Thomas Wyss[1], Bernard Marti[1], Sandro Rossi[1], Urs Kohler[2], Urs Mäder[1]

[1] Swiss Federal Institute of Sports, Magglingen, Switzerland
[2] Land Forces Training Units, Logistic Training Unit, Swiss Armed Forces, Switzerland

# Assembling and Verification of a Fitness Test Battery for the Recruitment of the Swiss Army and Nation-wide Use

## Abstract

The aim of this study was to assess the reliability, validity and feasibility of selected physical performance tests, to compile a new fitness test battery based on these results and to obtain standard values for young men. 79 men (20.3 ± 1.1 y) performed the tests for the reliability part, while 60 men (20.3 ± 1.1 y) completed the tests for the validity part of the study. Feasibility was confirmed by 25 sport experts who conducted the test battery among 1704 men (19.5 ± 1.0 y). For standard values, the data of 12 862 men (19.9 ± 1.0 y) were collected. Based on the reliability and validity data, the following 5 tests were selected for the fitness-test battery: 1) progressive endurance run, 2) seated 2-kg-shot put, 3) standing long jump, 4) trunk muscle strength test and 5) 1-leg standing test. The reliability and validity of the selected performance tests were sufficient to very good (r = 0.50–0.90 and r = 0.64–0.91, respectively). The suggested fitness-test battery can be applied among large groups.

## Zusammenfassung

Die vorliegende Studie hatte zum Ziel, die Wiederholbarkeit, Validität und Durchführbarkeit ausgewählter körperlicher Leistungstests zu erheben und daraus eine neue Fitness-Test-Batterie mit Normwerten für junge Männer zusammenzustellen. 79 Männer (20.3 ± 1.1 J.) absolvierten die Tests zweimal, um daraus die Wiederholbarkeit abzuleiten. Weitere 60 Männer (20.3 ± 1.1 J.) absolvierten die Validierungsmessungen. Die Durchführbarkeit wurde durch 25 Sportexperten beurteilt, welche bei insgesamt 1704 Männern (19.5 ± 1.0 J.) die Fitness-Test-Batterie anwendeten. Für die Normwerte wurden die Leistungen von 12 862 Männern (19.9 ± 1.0 J.) erfasst. Aufgrund der Resultate betreffend Wiederholbarkeit und Validierung wurden die folgenden 5 Disziplinen für die Fitness-Test-Batterie ausgewählt: 1) progressiver Ausdauerlauf, 2) 2-kg-Medizinballstoss aus dem Sitzen, 3) Standweitsprung, 4) globaler Rumpfkrafttest und 5) Einbeinstand. Die Wiederholbarkeit und Validität der ausgewählten Tests war genügend bis sehr gut (r = 0.50–0.90 bzw. r = 0.64–0.91). Die vorgeschlagene Fitness-Test-Batterie kann in grossen Gruppen durchgeführt werden.

Schweizerische Zeitschrift für «Sportmedizin und Sporttraumatologie» 55 (4), 126–131, 2007

## Introduction

In terms of public health, the monitoring of the population's fitness level is important because it is positively related to health outcomes. Since 1874 a physical performance test has been part of the recruitment process in the compulsory Swiss Army. Every 19 year-old male Swiss citizen has to participate. The former performance tests have never been evaluated scientifically. The need for a feasible, valid and standardised physical fitness test battery for population monitoring, education and occupational medicine has been growing lately in Switzerland. The previously used and other known physical fitness test batteries are either too time-consuming, or they contain tests with limited reliability or validity. Therefore, a Swiss physical fitness test battery (SPFTB) for nation-wide use in young men was developed. SPFTB should a) permit an evaluation of large groups of young men in a limited time period, b) be feasible with a minimum of material, c) comply with scientific criteria of validity and reliability, and d) contain relevant performance- and health-related components of physical fitness.

In 1951, Cureton (1951) defined physical fitness as the degree of balance, flexibility, agility (speed), strength, power and endurance. Several definitions of physical fitness have since been published. Miller et al. (1991) defined physical fitness in more general terms as the level of ability to perform sustained physical work characterized by an effective integration of cardiorespiratory endurance, strength, flexibility, coordination and body composition. Physical fitness is a multidimensional construct and therefore cannot be assessed by a single test. Thus a battery of different tests is needed.

According to our definition of physical fitness, the test battery should measure health-related factors and aspects of performance. A complete fitness-test battery should therefore assess a) cardiorespiratory endurance, b) muscle strength and endurance, and c) agility and balance.

A reliable and valid measure of cardiorespiratory endurance is the maximal oxygen consumption ($VO_2$max; Safrit et al., 1988), although it is not immune to inaccuracy (Shepard, 1984). In larger population groups, direct measurement of $VO_2$ is not feasible since expensive equipment is required. Therefore, a variety of less complex tests to measure cardiorespiratory endurance was developed. The frequently used 12-min run test (12-MRT), also referred to as Cooper test, is strongly related to the criterion measure of $VO_2$max in adults (r = 0.84–0.92; Cooper, 1968; Grant et al., 1995; McCutcheon et al., 1990). The 12-MRT is appropriate for individuals with a sufficient fitness level and requires considerable motivation and experience for self-pacing. The multistage 20-m shuttle run test (MST) is more appropriate for individuals without experience for self-pacing. The MST paces the participants by an acoustic signal and has shown to be an accurate method to estimate $VO_2$max in adults (r = 0.79–0.90; Cooper, 1968; Grant et al., 1995; Leger et al., 1989; McNaughton et al., 1998; Ramsbottom et al., 1988).

However, the frequent stopping and starting may limit the application for individuals with less developed motor skills, especially at higher speed levels. A paced progressive endurance run on a track could be a more appropriate method to measure cardiorespiratory endurance among heterogeneous population groups.

To estimate the muscle power in the upper and lower extremities, peak power during bench press throw, squat jump (SJ) and counter movement jump (CMJ) are evaluated. In larger population groups, this direct measurement of muscle power is not feasible. Therefore, less complex tests were developed to measure muscle power. The simple shot put is a feasible and valid test to measure the power of upper extremities in larger groups. The seated 4.5-kg-shot put, conducted by members of a weight training class, correlated positively (r = 0.75) with their power during bench press throw of 60% of their 1-repetition maximum (Mayhew et al., 1991). The distance of the seated 0.4 kg chest pass, conducted by women of a netball team, correlated significantly with peak power during bench press throw of 10 kg (r = 0.80; Cronin and Owen, 2004). Based on these results, we suggest that a weight of 2 kg could be appropriate for seated shot put in a heterogeneous population of young men.

To measure the power of lower extremities, high or long jumps are feasible tests. The vertical jump-and-reach score is a good predictor for the power in the leg extensor muscles (r = 0.93 resp. r = 0.91; Sayers et al., 1999). The standing long jump (SLJ) is widely used because of the good feasibility and test-retest reliability (r = 0.89–0.95; Markovic et al., 2004; Tsigilis et al., 2002). The correlation between SLJ and the principal component of explosive power is good (r = 0.76; Markovic et al., 2004). Concerning the assessment of maximal running speed, a pendulum sprint is a well known method, which can be applied in a gym hall. However, its validity has yet to be demonstrated.

Sit-ups are often used to measure the muscular strength and endurance of the abdominal muscle groups. Some studies showed limited reliability of dynamic or isometric sit-up tests (r < 0.50; Sparling et al., 1997; Suni et al., 1996), while others indicated satisfactory reliability (r = 0.72–0.84; DiNucci et al., 1990; Erbaugh, 1990; Tsigilis et al., 2002). Available data suggest that sit-ups yield limited to acceptable measurements of trunk muscle strength and endurance (r = 0.23–0.66; Knapik, 1989). Sit-ups may involve varying accessory muscles besides abdominal muscles, such as the hip flexors. Therefore curl-up testing was selected to minimise the use of the hip flexors. While the reliability of dynamic or isometric sit-up tests seems to be limited, the curl-up test reached a good reliability (r = 0.92; Sparling et al., 1997). However, curl-up tests were criticized because it can be difficult to judge whether they are carried out correctly. The trunk muscle strength test *(figure 1)* could be an interesting alternative to measure global muscular strength and endurance of the trunk. It is a part of the standardised dynamic trunk muscle test battery of the Swiss Olympic Medical Centres (Bourban et al., 2001; Tschopp et al., 2001). Its reliability was determined among athletes only (r = 0.87) and the authors judged the trunk muscle strength test to be valid to acquire health-related minimum requirements for elite athletes (Tschopp et al., 2001).

Motor skills are evaluated by determining test-retest reliability as for these test items a gold standard for validation is missing. The 1-leg standing test (OLS) is a feasible test to assess balance as a motor ability. Its interrater reliability (kappa value = 0.90) and the test-retest reliability (r = 0.73) are good and this test is further more valid for predicting ankle sprains in college students (Trojian and McKeag, 2006; Tsigilis et al., 2002). One limitation in evaluating balance is its specificity (Tsigilis et al., 2002). Therefore, we assume a bipart-balance test (static and dynamic) would be even more valid to predict injuries on lower extremities.

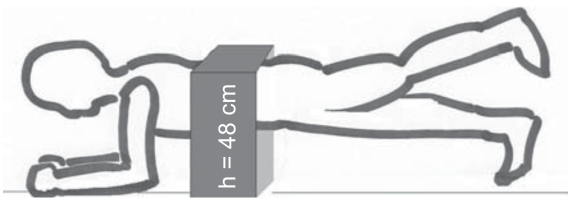


*Figure 1:* Trunk muscle strength test: standardised body position with a laterally open box.

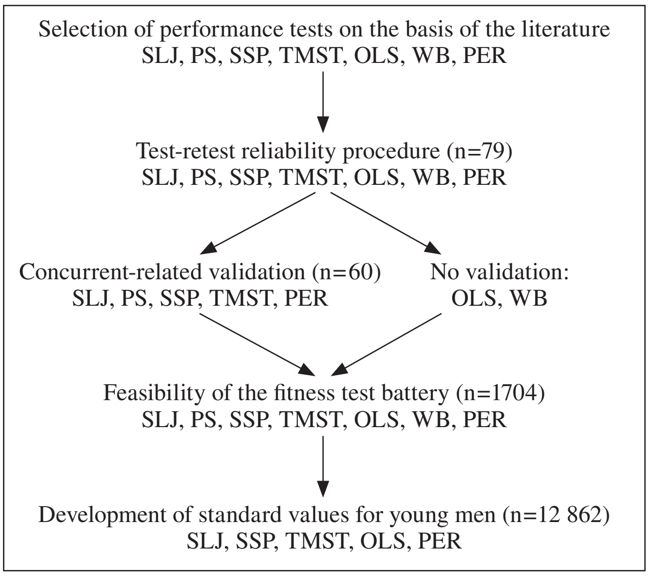


*Figure 2:* Schematic of the study design. SLJ, Standing long jump; PS, pendulum sprint; SSP, seated shot put; TMST, trunk muscle strength test; OLS, 1-leg standing; WB, walking on a beam; PER, progressive endurance run.

*Physical fitness-test battery*

To measure physical fitness among larger population groups, a feasible fitness-test battery is needed. Widely used fitness-test batteries for young adults are the health-related physical fitness test (HRPFT), the Eurofit test battery, the US Army Physical Fitness Test (APFT) and the health-related fitness test battery (HRFI). HRPFT contains assessments of cardiorespiratory endurance, abdominal muscle strength, flexibility and body composition (AAHPERD, 1980). The Eurofit test battery includes 9 motor fitness tests (cardiorespiratory endurance, muscular strength, endurance and speed, flexibility and balance) and 5 anthropometric measurements (Adam et al., 1988). APFT consists of cardiorespiratory endurance, abdominal and upper body strength and endurance (Knapik, 1989). HRFI contains cardiorespiratory endurance, muscular power and strength, trunk muscular endurance and balance (Suni et al., 1996). These test batteries are either focused on health-related or performance-related outputs, they are too time consuming or they contain tests with limited reliability or validity. Therefore, SPFTB to assess health- and performance-related outputs in larger groups of young men was developed.

The aim of this study was to assess reliability, validity and feasibility of selected physical performance tests and to assemble a feasible fitness-test battery for young men.

## Methods

***Study design***

The evaluation of SPFTB took place in 4 parts *(figure 2)*. First, the test-retest reliability of each performance test was assessed with a time interval of 7 days between measurements. Second, their concurrent validity was assessed with a time interval of 7 days between assessments with the sequence of field and laboratory tests being randomised. Then, the feasibility of the performance tests was assessed at a military recruitment centre in the French and German speaking parts of Switzerland. Last of all, standard values for young men were developed during compulsory Swiss Army recruitment with the data of all conscripts during 6 months.

***Subjects***

In the reliability part of the study, 79 men (20.3 ± 1.1 y, 76.8 ± 13.5 kg, 179.9 ± 7.1 cm) completed all performance tests twice. 60 men (20.3 ± 1.1 y, 76.7 ± 15.0 kg, 179.5 ± 6.6 cm) completed all performance tests in the validity part of the study. The feasibility of SPFTB was investigated among 1704 male draftees (19.5 ± 1.0 y, 72.7 ± 11.8 kg, 177.9 ± 6.5 cm).

Finally, standard values were obtained by the data of 15 794 conscripts who had to pass their recruitment to the Swiss Army. The entire fitness-test battery was completed by 81.4% of the conscripts (n = 12 862, 19.9 ± 1.0 y, 178.3 ± 15.9 cm, 72.8 ± 12.0 kg), while others were fully or partially exempted from fitness testing due to medical reasons.

***Physical performance tests***

*Cardiorespiratory endurance*

The progressive endurance run (PER) was conducted on an outdoor track. Every 10 m, a marker was placed on the track. Every subject started from another 10 m marker at the same time. An acoustic signal paced the running velocity. The subjects had to pass the next 10 m marker simultaneously with the acoustic signal. Paced velocity started at 8.5 km/h and increased 0.5 km/h every 200 m. Total running time was registered when the subject could no longer hold the given pace.

11 subjects of the reliability part and 16 subjects of the validity part of the study had to be excluded from endurance assessment, because they refused to run on one or both courses.

*Muscle power*

The power of upper extremities was assessed by a shot-put performance test. The seated 2-kg-shot put (SSP) was performed as a chest pass. The subjects were sitting upright on a bench of 38 cm height and their back was in contact with a vertical wall. They had to hold the position, while performing the shot put. The distance between the wall and the landing point was registered. The best of 3 trials was valued with an accuracy of 1 cm.

SLJ was performed from the gym hall floor onto a mat of 7 cm height to assess the power of lower extremities. The distance was measured from the scratch line to the closest point of body-contact on the landing mat. The best of 3 trials was valued with an accuracy of 1 cm.

The results of the pendulum sprint (PS) (4 x 10 m) were used to determine running speed. Each time the subjects had to step over the 10-m line before turning around. The better of 2 trials was valued with an accuracy of 0.1 s.

*Trunk muscle strength*

In the trunk muscle strength test (TMST), subjects had to support their body on forearms and feet, while keeping the upper body and the legs in a straight line as long as possible. They had to lift their feet alternately by the 1 Hz rhythm of a metronome. The body position was standardised and controlled with a laterally opened box *(figure 1)*. The test ended as soon as the subjects were not able to keep the prescribed body position. Time was recorded with an accuracy of 1 s. This test was a simplified adaptation (without original height adjustable positioning-rack and head restraint) of the trunk muscle strength test published elsewhere (Tschopp et al., 2001).

*Balance*

Static balance was assessed with OLS. The free foot had to be in contact with the hollow of the knee of the standing leg and the hands had to hold each other behind the back. After 10 s, the eyes had to be closed. After another 10 s, the head had to be laid back without opening the eyes. Time was stopped, when another part of the body, other than the standing foot, had contact with the floor or the standing foot lost contact with the floor or the eyes were opened or the hands were released. For those who did not lose balance for 1 min, maximal time of 60 s was registered. Time was measured for both legs and valued with an accuracy of 0.1 s.

Dynamic balance was tested by walking forward and backward on a beam (WB) (length: 2.6 m, width: 0.1 m, height: 0.38 m). Subjects had to walk as fast as possible forward to the end of the beam and backward over the middle of it. Subjects who lost balance and descended were immediately showed where to step back on the beam. Time was measured with an accuracy of 0.1 s.

***Validation of the test battery***

*Cardiorespiratory endurance*

To validate PER, a 12-MRT and a $VO_2$-peak-test were conducted. Previously, all subjects had performed the 12-MRT at least twice. 18 subjects (20.4 ± 1.3 y, 73.3 ± 10.8 kg, 180.0 ± 7.6 cm) from the validation part of the study were randomly selected to measure $VO_2$-peak on a treadmill. These additional assessments were conducted during the following 3 weeks after the validation of the test battery. The start velocity was chosen individually, related to the performance in PER, between 8.5 and 14.0 km/h. Treadmill velocity was increased 0.5 km/h every min. Subjects were asked to run until exhaustion. $VO_2$ was measured with Oxycon Pro (Jäger, Hoechberg, Germany). The maximal value was recorded relative to body weight as $VO_2$peak. A linear regression between the peak running velocity of PER and $VO_2$peak of the treadmill test was calculated to estimate the $VO_2$peak of all subjects.

*Muscle power*

The results of SSP were validated against maximal power performance during free-weight bench press. Subjects had to push a 15 kg and a 30 kg barbell. The barbell was lowered slowly to touch the chest, held there for 1 s and then pushed to full arm extension as fast as possible. Force plates (MLD2, SPSport, Innsbruck, Austria) were attached under the bench to calculate the maximal power output during bench press. The best of 3 trials was registered for each weight.

Maximal power relative to body weight was assessed during SJ and CMJ on a force plate MLD2 and related to the results of SLJ. The best of 3 results was recorded for each jump.

The performance of PS was compared to the running speed during a straight 40-m sprint. Light barriers recorded the sprint time between 30 and 40 m as a value for calculating the running speed. The best of 3 trials was valued with an accuracy of 0.001 s.

*Trunk muscle strength*

TMST was conducted after the protocol of Swiss Olympic Medical Centres (Tschopp et al., 2001) with specially trained physiotherapists, using a height adjustable positioning rack and head restraint.

***Feasibility Questionnaire***

A questionnaire was used in the feasibility part of the study. 25 responsible sport experts in recruitment centres were asked to rate the given statement, „the performance test is easily practicable“, for each performance test individually. In a second part, more feedback including information on the duration of the complete test battery was collected through open questions.

***Data analysis***

For every discipline, the measured values were directly used for data analysis, except for balance. The total score for balance was

| n | performance test 1 | vs | performance test 2 | difference | correlation |
|---|---|---|---|---|---|
| 68 | progressive endurance run (PER) 1<br>14.08 ± 2.02 km/h | | PER 2<br>14.19 ± 2.08 km/h | 0.11 km/h<br>($p = 0.14$) | $r = 0.89$ ($p < 0.001$) |
| 79 | standing long jump (SLJ) 1<br>2.26 ± 0.22 m | | SLJ 2<br>2.26 ± 0.23 m | 0.00 m<br>($p = 0.96$) | $r = 0.90$ ($p < 0.001$) |
| 79 | pendulum sprint (PS) 1<br>10.99 ± 0.72 s | | PS 2<br>10.80 ± 0.78 s | 0.19 s<br>($p < 0.001$) | $r = 0.84$ ($p < 0.001$) |
| 79 | seated shot put (SSP) 1<br>6.58 ± 0.65 m | | SSP 2<br>6.60 ± 0.67 m | 0.02 m<br>($p = 0.64$) | $r = 0.83$ ($p < 0.001$) |
| 79 | trunk muscle strength test (TMST) 1<br>01:35 ± 00:45 [min:sec] | | TMST 2<br>01:32 ± 00:49 [min:sec] | –3 s<br>($p = 0.44$) | $r = 0.77$ ($p < 0.001$) |
| 60 | 1-leg standing (OLS) 1<br>38.88 ± 8.14 s | | OLS 2<br>43.23 ± 12.08 s | 4.35 s<br>($p < 0.01$) | $r = 0.50$ ($p < 0.001$) |
| 60 | static and dynamic balance 1<br>27.78 ± 9.44 s | | static and dynamic balance 2<br>29.35 ± 14.03 s | 1.56 s<br>($p = 0.32$) | $r = 0.57$ ($p < 0.001$) |

*Table 1:* Reliability of the physical performance tests, means, standard deviations, differences and correlation coefficients.

| n | performance test 1 | vs | performance test 2 | correlation |
|---|---|---|---|---|
| 44 | progressive endurance run (PER)<br>13.60 ± 2.03 km/h | | 12-min run test (12-MRT)<br>2303.45 ± 414.95 m | $r = 0.91$ ($p < 0.001$) |
| 18 | PER<br>14.36 ± 2.26 km/h | | $VO_2$peak<br>49.93 ± 6.01 ml/min/kg | $r = 0.84$ ($p < 0.001$) |
| 60 | standing long jump (SLJ)<br>2.14 ± 0.24 m | | jump on force plate<br>SJ: 44.08 ± 6.03 W/kg<br>CMJ: 46.35 ± 6.98 W/kg | <br>$r = 0.64$ ($p < 0.001$)<br>$r = 0.61$ ($p < 0.001$) |
| 60 | SLJ<br>2.14 ± 0.24 m | | sprint 30–40m<br>1.33 ± 0.14 s | $r = -0.73$ ($p < 0.001$) |
| 60 | SLJ<br>2.14 ± 0.24 m | | pendulum sprint (PS)<br>10.98 ± 0.81 s | $r = -0.73$ ($p < 0.001$) |
| 60 | pendulum sprint (PS)<br>10.98 ± 0.81 s | | sprint 30–40m<br>1.33 ± 0.14 s | $r = 0.85$ ($p < 0.001$) |
| <br>57<br>45 | seated shot put (SSP)<br>6.10 ± 0.69 m<br>6.12 ± 0.73 m | | bench press power<br>15kg: 369.06 ± 80.67 W<br>30kg: 362.83 ± 108.17 W | <br>$r = 0.54$ ($p < 0.001$)<br>$r = 0.65$ ($p < 0.001$) |
| 60 | trunk muscle strength test (TMST)<br>01:19 ± 00:54 [min:sec] | | trunk muscle strength test SOMC<br>01:22 ± 00:59 [min:sec] | $r = 0.85$ ($p < 0.001$) |

*Table 2:* Validity of the physical performance tests, means, standard deviations and correlation coefficients, SJ = squat jump, CMJ = counter movement jump, SOMC = Swiss Olympic Medical Centres.

calculated adding the time of left and right OLS and then subtracting twice the time for WB ($t_l + t_r - 2t_{WB}$).

All statistical analyses were done with the program SPSS 13.0 (SPSS, Chicago, Illinois, USA). Pearson's correlation coefficients were calculated to estimate the relation between datasets. The t-test was used to estimate the significance of differences in repeated measurements. Descriptive analyses were done on the questionnaire data.

## Results

### *Physical performance tests*

#### *Cardiorespiratory endurance*

PER was highly reproducible *(table 1)*. Additionally, PER time correlated positively with the distance run in the 12-MRT and with the $VO_2$peak in the maximal treadmill exercise test *(table 2)*.

$VO_2$peak may be estimated by the peak running velocity of the PER with the following regression: $VO_2$peak [ml*kg$^{-1}$*min$^{-1}$] = 2.309 * Velocity$_{peak}$ [km/h] + 16,549, ($p<.001$).

#### *Muscle power*

SSP yielded a good reliability *(table 1)*. In 15 subjects, the force plates were not able to register the maximal power output during the bench press since the weight was never accelerated fast enough. SSP had a positive correlation with the maximal bench press power *(table 2)*.

SLJ was highly reproducible *(table 1)*. SLJ correlated positively with relative maximal power during normalised jumps. PS was inversely correlated with SLJ *(table 2)*. While the performance of SLJ did not differ between both trials, the performance in PS was better in the second trial *(table 1)*.

#### *Trunk muscle strength*

TMST yielded a good reliability *(table 1)*. The simplified TMST and the original one correlated positively *(table 2)*.

#### *Balance*

The repeated combination of static and dynamic balance tests generated a moderate correlation coefficient *(table 1)*. The results in both cycles were not different. The retest of OLS alone yielded

| percentile | BMI [kg/m²] | PER [min:s] | pdt $VO_2$peak [ml*kg$^{-1}$*min$^{-1}$] | SSP [m] | SLJ [m] | TMST [s] | OLS ($t_l$+$t_r$) [s] |
|---|---|---|---|---|---|---|---|
| 5 | 18.61 | 06:21 | 39.94 | 5.30 | 1.93 | 41.00 | 30.00 |
| 10 | 19.32 | 07:58 | 42.12 | 5.55 | 2.03 | 54.00 | 32.50 |
| 15 | 19.84 | 09:10 | 43.74 | 5.70 | 2.10 | 63.00 | 35.00 |
| 20 | 20.28 | 09:45 | 44.53 | 5.83 | 2.15 | 71.00 | 37.00 |
| 25 | 20.66 | 10:27 | 45.47 | 5.95 | 2.20 | 79.00 | 38.30 |
| 30 | 20.98 | 11:13 | 46.51 | 6.05 | 2.23 | 87.00 | 40.00 |
| 35 | 21.31 | 11:30 | 46.86 | 6.15 | 2.26 | 96.00 | 41.00 |
| 40 | 21.67 | 12:15 | 47.90 | 6.25 | 2.30 | 105.00 | 42.40 |
| 45 | 22.01 | 12:35 | 48.35 | 6.35 | 2.32 | 108.00 | 43.80 |
| 50 | 22.34 | 13:07 | 49.07 | 6.45 | 2.35 | 113.00 | 45.00 |
| 55 | 22.71 | 13:20 | 49.36 | 6.52 | 2.39 | 119.00 | 46.30 |
| 60 | 23.06 | 13:56 | 50.17 | 6.60 | 2.40 | 125.00 | 47.70 |
| 65 | 23.46 | 14:15 | 50.60 | 6.70 | 2.43 | 132.00 | 49.00 |
| 70 | 23.94 | 14:44 | 51.25 | 6.80 | 2.45 | 140.00 | 50.00 |
| 75 | 24.49 | 15:10 | 51.83 | 6.90 | 2.49 | 151.00 | 52.00 |
| 80 | 25.14 | 15:31 | 52.31 | 7.00 | 2.52 | 163.00 | 54.00 |
| 85 | 25.93 | 16:16 | 53.32 | 7.17 | 2.56 | 182.00 | 56.60 |
| 90 | 27.12 | 17:00 | 54.31 | 7.38 | 2.60 | 202.00 | 61.00 |
| 95 | 29.37 | 17:46 | 55.34 | 7.70 | 2.68 | 240.00 | 70.00 |

*Table 3:* Fitness-test battery standard values for young men (n = 12 862, 19.9 ± 1.0 y, 178.3 ± 15.9 cm, 72.8 ±12.0 kg). BMI, body mass index; SLJ, standing long jump; SSP, seated 2-kg-shot put; TMST, trunk muscle strength test; OLS, 1-leg standing; PER, progressive endurance run; pdt $VO_2$peak, predicted peak oxygen consumption.

a moderate reliability *(table 1)*. OLS results tended to be better in the repeated test *(table 1)*.

### *Feasibility*

Based on the results of the reliability- and validity part of this study PER, SLJ, SSP, TMST and bipart-balance test (OLS and WB) were selected for SPFTB. 3 sport experts needed less than 90 min to conduct the fitness test battery with 30 subjects, including information and warm up. The tests were rated as easily practicable or fairly easily practicable from 68% for SSP to 100% for SLJ. As an exception, the feasibility of WB was rated to be poor. In the open questions, 10 sport experts (40%) described WB as potentially dangerous or not well standardized.

### *Standard values*

The representative SPFTB standard values (n = 12 862) for young men are presented in *table 3*.

## Discussion

### *Physical performance tests*

#### *Cardiorespiratory endurance*

Peak treadmill running velocity, during a speed-incremented $VO_2$peak test, is an effective predictor of endurance performance (Harling et al., 2003). According to Noakes et al. (1990), peak treadmill running velocity is the best laboratory-measured predictor of running performance. The present study shows that peak running velocity in a progressive endurance run also reaches good reproducibility and validity if assessed as a field test on a track. PER was rated as feasible by sport experts. Additionally, PER may be more appropriate for individuals with low fitness levels or no experience in self-pacing compared to 12-MRT. However, further studies are needed to proof this assumption. We conclude that PER is an appropriate assessment of endurance capacity, especially for larger heterogeneous population groups.

#### *Muscle power*

SSP generated a good correlation with maximal power during bench press. Our results show that SSP is feasible, reliable and valid for young men. However, further research is needed to investigate the relationship between SSP of varying loads and the bench press power.

SLJ and PS are both valid and reliable. PS seems to be sensitive to learning effects, as the performance was better in the second measurement. This could be due to a learning effect in terms of agility. The comparison of PS with SLJ shows a strong relationship *(table 2)*. Baker and Nance (1999) and Cronin and Hansen (2005) also found an inverse correlation between power and speed by comparing the jump height in CMJ with the 30 m sprint time ($r = -0.56$, $p < 0.05$) and the relative leg power with the 40 m sprint time ($r = -0.76$, $p < 0.05$). For an inexpensive physical performance test it is therefore reasonable to conduct either SLJ or PS. SLJ has previously been widely used and validated ($r = 0.76$; Markovic et al., 2004). Hence the power of lower extremities and running speed can be assessed with the reproducible, valid and feasible SLJ.

#### *Trunk muscle strength*

TMST is, due to its good reliability, validity and feasibility, suitable for use among larger groups of young men. However, the good validity has to be interpreted carefully because no gold standard for trunk muscle strength is available. In this study, TMST was compared with the established but more expensive TMST of the Swiss Olympic Medical Centres (Tschopp et al., 2001). We recommend using a height adjustable positioning-rack for anthropometrically heterogeneous groups.

#### *Balance*

To assess balance, a static and dynamic balance test was combined. This bipart-balance test is reproducible but WB seems not to be feasible. Therefore, we suggest using OLS alone. Our OLS is reproducible and a similar 1-leg balance test was shown to be valid for predicting ankle sprains (Trojian and McKeag, 2006). Therefore, only OLS was included in SPFTB. Further research is needed to find a feasible and reliable dynamic balance test.

### *Limitations and strength of the study*

The subject's motivation in the reliability- and validity part of this study may have been heterogeneous. It can be expected that results would be even better with highly motivated subjects. Especially, for the cardiorespiratory endurance tests, where 11 (reliability-part) and 16 (validity-part) subjects refused maximal performance, motivation is crucial.

2 different samples of subjects were used for the validity and reliability part of the study. This limitation was accepted in order to obtain more subjects for participation in the study. The 2 study groups are comparable. They do not differ in age, weight, height and all performance tests except SSP and SLJ (data not shown).

The balance test could not be validated because no gold standard to assess motor skills is available according to the authors' knowledge. For a balance test a good reproducibility is already a challenge.

All measurements were done with a high number of subjects but exclusively with young men. Therefore, no data for other population groups are available so far. The replication of the present reliability, validity and feasibility studies for younger boys and girls will be the object of further research.

A major strength of the present study is that 3 important aspects of the fitness tests were assessed: reliability, validity and feasibility. In addition, standard values for future reference were collected with a representative sample among 19 year old men.

## Conclusion

The new health- and performance-related SPFTB is qualified for nation wide use. SPFTB meets the previously specified demands; a) 3 sport experts are able to assess 30 subjects in 90 min, b) no expensive material is needed, c) the tests are valid and reliable and d) the most relevant components of fitness are included. With the 5 disciplines of SPFTB, changes in the physical fitness of specific population groups can be monitored. The new fitness-test battery may be attractive for epidemiological research, physical education, sport clubs and could be of interest in the field of occupational medicine (selection and control of employees in physically demanding jobs such as public service personnel and military).

Address for correspondence:

Thomas Wyss, Swiss Federal Institute of Sports, 2532 Magglingen, Tel. +41 (0)32 327 64 16, e-mail: thomas.wyss@baspo.admin.ch

## Literature

AAHPERD (1980): Health Related Physical Fitness Test Manual. American Alliance for Health, Physical Education, Recreation and Dance, Reston, Virginia.

Adam C., Klissouras V., Ravazzolo M., Renson R., Tuxworth W. (1988): Eurofit, European test of physical fitness. Council of Europe, Committee for the Development of Sport, Rome, Italy.

Baker D., Nance S. (1999): The relation between running speed and measures of strength and power in professional rugby league players. J. Strength Cond. Res. 13: 230–235.

Bourban P., Hübner K., Tschopp M., Marti B. (2001): Grundkraftanforderungen im Spitzensport: Ergebnisse eines 3-teiligen Rumpfkrafttests. Schweiz. Zschr. Sportmed. Sporttraumat. 49: 73–78.

Cooper K.H. (1968): A means of assessing maximal oxygen intake. Correlation between field and treadmill testing. J. Am. Med. Assoc. 203: 201–204.

Cooper S.M., Baker J.S., Tong R.J., Roberts E., Hanford M. (2005): The repeatability and criterion related validity of the 20 m multistage fitness test as a predictor of maximal oxygen uptake in active young men. Br. J. Sports Med. 39: e19.

Cronin J.B., Hansen K.T. (2005): Strength and power predictors of sports speed. J. Strength Cond. Res. 19: 349–357.

Cronin J.B., Owen G.J. (2004): Upper-body strength and power assessment in women using a chest pass. J. Strength Cond. Res. 18: 401–404.

Cureton T. (1951): Physical fitness of champion athletes. Urbana: University of Illinois Press. 1–458.

DiNucci J., McCune D., Shows D. (1990): Reliability of a modification of the health-related physical fitness test for use with physical education majors. Res. Q. Exerc. Sport 61: 20–25.

Erbaugh S.J. (1990): Reliability of physical fitness tests administered to young children. Percept Mot Skills 71: 1123–1128.

Grant S., Corbett K., Amjad A.M., Wilson J., Aitchison T. (1995): A comparison of methods of predicting maximum oxygen uptake. Br. J. Sports Med. 29: 147–152.

Harling S.A., Tong R.J., Mickleborough T.D. (2003): The oxygen uptake response running to exhaustion at peak treadmill speed. Med. Sci. Sports Exerc. 35: 663–668.

Knapik J. (1989): The Army Physical Fitness Test (APFT): a review of the literature. Mil. Med. 154: 326–329.

Leger L., Gadoury C. (1989): Validity of the 20 m shuttle run test with 1 min stages to predict $VO_2$max in adults. Can. J. Sport Sci. 14: 21–26.

Markovic G., Dizdar D., Jukic I., Cardinale M. (2004): Reliability and factorial validity of squat and countermovement jump tests. J. Strength Cond. Res. 18: 551–555.

Mayhew J., Bemben M., Rohrs D., Ware J., Bemben D. (1991): Seated shot put as a measurement of upper body power in college males. J. Human Mov. Stud. 21: 137–148.

McCutcheon M.C., Sticha S.A., Giese M.D., Nagle F.J. (1990): A further analysis of the 12-minute run prediction of maximal aerobic power. Res. Q. Exerc. Sport 61: 280–283.

McNaughton L., Hall P., Cooley D. (1998): Validation of several methods of estimating maximal oxygen uptake in young men. Percept. Mot. Skills 87: 575–584.

Miller A.J., Grais I.M., Winslow E., Kaminsky L.A. (1991): The definition of physical fitness. A definition to make it understandable to the laity. J. Sports Med. Phys. Fit. 31: 639–640.

Noakes T.D., Myburgh K.H., Schall R. (1990): Peak treadmill running velocity during the $VO_2$ max test predicts running performance. J. Sports Sci. 8: 35–45.

Ramsbottom R., Brewer J., Williams C. (1988): A progressive shuttle run test to estimate maximal oxygen uptake. Br. J. Sports Med. 22: 141–144.

Safrit M.J., Glaucia Costa M., Hooper L.M., Patterson P., Ehlert S.A. (1988): The validity generalization of distance run tests. Can. J. Sport Sci. 13: 188–196.

Sayers S.P., Harackiewicz D.V., Harman E.A., Frykman P.N., Rosenstein M.T. (1999): Cross-validation of three jump power equations. Med. Sci. Sports Exerc. 31: 572–577.

Shepard R.J. (1984): Tests of maximum oxygen intake. A critical review. Sports Med. 1: 99–124.

Sparling P.B., Millard-Stafford M., Snow T.K. (1997): Development of a cadence curl-up test for college students. Res. Q. Exerc. Sport 68: 309–316.

Suni J.H., Oja P., Laukkanen R.T., Miilunpalo S.I., Pasanen M.E., Vuori I.M., Vartiainen T.M., Bos K. (1996): Health-related fitness test battery for adults: aspects of reliability. Arch. Phys. Med. Rehabil. 77: 399–405.

Trojian T.H., McKeag D.B. (2006): Single leg balance test to identify risk of ankle sprains. Br. J. Sports Med. 40: 610–613.

Tschopp M., Bouban P., Hübner K., Marti B. (2001): Messgenauigkeit eines 4-teiligen, standardisierten dynamischen Rumpfkrafttests: Erfahrungen mit gesunden männlichen Spitzensportlern. Schweiz. Zschr. Sportmed. Sporttraumat. 49: 67–72.

Tsigilis N., Douda H., Tokmakidis S.P. (2002): Test-retest reliability of the Eurofit test battery administered to university students. Percept. Mot. Skills 95: 1295–1300.