T, Riewald S. A descriptive profile of age-specific knee extension flexion strength in elite junior tennis players. *Br J Sports Med*. 2007;41(11):728–732.
6. Pasanen K, Parkkari J, Pasanen M, et al. Neuromuscular training and the risk of leg injuries in female floorball players: cluster randomised controlled study. *BMJ*. 2008;337:a295.
7. Lehance C, Binet J, Bury T, Croisier JL. Muscular strength, functional performances and injury risk in professional and junior elite soccer players. *Scand J Med Sci Sports*. 2009;19(2):243–251.
8. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. Upper Saddle River, NJ: Pearson Prentice Hall; 2009.
9. Muller E, Benko U, Raschner C, Schwameder H. Specific fitness training and testing in competitive sports. *Med Sci Sports Exerc*. 2000;32(1):216–220.
10. Alricsson M, Harms-Ringdahl K, Werner S. Reliability of sports related functional tests with emphasis on speed and agility in young athletes. *Scand J Med Sci Sports*. 2001;11(4):229–232.
11. Kovacs MS, Pritchett R, Wickwire PJ, Green JM, Bishop P. Physical performance changes after unsupervised training during the autumn/spring semester break in competitive tennis players. *Br J Sports Med*. 2007;41(11):705–710.
12. Kuzmits FE, Adams AJ. The NFL combine: does it predict performance in the National Football League? *J Strength Cond Res*. 2008;22(6):1721–1727.
13. Sekulic D, Spasic M, Mirkov D, Cavar M, Sattler T. Gender-specific influences of balance, speed, and power on agility performance. *J Strength Cond Res*. 2013;27(3):802–811.
14. Stewart PF, Turner AN, Miller SC. Reliability, factorial validity, and interrelationships of five commonly used change of direction speed tests. *Scand J Med Sci Sports*. 2014;24(3):500–506.
15. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Oxford: Oxford University Press; 2008.
16. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–310.
17. Currell K, Jeukendrup AE. Validity, reliability and sensitivity of measures of sporting performance. *Sports Med*. 2008;38(4):297–316.
18. Mirkov D, Nedeljkovic A, Kukulj M, Ugarkovic D, Jaric S. Evaluation of the reliability of soccer-specific field tests. *J Strength Cond Res*. 2008;22(4):1046–1050.
19. Sporis G, Jukic I, Milanovic L, Vucetic V. Reliability and factorial validity of agility tests for soccer players. *J Strength Cond Res*. 2010;24(3):679–686.
20. Markovic G, Dizdar D, Jukic I, Cardinale M. Reliability and factorial validity of squat and countermovement jump tests. *J Strength Cond Res*. 2004;18(3):551–555.
21. Vicente-Rodriguez G, Rey-Lopez JP, Ruiz JR, et al. Interrater reliability and time measurement validity of speed-agility field tests in adolescents. *J Strength Cond Res*. 2011;25(7):2059–2063.
22. Mayhew JL, Houser JJ, Briney BB, Williams TB, Piper FC, Brechue WF. Comparison between hand and electronic timing of 40-yd dash performance in college football players. *J Strength Cond Res*. 2010;24(2):447–451.
23. Kovacs MS, Roetert EP, Ellenbecker TS. Efficient deceleration: the forgotten factor in tennis-specific training. *J Strength Cond Res*. 2008;30:50–69.
24. Dugan SA. Sports-related knee injuries in female athletes: what gives? *Am J Phys Med Rehabil*. 2005;84(2):122–130.

Open Access Journal of Sports Medicine

Publish your work in this journal

Open Access Journal of Sports Medicine is an international, peer-reviewed, open access journal publishing original research, reports, reviews and commentaries on all areas of sports medicine. The manuscript management system is completely online and includes a very quick and fair peer-review system.

Submit your manuscript here: <http://www.dovepress.com/open-access-journal-of-sports-medicine-journal>

Dovepress

Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.