Doctoral Thesis

Physical activities and demands in Swiss soldiers

Author(s):
Wyss, Thomas

Publication Date:
2010

Permanent Link:
https://doi.org/10.3929/ethz-a-006319065

Rights / License:
In Copyright - Non-Commercial Use Permitted
PHYSICAL ACTIVITIES AND DEMANDS IN SWISS SOLDIERS

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
Thomas Wyss
Dipl. Natw. ETH Zurich
Born August 16, 1977
Citizen of Kappel, Solothurn

accepted on the recommendation of

Prof. Dr. U. Boutellier, examiner
Dr. U. Mäder, co-examiner
Prof. Dr. DP. Wolfner, co-examiner

2010
Acknowledgments

This thesis would not have been possible unless the support of many people. Herewith, I would like to thank everyone, who contributed to the studies and supported me during the last years. My cordial thanks to:

Dr. Urs Mäder, who served as my advisor and mentor, for scientific support from study design until publication. Sincere thanks for many hours of interesting discussions and helpful advices. I have learned so much from you.

Prof. Dr. Urs Boutellier, my dissertation supervisor, and Prof. Dr. David Paul Wolfer, who completed my doctoral thesis committee, for helpful advices during the past years.

Prof. Dr. Ueli Suter, the chair of my dissertation examination, for carefully reading my doctoral thesis.

All recruits of the Swiss Army who volunteered in one of the studies.

The collaborators of the Swiss Army (Reto Albert, Michael Bonadimann, Franz Frey, Paul Greuter, Jean-Jacques Joss, Claude Kaiser, Daniel Koch, Urs Kohler, Philippe Rebord, Sandro Rossi, Christoph Schwager, Rodo von Vigier, and Stephan Zehr, to name the most important exponents) who helped in administrative organization of investigations and for their experts appraisals in discussions of study results.

The Swiss Federal Institute of Sports, Magglingen (SFISM) represented by the rector Walter Mengisen for the opportunity to work in this interesting scientific field.

The members of the Physical Activity and Health Branch at the SFISM for their support during data acquisition, for inspiring methodological discussions, proofreading, and especially for their friendship.

And most important, my wife Bettina, my family, and friends for their patience.
## Contents

1. Summary / Zusammenfassung 2 / 4
2. General introduction 7
4. Assembling and verification of a fitness-test battery for the recruitment of the Swiss Army and nation-wide use 18
5. Physical fitness predicts risk of overuse injuries among Swiss Army recruits 34
6. Recognition of military-specific physical activities with body-fixed sensors 48
7. Energy expenditure estimation during daily military routine with body-fixed sensors 64
8. Estimation of gait velocity based on continuous step count data 77
9. Physical activities and demands in boot camps of different Swiss Army occupational specialties 83
10. General discussion 94
11. References 101
12. Abbreviations 115
13. Curriculum vitae 116
1. Summary

In a military setting, a mismatch between physical fitness and physical job requirements can increase the risk of injury, jeopardize unit military performance, and decrease overall morale. Therefore, an accurate description of physical job requirements and individual physical capacities is crucial. The ideal method to objectively assess physical job requirements is not yet available. In terms of measuring soldiers’ physical capabilities, armed forces have long traditions of physical performance tests. Over time, the physical performance tests in the Swiss Army have been modified repeatedly. However, the former performance tests have never been evaluated scientifically.

It was the aim of the present thesis to develop new methods and adapt existing ones to objectively and reliably assess soldiers’ physical capacities and military job-specific physical demands.

For the assessment of daily physical demands, objective methods were adapted to the military setting to assess activity types, energy expenditure, and distance covered on foot. Using heart rate, acceleration, and step frequency (measured with body-fixed sensors), the most relevant military-specific, physically demanding activities were recognized in one-minute intervals. The sensor-based system achieved an overall recognition rate of 87.5% in a laboratory setting and 85.5% during daily military routine. Based on the activity identification, activity class-specific multiple linear regressions with hip acceleration and heart rate as independent variables were calculated to estimate energy expenditure. Based on step frequency, an algorithm was developed to estimate gait velocity and distance covered on foot. These approaches were considered appropriate. Applied to five different Swiss Army occupational specialties, daily physical demands were quantified. Physical demands in Swiss Army boot camps vary among occupational specialties. However, the average daily total energy expenditure values in all investigated occupational specialties are in the range of values found in other nations’ armed forces as well as in professional sportsmen. An important contrast between elite sports and armed forces is the lack of soldiers’ adequate preparation for their physically demanding daily routines. We
demonstrated that in the investigated Swiss Army occupational specialties, the physical demands (quantified as daily energy expenditure) decreased between the second and the eighth week of basic training.

To develop a new physical fitness-test battery, test-retest reliability, concurrent validity, and feasibility of promising performance tests were investigated. Based on these results, the progressive endurance run, trunk muscle strength test, one-leg standing test, standing long jump, and seated shot-put were chosen as the new test battery. Representative standard values for young men and descriptive power to predict overuse injuries during military service were assessed for the five performance tests. The trunk muscle strength test and the progressive endurance run have shown the strongest discriminative power to predict overuse injuries, while the seated shot-put was the only test that failed to prove its discriminative power to predict overuse injuries in any study group. For different occupational specialties, varying performance tests were discriminative to predict overuse injuries.

We conclude that the fitness-test battery developed in this study is appropriate to detect conscripts with enhanced risk of overuse injuries in physically demanding military occupational specialties. With the used sensors and the developed algorithms, military-specific activities can be recognized. Furthermore, energy expenditure and distance covered on foot can be estimated in one-minute intervals over six continuous days. The presented methods allow investigators to objectively assess soldiers’ physical capacities and job-specific physical demands in terms of type, intensity, duration, and frequency of military-specific physical activities. Therefore, the methods developed in the present thesis may contribute important information to future efforts in a military setting to balance between physical demands, physical fitness, injuries and military performance.
1. Zusammenfassung


Das Ziel der vorliegenden Arbeit ist es, neue Methoden zu entwickeln und bestehende Methoden ans militärische Setting anzupassen, so dass die körperliche Leistungsfähigkeit von Soldaten und deren körperliche Beanspruchung im militärischen Alltag objektiv und zuverlässig erhoben werden können.

Um die körperliche Beanspruchung im militärischen Alltag objektiv zu erheben, wurden Methoden an den militärischen Kontext angepasst, welche dem Erkennen von Aktivitätsklassen und dem Abschätzen von Energieverbrauch sowie zu Fuss zurückgelegter Distanz dienen. Mittels am Körper getragener Sensoren wurden Herzfrequenz, Beschleunigungen und Schrittfrequenz erhoben. Daraus konnten die wichtigsten militärspezifischen, körperlich anspruchsvollen Aktivitäten in 60s-Intervallen abgeleitet werden. In Laborerhebungen erreichte diese Methode eine Gesamterkennungsrate von 87.5% und in Feldmessungen von 85.5%. Basierend auf der Aktivitätserkennung wurden für jede militärspezifische Aktivitätsklasse eine multiple lineare Regression hergeleitet, welche es ermöglicht, den Energieverbrauch aus Herzfrequenz und Beschleunigungen der Hüfte abzuschätzen. Ein weiterer Algorithmus wurde entwickelt, um aus der Schrittfrequenz kontinuierlich die Ganggeschwindigkeit und die zu Fuss zurückgelegte Distanz abzuschätzen. Die Validität dieser Ansätze konnte


Methoden in einem militärischen Umfeld in Zukunft wichtige Informationen, für Bestrebungen zum Erreichen eines Gleichgewichtes zwischen körperlichen Anforderungen, körperlicher Fitness, Verletzungen und militärischer Leistungsfähigkeit liefern.
2. General introduction

Physical fitness, health complaints, physical demands, and military performance are important research fields in a military setting (Figure 2.1). Knowledge about the four single fields and their interrelations is important for the description of the physical job requirements in military occupational specialties. Due to limited resources, the present thesis focused on the three fields of physical fitness, injuries, and physical demands during the Swiss Army’s daily military routine.

![Figure 2.1](image_url)

Figure 2.1: The research fields of individuals’ physical fitness, health complaints, physical demands, and military performance are interrelated and are relevant for the description of physical job requirements for different military occupational specialties.

It is established that a mismatch between physical fitness and physical job requirements can increase the risk of injury, jeopardize unit military performance, and decrease overall morale in a military setting (Gledhill and Jannik, 1992; Rayson, 1998; Rosendal et al.,
Adequately described job requirements for occupational specialties are especially important in the compulsory Swiss Army. In contrast to many other military organizations (National Research Council, 2006; NATO, 1997; Stevenson et al., 1992), the Swiss Army uses job-specific minimum physical fitness standards to assign the conscripts to a suitable job. However, today’s Swiss Army job-specific minimum fitness standards only contain references to the total score achieved in the fitness-test battery and are based solely on an expert’s appraisal. However, for more sophisticated, job-specific minimum fitness standards, an explicit description of job requirements is needed first. Therefore, the relationship between the parameters shown in Figure 2.1 should be investigated. The present thesis focused on the adoption and development as well as validation of methods to objectively investigate the respective fields in the future.

**Physical demands on soldiers**

The following activities were identified as the key common tasks performed in recent and current armed forces missions: walking, marching with a backpack, running, and physically demanding materials-handling including lifting and lowering loads, lifting and carrying loads, and digging (Jaenen, 2009). Eighty nine percent of physically demanding tasks in British Army occupations involved either lifting or carrying loads (Rayson, 1998). To the author’s knowledge, no objective data on duration, frequency, or intensity of these activities during daily military routines has been published so far. Most often, the total daily energy expenditure is assessed to quantify soldiers’ physical demands. Total energy expenditure (TEE) in armed forces ranges from moderate to high values during daily military routine (14.1 MJ/d in US Army support soldiers to 17.2 MJ/d in US special forces; Tharion et al., 2004) to very high values during specific field training exercises (25.7 MJ/d in US marines; Castellani et al., 2006). In their meta-analysis, Tharion et al. (2005) summarized TEE values of 16.8 to 17.1 MJ/d in long-duration boot camps (approximately 60 days). Similar data were found in training schools of Australian infantry.
and British parachute recruits (17 and 18.3 MJ/d, respectively; Forbes-Ewan, 2004; Wilkinson et al., 2008).

Published procedures to assess physical demands in military or fire fighting occupational specialties include self-report questionnaires, interviews, observations, and physical measurements (Chahal et al., 1992; Gledhill and Jamnik, 1992; Rayson, 1998; Stevenson et al., 1992). To assess the duration and frequency of physical activities, direct observation or video observation are most precise, but these methods are impractical for large groups of participants. Using self-report questionnaires is the most practical approach in large-scale studies, but their reliability, validity, and objectivity is low (LaPorte et al., 1985; Lester et al., 2006). The doubly labeled water (DLW) method for energy expenditure estimation is expensive and yields no information on activities characteristics or on its progression across the day. Still, the ideal method to objectively assess physical job requirements in a military setting is not yet available (Almeida et al., 1999; Bos et al., 2002). The most promising method to assess the duration, frequency, and intensity of physical activities in large groups of participants, as recommended by Bos et al. (2002), is the use of body-fixed sensors. Several approaches have shown to be effective in recognizing specific activities and estimating energy expenditure based on data of diverse body-fixed sensors (Aminian et al., 1999; Brage et al., 2004; Pärkkä et al., 2004; Pober et al., 2006). However, none of these approaches has been adapted for the application in a military setting.

Therefore, in Chapters 6 to 8 of the present thesis, objective methods were adapted to the military setting to assess activity types, energy expenditure, and distance covered on foot based on body-fixed sensors. In Chapter 9, the methods were applied to quantify physical activities and demands in five different occupational specialties of the Swiss Army.

Physical fitness tests

In terms of measuring their employees’ physical capabilities, all military organizations participating in the Research Technical Group of the North Atlantic Treaty Organization (NATO) test their soldiers’s physical fitness at least once a year (NATO, 1997; Williamson et al., 2009). The Swiss Army as well has a long tradition of testing the physical capabilities of its personnel. Over time, the physical performance test has been modified
frequently, but it is still an important part of the recruitment procedure and soldiers’ performance evaluations during military service.

The fitness-test batteries of the armed forces of different nations described in NATO (1997) all contain an assessment of aerobic endurance capacity, namely a distance run of 1.5 to 2.0 miles in the United Kingdom (UK), the United States of America (USA), Czech Republic (CZ), Austria (AT), and Georgia (GE); a 20-meter shuttle run in Canada (CA); and a 12-minute running test in Finland (FI), the Netherlands (NL), and Germany (DE). Most of those fitness-test batteries further contain sit-ups to assess abdominal muscle fitness (UK, CA, FI, NL, USA, CZ, and DE). Additionally, diverse tests for muscle power and agility are conducted, namely push-ups (GB, CA, AT, FI, NL, USA, CZ, GE, and DE), hand grip tests (CA and FI), standing long jumps (DE), timed squats (FI), and 4x10m shuttle run tests (DE and CZ). For further details, see NATO (1997).

In Chapter 3 of this thesis, physical fitness profiles in Swiss conscripts during the years 1982 to 2005 are presented. In Chapter 4, the development of a new physical performance test to assess Swiss soldiers’ fitness is described. In Chapter 5, its ability to predict injury risk is quantified and discussed.

**Injuries in soldiers**

The US Armed Forces Epidemiological Board (AFEB) Injury Prevention and Control Work Group concluded in its report that injuries impose a greater negative impact on the health and readiness of the US Armed Forces than any other category of medical complaint (Jones and Hansen, 2000). Today, non-battle injury rates are the major health problem of the armed forces (Peake, 2000). Meta-analyses have quantified non-battle injury rates to range from 10 to 15 per month per 100 male recruits (Kaufman et al., 2000; National Research Council, 2006). In basic training of different armed forces, it has been shown that 58 to 78% of injuries concern the lower extremities (Almeida et al., 1999; Heir and Glomsaker, 1996; Knapik et al., 1993). The knee and foot are the anatomical sites most often affected, followed by the ankle and the back (Almeida et al., 1999; Jones et al., 1993; Knapik et al., 1993; Popovich et al., 2000).
Injuries are important in terms of lost time at work or training, recruit attrition, decreased job readiness and morale, reduced operational performance of the entire troop, as well as high medical costs (Almeida et al., 1999; Bilzon et al., 2005; Heir and Glomsaker, 1996; Kaufman et al., 2000; National Research Council, 2006; Ruscio et al., 2010; Snoddy and Henderson, 1994). In military populations, a number of risk factors for injuries have been identified, such as low level of past physical activity, low level of physical fitness, previous injury history, high running mileage, high amount of weekly exercise, smoking, age, gender, and psychosocial demands (Almeida et al., 1999; Blacker et al., 2008; Gilchrist et al., 2000; Jones et al., 1993; Jones and Knapik, 1999; Kaufman et al., 2000; Knapik et al., 1993; National Research Council, 2006). Among these, low level of physical fitness, especially low level of aerobic endurance, is one of the most relevant risk factors for injuries in a military population (Blacker et al., 2008; Craig et al., 1998; Jones et al., 1993; Jones and Knapik, 1999; Knapik et al., 1993; Lee et al., 1997; Rosendal et al., 2003; Snoddy and Henderson, 1994). Several approaches have aimed to lower these risk factors and have been effective in reducing injuries and attrition. Modified training programs in terms of reduction of running mileage (Jones and Knapik, 1999; Jones et al., 2000; Knapik et al., 2004; Rudzki, 1997; Rudzki and Cunningham, 1999) and progressive load in physical readiness training (Knapik et al., 2009), preconditioning and individualized fitness training (Knapik et al., 2006; Lee et al., 1997; McGuine and Keene, 2006), special training units for individuals with low physical fitness (Knapik et al., 2004), and multiple interventions (Kelly and Bradway, 1997; Knapik et al., 2004; Stacy and Hungerford, 1984) have been reported as successful approaches.

In Chapter 5 of the present thesis, injuries in four different Swiss Army occupational specialties are described. Further, those injuries are related to the individual’s physical fitness by the use of receiver operating curve analysis to detect the fitness parameters with discriminative power to predict overuse injuries in each military occupational specialty.
Main aims

The main aims of the present thesis are as follows:

- to assemble a reliable and feasible fitness-test battery to assess physical capabilities in larger groups of young men.
- to demonstrate that the Swiss Army physical fitness-test battery is a valid instrument to predict risk of injuries in physically demanding military occupational specialties.
- to develop and apply a new military-specific method to objectively assess physical demands in terms of activity types, intensity, duration, and frequency.

Ethics

All participants in the presented studies received comprehensive oral and written information on the study, and they provided informed consent for their participation as approved by the Swiss Army Sports and Preventions Competence Centre and the Cantonal Ethics Committee of Bern, Switzerland (approved 28.09.2007, registry number 195/07). Volunteers’ data were always made anonymous after data collection by the use of a random reference number for every subject. Therefore, a table with names of the volunteers and respective reference numbers was deposited at the Swiss Army Medical Service Department, Ittigen, while only anonymous sensor, fitness, and injury data were used for analyses at the Swiss Federal Institute for Sports, Magglingen.
3. Physical performance in young men at Swiss Army recruitment 1982 to 2005

Published in Schweiz Z Sportmed Sporttraumat, 2009, 57, 75-77.

Introduction

Obesity in young people has increased in Switzerland and other western countries in the last 10 to 30 years (Blair and Church, 2004; Faeh et al., 2008; Santtila et al., 2006). Actually over 50% of young Swiss citizens are either totally inactive or not active enough, based on the minimal recommendation of at least half an hour of moderate physical activities per day (Lamprecht and Stamm, 2006). However, little is known about the changes in physical fitness in young Swiss citizens. In young Norwegian, Danish, Swedish and Finnish men, a trend of decreased aerobic endurance performance in the last 20 years was observed (Rasmussen et al., 1999; Santtila et al., 2006; Sharp et al., 2002; Sorensen et al., 1997). In young Finnish men also a decrease in muscle fitness in the last 15 years was observed (Santtila et al., 2006).

In the year 1905 a physical performance test was implemented at the recruitment of the Swiss Army. Over time the performance test has been modified frequently. Nevertheless between 1982 and 2005 the 12-minutes running test (12-MRT) and the standing long jump (SLJ) were continuously part of the physical performance test battery. The five-meter pole climbing (PC) was included until 2002 only. It was the aim of this study to describe changes in physical performances in young Swiss men over a time of 21 to 24 years, based on 12-MRT, SLJ and PC results.
Methods

Every year, all 19 year old male Swiss citizens take part in the compulsory Swiss Army recruitment. Physical performance data used in this study were assessed during the years 1982 to 2005. Until 1987, the Swiss Federal Statistical Office published the aggregated results of only every fifth year. From 1988 on, the Swiss Federal Institute of Sport assumed this task and published the results annually. In those aggregated data sets the numbers of exempted and participating conscripts as well as the mean results of the participating conscripts were given.

The numbers of the exempted and participating conscripts were presented as means and standard deviations. For analysis, the relative attendance in the three disciplines 12-MRT (also referred to as Cooper test; Cooper, 1968), SLJ (Bosco et al., 1983) and PC were calculated and presented with the respective mean performances. For PC no descriptions were published in a scientific journal. However, this part of the performance test was strictly standardized as well. Only descriptive statistics were used in the present study.

Results

The mean distance in the 12-MRT decreased during the years 1987-2002 constantly from 2601 to 2495 m (-4.1%). During 1982-1987 and after 2002 the 12-minutes running distance was constant or even tended to increase (Figure 3.1). The results in the SLJ did not change over the 24 years. After a small increase in distance during 1982-1990 from 2.38 to 2.43 m (+2.1%) the results were stable until 2000 when they started to decrease slightly from 2.43 to 2.38 m (-2.1% in 5 years; 3.2). Time in PC increased during 1982-2002 constantly from 4.80 to 5.75 s (+19.8%). Every year 33'140 ± 4'252 conscripts attended the recruitment. Therefrom 4'482 ± 1'150 conscripts were excluded from the 12-minutes running test, 4'491 ± 1'593 from the standing long jump and 3'524 ± 1'742 from pole climbing due to medical reasons. During 1900-1993 and 2003-2004 the percentage of conscripts participating in the performance tests decreased noticeably from 92% to 83% respectively from 82% to 76% (Figures 3.1 and 3.2, grey columns).
Figure 3.1: Performance and participation in 12-minutes running test (Cooper test) at the compulsory Swiss Army recruitment during the years 1982-2005. Dots: Distance of the Cooper test; bars: Participation.

Figure 3.2: Performance and participation in standing long jump at the compulsory Swiss Army recruitment during the years 1982-2005. Dots: Distance of the standing long jump; bars: Participation.

Discussion

The indicated decrease in 12-MRT performances could be the result of lower physical activity levels among Swiss adolescents. However, questionnaire data between 1992 and 2002 do not show a relevant decrease in physical activity in Swiss citizens (Lamprecht and Stamm, 2006). Other possible reasons for the decline of 12-MRT and for climbing performances might be the continuous gain in body mass (Faeh et al., 2008) and loss of specific practice at school. The effect of the changing motivation to perform an endurance test in a military setting cannot be determined with the existing data. It is likely that with changing acceptance of the Army in the population motivation for maximal performances at the recruitment has changed as well.

In Switzerland and Finland changes in 12-MRT values are similar (Santtila et al., 2006). However, in 1982 the aerobic performance of young Finnish men was considerably higher than the one of Swiss men. Therefore the decrease in 12-MRT performance in young Swiss men between 1987 and 2002 was less dramatic than in young Finnish men. In Switzerland a reversal of the trend may have occured after 2002 from decreasing to by trend increasing 12-MRT performances. This change of trend is not visible in the data of young Finnish men. The indicated increase in 12-MRT performances after 2002 cannot be explained with available data. Only the decrease in participation after 2003 may have influenced the mean performance.

While in the last 15 years a decrease in muscle fitness was observed in young Finnish men (Santtila et al., 2006), in young Swiss men this trend was not confirmed. SLJ performances were stable until 2000 and decreased slightly thereafter. Physical activity level and motivation had probably a lower impact on SLJ than on the 12-MRT performances.

To a certain extent, muscle fitness was determined by the PC test too. However PC performance is strongly related to movement coordination and requires experiences in climbing technique. Therefore it remains unclear if the indicated decrease of climbing performance was related to decreasing muscle fitness or to limited climbing technique. After 2002 the PC was eliminated from the performance test, because climbing poles had been taken down in numerous Swiss schools and the possibilities to instruct male adolescents in the climbing technique were more and more limited.

After changes in recruiting procedures the rate of conscripts exempted from performance tests, due to medical reasons, rose in the years 1993 and 2004. The decrease in participation probably influenced the mean physical performances, based on a selection
effect. However, performance values were still comparable with previous measurement
data because more than 75% of all 19 years old Swiss men took part in these assessments.

For more sophisticated analyses individual performance data, not only aggregated data,
and information about conscripts' motivation would be necessary. Unfortunately these
data were not available any more.

After 2005 a new physical performance test battery was established at the recruitment of
the Swiss Army. Therefore it will be difficult to continue the monitoring of 12-MRT
performances in young Swiss men. However the new endurance test (a progressive
endurance run) may be more appropriate especially for individuals with low fitness levels
or no experience in self-pacing compared to 12-MRT (see Chapter 4). The new test battery
measures aerobic endurance (progressive endurance run), explosive muscle power (seated
2-kg-shot put and standing long jump), trunk muscle fitness (trunk muscle strength test)
and balance (one-leg standing test). Additionally daily physical activity level, body weight
and height of all conscripts are assessed.

Conclusion

Aerobic endurance performances and PC performances indicated a decrease since 1987
until 2002, while SLJ performances were stable. The inconsistent data do not prove the
common opinion of a decreasing general physical fitness level among young Swiss men
over the last decades. However, for future monitoring activities individual performance
data are needed to better understand the physical fitness profiles among young Swiss men.
4. Assembling and verification of a fitness test battery for the recruitment of the Swiss Army and nation-wide use

Published in Schweiz Z Sportmed Sporttraumat, 2007, 55, 126-31.

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 1905 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 standardized 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 the Swiss Army and 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 \( \text{VO}_{2\text{max}} \) (Safrit et al., 1988), although it is not immune to inaccuracy (Shepard, 1984). In larger population groups, direct measurement of \( \text{VO}_2 \) 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-minutes run test (12-MRT), also referred to as Cooper test, is strongly related to the criterion measure of \( \text{VO}_{2\text{max}} \) in adults \( (r = 0.84 - 0.92; \text{Cooper, 1968; Grant et al., 1995; McCutcheon et al., 1990}) \). 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 \( \text{VO}_{2\text{max}} \) in adults \( (r = 0.79 - 0.90; \text{Cooper, 1968; Grant et al., 1995; Leger and Gadoury, 1989; McNaughton et al., 1998; Ramsbottom et al., 1988}) \). 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 \( (r = 0.80; \text{Cronin and Owen, 2004}) \). 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 rsp. 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 4.1) 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 one-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.
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 nine motor fitness tests (cardiorespiratory endurance, muscular strength, endurance and speed, flexibility and balance) and five 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 four parts (Figure 4.2). 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.

Figure 4.2: Schematic of the study design. SLJ: Standing long jump, PS: pendulum sprint, SSP: seated shot put, TMS: trunk muscle strength test, OLS: one-leg standing, WB: walking on a beam, PER: progressive endurance run.

Subjects

All subjects were recruits of the logistic corps of the Swiss Army. The recruits in the logistic corps represent subjects with a wide range of physical fitness levels. They are representative for Swiss men at the age of 20 years, therefore they were chosen to
participate in this study. 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. Sixty 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 15794 conscripts who had to pass their recruitment for the Swiss Army. The entire fitness-test battery was completed by 81.4% of the conscripts (n = 12862, 19.9 ± 1.0 y, 72.8 ± 12.0 kg, 178.3 ± 15.9 cm), 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.

Eleven 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 three 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 three 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 two trials was valued with an accuracy of 0.1 s.

**Trunk muscle strength**

In the trunk muscle strength test (TMS), 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 standardized and controlled with a laterally opened box (Figure 4.1). 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 one minute, 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_{2peak}-test were conducted. Previously, all subjects had performed the 12-MRT at least twice. Eighteen 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_{2peak} on a treadmill. These additional assessments were conducted during the following three 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 minute. 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_{2peak}. A linear regression between the peak running velocity of PER and VO_{2peak} of the treadmill test was calculated to estimate the VO_{2peak} 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 one second 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 three 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 three 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 three trials was valued with an accuracy of 0.001 s.

Trunk muscle strength
TMS 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. Twenty-five 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 calculated adding the time of left and right OLS and then subtracting twice the time for 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 4.1). Additionally, PER time correlated positively with the distance run in the 12-MRT and with the VO$_{2peak}$ in the maximal treadmill exercise test (Table 4.2).

VO$_{2peak}$ may be estimated by the peak running velocity of the PER with the following regression ($p = 0.000$):

$$\text{VO}_{2peak} \left[ \text{ml kg}^{-1} \text{min}^{-1} \right] = 2.309 \cdot \text{Velocity-peak} \left[ \text{km h}^{-1} \right] + 16,549.$$

Muscle power

SSP yielded a good reliability (Table 4.1). 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 4.2).

SLJ was highly reproducible (Table 4.1). SLJ correlated positively with relative maximal power during normalised jumps. PS was inversely correlated with SLJ (Table 4.2). While the performance of SLJ did not differ between both trials, the performance in PS was better in the second trial (Table 4.1).

Trunk muscle strength

TMS yielded a good reliability (Table 4.1). The simplified TMS and the original one correlated positively (Table 4.2).

Balance

The repeated combination of static and dynamic balance tests generated a moderate correlation coefficient (Table 4.1). The results in both cycles were not different. The retest of OLS alone yielded a moderate reliability (Table 4.1). OLS results were better in the repeated test (Table 4.1).
Table 4.1: Reliability of the physical performance tests (number of subjects, means, standard deviations, differences and correlation coefficients).

<table>
<thead>
<tr>
<th>n</th>
<th>performance test 1</th>
<th>vs.</th>
<th>performance test 2</th>
<th>difference</th>
<th>correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>68</td>
<td>progressive endurance run (PER) 1</td>
<td>progressive endurance run (PER) 2</td>
<td>0.11 km/h</td>
<td>r = 0.89 (p = 0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.08 ± 2.02 km/h</td>
<td>14.19 ± 2.08 km/h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>standing long jump (SLJ) 1</td>
<td>standing long jump (SLJ) 2</td>
<td>0.00 m</td>
<td>r = 0.90 (p = 0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.26 ± 0.22 m</td>
<td>2.26 ± 0.23 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>pendulum sprint (PS) 1</td>
<td>pendulum sprint (PS) 2</td>
<td>-0.19 s</td>
<td>r = 0.84 (p = 0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.99 ± 0.72 s</td>
<td>10.80 ± 0.78 s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>seated shot put (SSP) 1</td>
<td>seated shot put (SSP) 2</td>
<td>0.02 m</td>
<td>r = 0.83 (p = 0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.58 ± 0.65 m</td>
<td>6.60 ± 0.67 m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>trunk muscle strength test (TMS) 1</td>
<td>trunk muscle strength test (TMS) 2</td>
<td>-3 s</td>
<td>r = 0.77 (p = 0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>01:35 ± 00:45 [min:sec]</td>
<td>01:32 ± 00:49 [min:sec]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>one-leg standing (OLS) 1</td>
<td>one-leg standing (OLS) 2</td>
<td>4.35 s</td>
<td>r = 0.50 (p = 0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>38.88 ± 8.14 s</td>
<td>43.23 ± 12.08 s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>static and dynamic balance 1</td>
<td>static and dynamic balance 2</td>
<td>1.56 s</td>
<td>r = 0.57 (p = 0.000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>27.78 ± 9.44 s</td>
<td>29.35 ± 14.03 s</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Feasibility and standard values

Based on the results of the reliability- and validity part of this study PER, SLJ, SSP, TMS and bipart-balance test (OLS and WB) were selected for SPFTB. Three 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, ten sport experts (40%) described WB as potentially dangerous or not well standardized. The representative SPFTB standard values (n=12’862) for young men are presented in Table 4.3.
Table 4.2: Validity of the physical performance tests (number of subjects, means, standard deviations and correlation coefficients).

<table>
<thead>
<tr>
<th>n</th>
<th>performance test 1</th>
<th>vs.</th>
<th>performance test 2</th>
<th>correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>progressive endurance run (PER)</td>
<td>12-min run test (12-MRT)</td>
<td></td>
<td>r = 0.91 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>13.60 ± 2.03 km/h</td>
<td>2303.45 ± 414.95 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>progressive endurance run (PER)</td>
<td>VO₂peak</td>
<td></td>
<td>r = 0.84 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>14.36 ± 2.26 km/h</td>
<td>49.93 ± 6.01 ml/min/kg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>standing long jump (SLJ)</td>
<td>jump on force plate</td>
<td></td>
<td>r = 0.64 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>2.14 ± 0.24 m</td>
<td>SJ: 44.08 ± 6.03 W/kg</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CMJ: 46.35 ± 6.98 W/kg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>standing long jump (SLJ)</td>
<td>sprint 30-40m</td>
<td></td>
<td>r = -0.73 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>2.14 ± 0.24 m</td>
<td>1.33 ± 0.14 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>standing long jump (SLJ)</td>
<td>pendulum sprint (PS)</td>
<td></td>
<td>r = -0.73 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>2.14 ± 0.24 m</td>
<td>10.98 ± 0.81 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>pendulum sprint (PS)</td>
<td>sprint 30-40m</td>
<td></td>
<td>r = 0.85 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>10.98 ± 0.81 s</td>
<td>1.33 ± 0.14 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>seated shot put (SSP)</td>
<td>bench press power</td>
<td></td>
<td>r = 0.54 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>6.10 ± 0.69 m</td>
<td>15kg: 369.06 ± 80.67 W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>6.12 ± 0.73 m</td>
<td>30kg: 362.83 ± 108.17 W</td>
<td></td>
<td>r = 0.65 (p = 0.000)</td>
</tr>
<tr>
<td>60</td>
<td>trunk muscle strength test (TMS)</td>
<td>trunk muscle strength test SOMC</td>
<td></td>
<td>r = 0.85 (p = 0.000)</td>
</tr>
<tr>
<td></td>
<td>01:19 ± 00:54 [min:sec]</td>
<td>01:22 ± 00:59 [min:sec]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: SOMC = Swiss Olympic Medical Centre
Table 4.3: 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).

<table>
<thead>
<tr>
<th>percentile</th>
<th>BMI [kg/m²]</th>
<th>PER [min:s]</th>
<th>pdt VO₂peak [ml kg⁻¹ min⁻¹]</th>
<th>SSP [m]</th>
<th>SLJ [m]</th>
<th>TMS [s]</th>
<th>OLS (t+l+t) [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>18.61</td>
<td>06:21</td>
<td>39.94</td>
<td>5.30</td>
<td>1.93</td>
<td>41.00</td>
<td>30.00</td>
</tr>
<tr>
<td>10</td>
<td>19.32</td>
<td>07:58</td>
<td>42.12</td>
<td>5.55</td>
<td>2.03</td>
<td>54.00</td>
<td>32.50</td>
</tr>
<tr>
<td>15</td>
<td>19.84</td>
<td>09:10</td>
<td>43.74</td>
<td>5.70</td>
<td>2.10</td>
<td>63.00</td>
<td>35.00</td>
</tr>
<tr>
<td>20</td>
<td>20.28</td>
<td>09:45</td>
<td>44.53</td>
<td>5.83</td>
<td>2.15</td>
<td>71.00</td>
<td>37.00</td>
</tr>
<tr>
<td>25</td>
<td>20.66</td>
<td>10:27</td>
<td>45.47</td>
<td>5.95</td>
<td>2.20</td>
<td>79.00</td>
<td>38.30</td>
</tr>
<tr>
<td>30</td>
<td>20.98</td>
<td>11:13</td>
<td>46.51</td>
<td>6.05</td>
<td>2.23</td>
<td>87.00</td>
<td>40.00</td>
</tr>
<tr>
<td>35</td>
<td>21.31</td>
<td>11:30</td>
<td>46.86</td>
<td>6.15</td>
<td>2.26</td>
<td>96.00</td>
<td>41.00</td>
</tr>
<tr>
<td>40</td>
<td>21.67</td>
<td>12:15</td>
<td>47.90</td>
<td>6.25</td>
<td>2.30</td>
<td>105.00</td>
<td>42.40</td>
</tr>
<tr>
<td>45</td>
<td>22.01</td>
<td>12:35</td>
<td>48.35</td>
<td>6.35</td>
<td>2.32</td>
<td>108.00</td>
<td>43.80</td>
</tr>
<tr>
<td>50</td>
<td>22.34</td>
<td>13:07</td>
<td>49.07</td>
<td>6.45</td>
<td>2.35</td>
<td>113.00</td>
<td>45.00</td>
</tr>
<tr>
<td>55</td>
<td>22.71</td>
<td>13:20</td>
<td>49.36</td>
<td>6.52</td>
<td>2.39</td>
<td>119.00</td>
<td>46.30</td>
</tr>
<tr>
<td>60</td>
<td>23.06</td>
<td>13:56</td>
<td>50.17</td>
<td>6.60</td>
<td>2.40</td>
<td>125.00</td>
<td>47.70</td>
</tr>
<tr>
<td>65</td>
<td>23.46</td>
<td>14:15</td>
<td>50.60</td>
<td>6.70</td>
<td>2.43</td>
<td>132.00</td>
<td>49.00</td>
</tr>
<tr>
<td>70</td>
<td>23.94</td>
<td>14:44</td>
<td>51.25</td>
<td>6.80</td>
<td>2.45</td>
<td>140.00</td>
<td>50.00</td>
</tr>
<tr>
<td>75</td>
<td>24.49</td>
<td>15:10</td>
<td>51.83</td>
<td>6.90</td>
<td>2.49</td>
<td>151.00</td>
<td>52.00</td>
</tr>
<tr>
<td>80</td>
<td>25.14</td>
<td>15:31</td>
<td>52.31</td>
<td>7.00</td>
<td>2.52</td>
<td>163.00</td>
<td>54.00</td>
</tr>
<tr>
<td>85</td>
<td>25.93</td>
<td>16:16</td>
<td>53.32</td>
<td>7.17</td>
<td>2.56</td>
<td>182.00</td>
<td>56.60</td>
</tr>
<tr>
<td>90</td>
<td>27.12</td>
<td>17:00</td>
<td>54.31</td>
<td>7.38</td>
<td>2.60</td>
<td>202.00</td>
<td>61.00</td>
</tr>
<tr>
<td>95</td>
<td>29.37</td>
<td>17:46</td>
<td>55.34</td>
<td>7.70</td>
<td>2.68</td>
<td>240.00</td>
<td>70.00</td>
</tr>
</tbody>
</table>

Discussion

Physical performance tests

Cardiorespiratory endurance
Peak treadmill running velocity, during a speed-incremented VO$_{2peak}$ 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 and 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 4.2). 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 (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

TMS 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, TMS was compared with the established but more expensive TMS 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 strengths

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.

Two 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 two 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 three 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-years old men.

**Conclusion**

The new health- and performance-related SPFTB is qualified for nation-wide use. SPFTB meets the previously specified demands; a) three 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 five 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).
5. **Physical fitness predicts risk of overuse injuries among Swiss Army recruits**

**Introduction**

The awareness of the impact of injuries has grown in military organizations in the last two decades. In November 1996, the US Armed Forces Epidemiological Board (AFEB) Injury Prevention and Control Work Group concluded in their report, that injuries impose a greater negative impact on the health and readiness of US Armed Forces than any other category of medical complaint (Jones and Hansen, 2000). Today non-battle injury rates are the major health problem of armed forces (Peake, 2000). Meta analyses quantified non-battle injury rates to range from 10 to 15 per month per 100 male recruits (Kaufman et al., 2000; National Research Council, 2006). These injuries are important in terms of loss of time from work or training, recruit attrition, decreased job readiness and morale, reduced operational performance of the entire troop as well as high medical costs (Almeida et al., 1999; Bilzon et al., 2005; Heir and Glomsaker, 1996; Kaufman et al., 2000; National Research Council, 2006; Ruscio et al., 2010; Snoddy and Henderson, 1994). In military populations a number of risk factors for injuries have been identified, such as low level of past physical activity, low level of physical fitness, previous injury history, high running mileage, high amount of weekly exercise, smoking, age, gender and psychosocial demands (Almeida et al., 1999; Blacker et al., 2008; Gilchrist et al., 2000; Jones et al., 1993; Jones and Knapik, 1999; Kaufman et al., 2000; Knapik et al., 1993; National Research Council, 2006). Among these, low level of physical fitness (Jones et al., 1993; Jones and Knapik, 1999; Lee et al., 1997; Rosendal et al., 2003), especially of aerobic endurance (Blacker et al., 2008; Craig et al., 1998; Jones et al., 1993; Knapik et al., 1993; Snoddy and Henderson, 1994), is one of the most relevant risk factors for injuries in a military population.

To minimize physical overload and to prevent health complaints as well as loss of time from duty, it is crucial to find an optimal balance between job requirements and individual physical capacity. In terms of measuring their employees’ physical capabilities, all military
organizations participating in the Research Technical Group of the North Atlantic Treaty Organization (NATO) test their employees’ physical fitness at least once a year (NATO, 1997; Williamson et al., 2009). All those fitness-test batteries contain a running test to assess aerobic endurance capacity. Most of the test batteries contain sit-ups to assess abdominal muscle fitness and some contain further diverse tests to assess muscle power, like push-ups, hand grip test, standing long jump and others (NATO, 1997). The Swiss Army fitness-test battery consists of similar fitness parameters and includes additionally a test for balance. In contrast to many other military organizations (National Research Council, 2006; NATO, 1997; Stevenson et al., 1992), the compulsory Swiss Army uses job-specific minimum physical fitness standards to assign the conscripts to their jobs in the military service. However, today’s Swiss Army job-specific minimum fitness standards contain only references to the total score achieved of the fitness test battery and are based on subjective expert appraisal. The use of more reliable job-specific minimum physical fitness standards upon recruitment that are developed based on objective data may reduce injuries in military boot camps. For reliable job-specific minimum fitness standards the relationship between fitness parameters and job specific task performances as well as injury occurrence should be investigated. The relationship between fitness parameters and task performances has been studied before (Bernauer and Bonanno, 1975; Chahal et al., 1992; Popper et al., 1999; Rayson, 1998; Stevenson et al., 1992).

The present study determined the occurrence of injuries in four different Swiss Army boot camps and investigated the discriminative power of physical fitness tests for predicting the risk of injuries. It was the aim of the present study to demonstrate validity of different fitness tests as injury-prevention tool in selection of personal for physically demanding military occupational specialties.
Methods

Study design and participants

In the present study the physical fitness performance of volunteers prior to military service and their injuries during 18 weeks of boot camp were assessed. The four training schools investigated were chosen by the criteria of being physically demanding based on experts' appraisal and involving more than 400 recruits a year. All recruits in these boot camps (rescue technicians, armored infantry, fusilier infantry and reconnaissance infantry school) were asked to take part in the study (Table 5.1). The participants received comprehensive oral and written information and they provided written informed consent for their participation as approved by the Cantonal Ethics Committee of Bern, Switzerland and the Swiss Army Sports and Preventions Competence Center.

Data collection

Injuries

Injuries were defined as and registered if a subject who sustained physical damage to his body visited the medical care center for this reason. Injuries were continuously recorded on the individuals’ medical records by the medical staff. The data collected included calendar date, anatomical site, diagnosis, severity and whether the symptoms were of acute or overuse onset. Acute injuries were defined as those happened by a sudden traumatic event. Overuse injuries were defined as those that were associated with repetitive physical activities. Anatomical sites were categorized as head, shoulder, back, knee, foot, ankle, Achilles tendon, and others. Categories for diagnosis were inflammation or musculoskeletal pain, sprain, contusion, strain and others. Severity was categorized either as trivial (no consequences), as low (limited duty up to one week), as moderate (one or more full training-days lost or limited duty for more than one week), or as severe (instant discharge from service or permanent physical damage).
Anthropometry and physical fitness

Body height was measured to the nearest 0.1 cm using a stadiometer (Seca model 214, Seca GmbH, Hamburg, Germany) and body weight was measured to the nearest 0.1 kg on a calibrated digital balance (Seca model 877, Seca GmbH, Hamburg, Germany). The fitness-test battery contained a progressive endurance run (PER) to measure aerobic endurance capacity, a trunk muscle strength test (TMS) to measure trunk muscle fitness, a standing long jump (SLJ) and a seated shot put (SSP) to measure the muscle power of the lower and upper extremities, respectively, and a one-leg standing test (OLS) to measure balance ability. The PER is a paced running test, conducted according to the protocol developed by Conconi et al. (1982), evaluated using the final running velocity. In the TMS the subject had to hold an isometric body position (on forearms and feet with upper body and legs in a straight line) for as long as possible while lifting their feet alternately. The SLJ was performed from the gym hall floor onto a mat of 7 cm height. The SSP was performed as a 2-kg-ball chest pass while sitting upright on a bench with the back in contact with a solid wall. In OLS participants had to close their eyes after 10 s and they had to lay their head back after 20 s in position. Time was measured for the left and right leg separately and the sum of both was evaluated for balance. Precise descriptions of the five tests were published elsewhere (Wyss et al., 2007; Chapter 4).

Statistical Analysis

Injury risk (incidence proportion) is calculated as the number of recruits with one or more injuries during 18 weeks of military service divided by the total number of assessed recruits. The injury incidence rate is expressed as the total number of injuries per month per 100 recruits.

Statistical analysis was performed with SPSS for Windows (version 16.0, SPSS Inc., Chicago, IL) with an alpha level of 0.05 to indicate statistical significance. Descriptive statistics on fitness and injury data were produced for every military training school. To compare anthropometric data and the fitness performances of the recruits in the four different military training schools, a one-way analysis of variance (ANOVA) was conducted. A Tukey post-hoc analysis was used to evaluate pair wise differences among training schools. To compare injury risk between recruits above and below minimum fitness standards and between different military training schools, a chi-square distribution
was performed. Bonferroni adjustments were applied to correct the significance for multiple test comparisons among the four study groups.

The discriminative power of the physical fitness tests for predicting the risk of overuse injuries was tested with receiver operating characteristic (ROC) analysis (Zou et al., 2007). If the area under the ROC-curve was of significant size, the cut-off point on the curve with maximal sensitivity and specificity was used as exemplary injury-related minimum physical fitness standard (IMFIS) in this study (see example in Figure 5.1). Further, the overuse injury incidences of the group of recruits who exceeded the IMFIS and of the group of recruits who did not exceed the specific cut off value were calculated. Finally, the relative risk (RR) and respective 95% confidence interval (95%-CI) for overuse injuries in the group of recruits below the IMFIS was calculated.

Results

Descriptive fitness and injury data

A total of 473 recruits in the four selected military training schools were asked to volunteer in the present study. Five recruits did not participate and medical records of nine recruits were not available after their service. Finally, complete physical fitness and injury data of 459 volunteers were recorded (Table 5.1).
Table 5.1: Mean and standard deviations of age, weight, height and body mass index (BMI) of investigated recruits.

<table>
<thead>
<tr>
<th>Swiss Army military training school</th>
<th>rescue technicians</th>
<th>armored infantry</th>
<th>fusilier infantry</th>
<th>reconnaissance infantry</th>
<th>total</th>
<th>ANOVA (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>131</td>
<td>145</td>
<td>107</td>
<td>76</td>
<td>459</td>
<td></td>
</tr>
<tr>
<td>age [y]</td>
<td>20.83 ± 1.39</td>
<td>20.71 ± 0.91</td>
<td>20.85 ± 1.10</td>
<td>20.76 ± 1.08</td>
<td>20.85 ± 1.12</td>
<td>0.684</td>
</tr>
<tr>
<td>weight [kg]</td>
<td>74.27 ± 12.22</td>
<td>72.38 ± 7.99c</td>
<td>76.42 ± 12.83bd</td>
<td>71.67 ± 9.46c</td>
<td>73.74 ± 10.56</td>
<td>0.007</td>
</tr>
<tr>
<td>height [cm]</td>
<td>176.99 ± 6.91</td>
<td>177.20 ± 6.30</td>
<td>178.90 ± 6.28</td>
<td>179.29 ± 5.64</td>
<td>177.88 ± 6.36</td>
<td>0.015</td>
</tr>
<tr>
<td>BMI</td>
<td>23.71 ± 3.23d</td>
<td>23.05 ± 2.09</td>
<td>23.87 ± 3.62d</td>
<td>22.29 ± 2.17ad</td>
<td>23.27 ± 2.91</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Significant differences (p < 0.05) in Tuckey post-hoc tests are symbolized with a (different from rescue technicians), b (different from armored infantry), c (different from fusilier infantry) and d (different from reconnaissance infantry).

The body mass index (BMI) of recruits in the reconnaissance infantry school was significantly lower than in the rescue technician (p = 0.004) and fusilier infantry school (p = 0.001, Table 5.1). The physical fitness performances of all recruits within the four military training schools are shown in Table 5.2 and compared with standard values assessed by Wyss et al. (2007). The SLJ performances did not differ between recruits from the armored infantry and reconnaissance infantry school only. The SSP performance of recruits in the rescue technician school was significantly lower than in the other schools. The OLS, TMS and PER performance of recruits in the armored infantry and reconnaissance infantry school were higher than in the rescue technician and fusilier infantry school (Table 5.2).
Table 5.2: Performances in physical fitness tests during recruitment for the Swiss Army.

<table>
<thead>
<tr>
<th>test</th>
<th>Swiss Army military training school</th>
<th>ANOVA (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rescue technicians</td>
<td></td>
</tr>
<tr>
<td>SLJ [m]</td>
<td>2.29 ± 0.22 b,c,d (35-40%)</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>2.50 ± 0.16 a,c (75-80%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.36 ± 0.23 b,d (50-55%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.50 ± 0.17 a,c (75-80%)</td>
<td></td>
</tr>
<tr>
<td>SSP [m]</td>
<td>6.35 ± 0.67 b,c,d (45%)</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>6.93 ± 0.64 a (75-80%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.72 ± 0.77 a,c (65-70%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.92 ± 0.65 a (75-80%)</td>
<td></td>
</tr>
<tr>
<td>TMS [s]</td>
<td>119.80 ± 50.34 b,d (55-60%)</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>176.48 ± 57.35 a,c (80-85%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>134.25 ± 58.43 b,d (65-70%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>176.83 ± 54.99 a,c (80-85%)</td>
<td></td>
</tr>
<tr>
<td>OLS [s]</td>
<td>45.86 ± 13.69 b,d (50-55%)</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>52.48 ± 12.58 a,c (75-80%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48.02 ± 11.80 b,d (60-65%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>54.29 ± 12.10 b,c (80-85%)</td>
<td></td>
</tr>
<tr>
<td>PER [s]</td>
<td>757.73 ± 205.24 b,d (45-50%)</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>926.89 ± 138.11 a,c (75-80%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>786.66 ± 207.36 b,d (50%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>949.05 ± 174.40 a,c (80-85%)</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Mean values and standard deviations within every assessed military training school and in parentheses respective percentile compared to standard values assessed with 12'862 conscripts by Wyss et al. (2007) are tabled. BMI: body mass index, SLJ: standing long jump, SSP: seated 2-kg shot put, TMS: trunk muscle strength test, OLS: one-leg standing test, PER: progressive endurance run.

**abcd** Significant (p < 0.05) differences in Tukey post-hoc tests are symbolized with a (different from rescue technicians), b (different from armored infantry), c (different from fusilier infantry) and d (different from reconnaissance infantry).

The total injury risk (41.2% - 56.6%) in the four assessed training schools did not differ significantly whereas the risk of overuse injury (Table 5.3) in the rescue technician school (22.9%) was significantly lower than in all other boot camps (37.9% - 43.4%). The majority of injuries were of overuse origin (64.5%) registered in 70% of injured recruits. The anatomical site most often affected was the knee (27%) followed by the back and other lower extremity sites. Fifty percent of injuries were diagnosed as inflammation or musculoskeletal pain, 20% were classified as sprains, 3% as strains, and 3% as contusions. Many injuries (78%) were trivial or of low severity causing limited duty for up to one week. Only 2.5% of all injuries were severe, causing immediate discharge from service or permanent physical damage (Table 5.3).
Table 5.3: Injury incidence proportion, injury incidence rate, anatomical site, diagnosis, severity and discharge in four assessed military training schools.

<table>
<thead>
<tr>
<th></th>
<th>rescue technicians</th>
<th>armored infantry</th>
<th>fusilier infantry</th>
<th>reconnaissance infantry</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>total number of subjects</td>
<td>131</td>
<td>145</td>
<td>107</td>
<td>76</td>
<td>459</td>
</tr>
<tr>
<td>injury incidence rate (per month/100)</td>
<td>12.2</td>
<td>18.2</td>
<td>18.2</td>
<td>18.8</td>
<td>16.7</td>
</tr>
<tr>
<td>injury incidence proportion (%)</td>
<td>41.2</td>
<td>56.6</td>
<td>52.3</td>
<td>53.9</td>
<td>50.8</td>
</tr>
<tr>
<td>overuse injury incidence proportion (%)</td>
<td>22.9**</td>
<td>37.9</td>
<td>40.2</td>
<td>43.4</td>
<td>35.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>anatomical site [number (% of total injuries)]</th>
<th>knee</th>
<th>back</th>
<th>Achilles tendon</th>
<th>foot</th>
<th>ankle</th>
<th>shoulder</th>
<th>head</th>
<th>other lower extremity sites</th>
<th>others</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>15 (22.4)</td>
<td>5 (7.5)</td>
<td>6 (9.0)</td>
<td>7 (10.4)</td>
<td>3 (4.5)</td>
<td>2 (3.0)</td>
<td>2 (3.0)</td>
<td>7 (10.4)</td>
<td>20 (29.9)</td>
<td>86 (26.9)</td>
</tr>
<tr>
<td>knee</td>
<td>33 (29.7)</td>
<td>10 (9.0)</td>
<td>10 (9.0)</td>
<td>12 (10.8)</td>
<td>12 (10.8)</td>
<td>2 (1.8)</td>
<td>6 (5.4)</td>
<td>5 (4.5)</td>
<td>21 (18.9)</td>
<td>86 (26.9)</td>
</tr>
<tr>
<td>back</td>
<td>18 (22.0)</td>
<td>13 (15.9)</td>
<td>9 (11.0)</td>
<td>8 (9.8)</td>
<td>9 (11.0)</td>
<td>6 (7.3)</td>
<td>1 (1.2)</td>
<td>2 (2.4)</td>
<td>16 (19.5)</td>
<td>86 (26.9)</td>
</tr>
<tr>
<td>Achilles tendon</td>
<td>20 (33.3)</td>
<td>4 (6.7)</td>
<td>7 (11.7)</td>
<td>6 (6.7)</td>
<td>6 (10.0)</td>
<td>1 (1.7)</td>
<td>0 (0.0)</td>
<td>6 (10.0)</td>
<td>12 (20.0)</td>
<td>86 (26.9)</td>
</tr>
<tr>
<td>foot</td>
<td>86 (26.9)</td>
<td>32 (10.0)</td>
<td>32 (10.0)</td>
<td>31 (9.7)</td>
<td>30 (9.4)</td>
<td>11 (3.4)</td>
<td>9 (2.8)</td>
<td>20 (6.3)</td>
<td>69 (21.6)</td>
<td>86 (26.9)</td>
</tr>
<tr>
<td>ankle</td>
<td>33 (29.7)</td>
<td>23 (20.7)</td>
<td>18 (22.0)</td>
<td>18 (22.0)</td>
<td>18 (22.0)</td>
<td>8 (13.3)</td>
<td>6 (5.4)</td>
<td>5 (4.5)</td>
<td>21 (18.9)</td>
<td>86 (26.9)</td>
</tr>
<tr>
<td>shoulder</td>
<td>52 (46.8)</td>
<td>23 (20.7)</td>
<td>18 (22.0)</td>
<td>13 (15.9)</td>
<td>13 (15.9)</td>
<td>4 (6.7)</td>
<td>1 (1.2)</td>
<td>2 (2.4)</td>
<td>16 (19.5)</td>
<td>86 (26.9)</td>
</tr>
<tr>
<td>head</td>
<td>49 (59.8)</td>
<td>49 (59.8)</td>
<td>18 (22.0)</td>
<td>8 (9.8)</td>
<td>6 (10.0)</td>
<td>1 (1.7)</td>
<td>1 (1.2)</td>
<td>6 (10.0)</td>
<td>12 (20.0)</td>
<td>86 (26.9)</td>
</tr>
<tr>
<td>other lower extremity sites</td>
<td>36 (60.0)</td>
<td>36 (60.0)</td>
<td>36 (60.0)</td>
<td>36 (60.0)</td>
<td>36 (60.0)</td>
<td>36 (60.0)</td>
<td>36 (60.0)</td>
<td>36 (60.0)</td>
<td>36 (60.0)</td>
<td>160 (50.0)</td>
</tr>
<tr>
<td>others</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>diagnosis [number (% of total injuries)]</th>
<th>inflammation or pain</th>
<th>sprain</th>
<th>strain</th>
<th>contusion</th>
<th>others</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>23 (34.4)</td>
<td>15 (22.4)</td>
<td>3 (4.5)</td>
<td>4 (6.0)</td>
<td>22 (32.8)</td>
</tr>
<tr>
<td>inflammation or pain</td>
<td>52 (46.8)</td>
<td>23 (20.7)</td>
<td>4 (3.6)</td>
<td>2 (1.8)</td>
<td>30 (27.0)</td>
</tr>
<tr>
<td>sprain</td>
<td>49 (59.8)</td>
<td>18 (22.0)</td>
<td>2 (2.4)</td>
<td>2 (2.4)</td>
<td>11 (13.4)</td>
</tr>
<tr>
<td>strain</td>
<td>36 (60.0)</td>
<td>8 (13.3)</td>
<td>2 (3.3)</td>
<td>2 (3.3)</td>
<td>12 (20.0)</td>
</tr>
<tr>
<td>contusion</td>
<td>160 (50.0)</td>
<td>64 (20.0)</td>
<td>11 (3.4)</td>
<td>10 (3.1)</td>
<td>75 (23.4)</td>
</tr>
<tr>
<td>others</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
<td>160 (50.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>severity [number (% of total injuries)]</th>
<th>trivial</th>
<th>low</th>
<th>moderate</th>
<th>severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>32 (47.8)</td>
<td>30 (44.8)</td>
<td>5 (7.5)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>trivial</td>
<td>21 (18.9)</td>
<td>58 (52.2)</td>
<td>27 (24.3)</td>
<td>5 (4.5)</td>
</tr>
<tr>
<td>low</td>
<td>14 (17.1)</td>
<td>51 (62.2)</td>
<td>16 (19.5)</td>
<td>1 (1.2)</td>
</tr>
<tr>
<td>moderate</td>
<td>16 (26.7)</td>
<td>26 (43.3)</td>
<td>16 (26.7)</td>
<td>2 (3.3)</td>
</tr>
<tr>
<td>severe</td>
<td>83 (25.9)</td>
<td>165 (51.6)</td>
<td>64 (19.1)</td>
<td>8 (2.5)</td>
</tr>
</tbody>
</table>

| discharge                              | 8 (6.1) | 18 (12.4) | 10 (9.3) | 16 (21.1) | 52 (11.3) |

** significantly lower (p=0.008) than in the three other military training schools.

Physical fitness to predict risk of overuse injuries

The presented fitness tests had no discriminative power for predicting the risk of acute injuries. However, TMS had significant discriminative power for predicting the risk of overuse injuries in all four study groups. PER was discriminative for predicting overuse injuries in three, OLS in two, and SLJ in one of the study groups. Only SSP had no significant power for predicting overuse injuries in any study group. The overuse injury risk of recruits who achieved the present IMFIS and those who did not are shown in Table 5.4. The highest area under the ROC-curve (AUC) was found for the PER (AUC = 0.706,
p = 0.000) and the TMS (AUC = 0.677, p = 0.002, Figure 5.1) to predict risk of overuse injuries in the fusilier infantry school. The subgroup of recruits, who failed one or more of these theoretical IMFIS of their specific boot camp, had an enhanced overuse injury rate of about 2.6 (rescue technicians), 2.2 (armored infantry), 2.4 (fusilier infantry), and 2.0 (reconnaissance infantry), compared to recruits who achieved all IMFISs (Table 5.4).

Figure 5.1: Receiver operating characteristic (ROC) analysis curve for trunk muscle strength test performance of 107 fusilier infantry recruits prior to their military service and their overuse injuries during 18 weeks of military training school. The area under the curve = 0.677 (p = 0.002); maximal sensitivity and specificity (0.641 and 0.721, respectively) for predicting overuse injuries was found at 135 s in the trunk muscle strength test.
Table 5.4: Overuse injury risk above and below injury-related minimum physical fitness standards.

<table>
<thead>
<tr>
<th>military training school</th>
<th>achieved IMFIS</th>
<th>standing long jump</th>
<th>trunk muscle strength test</th>
<th>one-leg standing test</th>
<th>progressive endurance run</th>
<th>all minimum fitness standards</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>OUI</td>
<td>p</td>
<td>n</td>
<td>OUI</td>
<td>p</td>
</tr>
<tr>
<td>rescue technicians</td>
<td>cut off</td>
<td>nd</td>
<td>130s</td>
<td>37s</td>
<td>nd</td>
<td>both</td>
</tr>
<tr>
<td>(n=131; OUI=23%)</td>
<td>yes</td>
<td>44</td>
<td>11%</td>
<td>101</td>
<td>18%</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>87</td>
<td>29%</td>
<td>30</td>
<td>40%</td>
<td>0.011</td>
</tr>
<tr>
<td>RR (95%CI)</td>
<td>2.53 (1.04-6.15)</td>
<td>2.24 (1.22-4.12)</td>
<td>2.56 (0.96-6.83)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>armored infantry</td>
<td>cut off</td>
<td>nd</td>
<td>181s</td>
<td>nd</td>
<td>937s</td>
<td>both</td>
</tr>
<tr>
<td>(n=145; OUI=38%)</td>
<td>yes</td>
<td>67</td>
<td>28%</td>
<td>75</td>
<td>24%</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>78</td>
<td>46%</td>
<td>70</td>
<td>53%</td>
<td>0.000</td>
</tr>
<tr>
<td>RR (95%CI)</td>
<td>1.63 (1.04-2.56)</td>
<td>2.20 (1.39-3.49)</td>
<td>2.16 (2.12-4.16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fusilier infantry</td>
<td>cut off</td>
<td>2.4m</td>
<td>135s</td>
<td>46s</td>
<td>700s</td>
<td>all four</td>
</tr>
<tr>
<td>(n=107; OUI=40%)</td>
<td>yes</td>
<td>58</td>
<td>29%</td>
<td>55</td>
<td>24%</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>49</td>
<td>53%</td>
<td>0.013</td>
<td>52</td>
<td>58%</td>
</tr>
<tr>
<td>RR (95%CI)</td>
<td>1.81 (1.12-2.92)</td>
<td>2.19 (1.31-3.66)</td>
<td>2.44 (1.44-4.14)</td>
<td>2.58 (1.67-3.99)</td>
<td>2.38 (0.96-5.93)</td>
<td></td>
</tr>
<tr>
<td>reconnaissance infantry</td>
<td>cut off</td>
<td>nd</td>
<td>158s</td>
<td>nd</td>
<td>940s</td>
<td>both</td>
</tr>
<tr>
<td>(n=76; OUI=43%)</td>
<td>yes</td>
<td>48</td>
<td>31%</td>
<td>42</td>
<td>36%</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>28</td>
<td>64%</td>
<td>0.005</td>
<td>34</td>
<td>53%</td>
</tr>
<tr>
<td>RR (95%CI)</td>
<td>2.09 (1.25-3.40)</td>
<td>1.49 (0.89-2.48)</td>
<td>2.04 (1.01-3.90)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** Overuse injury incidence (OUI) proportion in all four study groups for subjects achieving the injury-related minimum physical fitness standard (IMFIS) defined for this study and for those who performed below the cut off values are tabled. Data are shown for all fitness tests which are discriminative to predict OUI (p < 0.05, based on receiver operating characteristic analysis). nd: not discriminative; RR: Risk ratio in the group of subjects below the cut off values, and p: statistical significance of OUI proportion-comparison between groups of volunteers above (yes) and below (no) cut off fitness levels.

Discussion

The majority of injuries registered during 18 weeks of Swiss Army boot camp were overuse injuries, sited in lower extremities, diagnosed as inflammation or musculoskeletal pain and leading to limited duty of up to one week. In the four investigated military training schools, different physical fitness parameters turned out to be discriminative to predict overuse injuries. The TMS was discriminative in all, the PER in three, the OLS in two and the SLJ in one military training school. Only for SSP was the discriminative power to predict overuse injuries weak in all study groups.
Description of injuries in Swiss Army basic training

For comparisons of injury incidences among studies, differences in injury definitions and methods of data collection must be considered. Only studies with similar recruits and methodology to assess and define injuries are considered here.

The injury incidence rate in three out of four training schools in the present study (18.2 - 18.8 injuries per month per 100 male recruits) is above the range of 10 - 15 injuries per month per 100 male recruits found in meta analyses of previous review studies (Kaufman et al., 2000; National Research Council, 2006). There are actually a few other studies among US and Norwegian trainees that also found injury rates of over 15 injuries per month per 100 male recruits (Heir and Glomsaker, 1996; Knapik et al., 2001; Popovich et al., 2000).

The lower extremities are most often affected (62% of injuries). This finding is consistent with comparative studies referring to 58% to 78% of injuries to lower extremities (Almeida et al., 1999; Heir and Glomsaker, 1996; Knapik et al., 1993). The knee is the anatomical site most often affected in Swiss Army military service (27%) and is in the upper range of data in comparative studies (10% - 28%; Almeida et al., 1999; Jones et al., 1993; Knapik et al., 1993; Popovich et al., 2000). Injuries to the back are more common in Swiss Army recruits (10%), than in US Army trainees (up to 8%; Jones et al., 1993; Knapik et al., 1993; Popovich et al., 2000). On the other hand, the foot (10%) and ankle (9%) are less often affected than in comparative studies (11% - 26%, and 10% - 13%, respectively; Almeida et al., 1999; Jones et al., 1993; Knapik et al., 1993; Popovich et al., 2000). Health complaints concerning the Achilles tendon are more common in Swiss recruits compared to US recruits (10% vs. up to 3%; Knapik et al., 1993; Popovich et al., 2000).

The outlined differences between injury data in the present study compared to the literature may be the result of diverse causes. Differences in the kind of armed forces (the Swiss Army is compulsory), physical training programs, material (as jackboots), physical capabilities of the trainees, and inhibition thresholds to visit the medical care center may cause the diversity.
Physical fitness to predict risk of overuse injuries

It may be assumed that the physical demands on individuals in the same boot camp are similar. Therefore, less fit trainees fatigue faster because they perform at a higher percentage of their maximal physical capacity. Consequently, injuries and discharges may be more likely in less fit trainees. Previous studies found the strongest association between a fitness parameter and injury incidence in military service for aerobic endurance (Blacker et al., 2008; Gilchrist et al., 2000; Jones and Knapik, 1999). Furthermore, less consistent and less significant associations between other physical fitness measures such as push-ups or sit-ups and risk of injury were found (Blacker et al., 2008; Jones et al., 1993; Jones et al., 1993; Knapik et al., 1993; Reynolds et al., 1994). Consistent with the previous studies the aerobic endurance test used in the present study showed a strong discriminative power to predict overuse injury risk. In contrast to previous studies, our data suggest that the test used for trunk muscle fitness is at least as powerful as aerobic endurance assessments to predict injury risk among recruits. TMS is an isometric test of trunk muscle fitness, utilizing all trunk muscles including supporting shoulder, hip and leg muscles at the same time. As Bourban et al. (2001) quantified, this test detects the weakest link in the chain of individuals’ trunk muscles. Consistent with the expectation that fitness of supporting trunk muscles is important for jobs which require substantial handling of heavy materials or marching with fully-loaded back packs the TMS, with its good test-retest reliability (Tschopp et al., 2001; Wyss et al., 2007), showed strong discriminative power to predict overuse injuries.

A negative association between balance ability and physical training- or sport injury risk has been shown before in college students (Hrysomallis, 2007). However, to our knowledge, the present study is the first to assess a balance test’s discriminative power to predict injury incidence in a military setting. The present results indicate that the OLS may be an even better predictor of overuse injuries than tests of muscle power in upper and lower extremities (SSP and SLJ). The SSP was the only fitness test which failed to prove its discriminative power to predict overuse injuries in all study groups.
Injury-related minimum physical fitness standards for injury prevention

The cut off values used in the present study represent the value with the highest accuracy of the test in predicting overuse injuries. However, these values are not necessarily the ideal IMFIS. They are used as an example to quantify the effect on injury-prevention by the use of minimum standards at the recruitment of military organizations. Ideal IMFIS have first to be identified. For maximum injury protection, cut off values shall be high, increasingly precluding assignment of recruits. On the other hand, to find enough recruits for every occupational specialty, cut off values shall be low, reducing the effect on injury protection. The present data demonstrate the potential to prevent injuries by the use of IMFIS. For institutions with tools of continuous and automated military medical and personnel databases as recommended by the AFEB (Jones and Hansen, 2000), it would be only a small effort to periodically examine validity of used fitness tests and respective cut-off values for injury-prevention by the use of the presented method.

Limitations

The numbers of recruits in the four study groups differ. While in the armored infantry school data of 145 volunteers were collected, in the reconnaissance infantry school only 76 volunteers were investigated. However, every volunteer of the formation cycle in the four designated military training schools was assessed.

Fitness is only one of many risk factors for injuries during daily military routine. In a follow-up intervention study it would be necessary to assess further risk factors. However, the aim of the present study was to question validity of different fitness tests as injury prevention tool in selection of military personal only.
Conclusion

Injury incidence among Swiss recruits in physically demanding jobs is in the upper range of values found in other nations, especially injuries to the knee, lower back and Achilles tendon. The present study shows that the fitness tests used during the Swiss Army recruitment, especially the TMS and the PER, have a strong discriminative power to predict overuse injuries during military service. We identified that for different occupational specialties, different fitness parameters were discriminative to predict overuse injuries. We conclude that the fitness test battery used during the recruitment of the Swiss Army is appropriate to detect conscripts with enhanced risk of overuse injuries and therefore provides a valid indicator to select suitable personnel for physically demanding military occupational specialties.
6. Recognition of military-specific physical activities with body-fixed sensors

Published in Military Medicine, 2010, 175, 858-64.

Introduction

A mismatch between physical capability and physical job requirements can increase the risk of injury, jeopardize unit performance and decrease overall morale (Gledhill and Jamnik, 1992; Rayson, 1998; Rosendal et al., 2003; Sharp et al., 1998). Therefore, the importance of obtaining an accurate description of the job requirements in physically demanding occupations cannot be overestimated (Gledhill and Jamnik, 1992; Rayson, 1998), particularly in military organizations. Commonly used procedures to assess military- or fire fighting-specific job requirements include self-report questionnaires, interviews, observations and physical measurements (Chahal et al., 1992; Gledhill and Jamnik, 1992; Rayson, 1998; Stevenson et al., 1992). Bos et al. (2002) recommend describing the exposure to work demands objectively in terms of duration, frequency and intensity. To assess the duration and frequency of physical activities, direct observation or video observation are most precise but are impractical for large groups of participants. Using self-report questionnaires is the most practical approach in large scale studies, but their reliability, validity and objectivity are low (LaPorte et al., 1985; Lester et al., 2006).

The most promising method of assessing the duration, frequency and intensity of physical activities in large groups of participants is the use of body-fixed sensors. Several approaches have shown to be effective in recognizing specific activities based on data of diverse body-fixed sensors (Aminian et al., 1999; Pärkkä et al., 2004; Pober et al., 2006). However, none of these approaches has been adapted for a specific application in a military setting.

The most widely used body-wearable sensors measure acceleration (ACC) or heart rate (HR). Of the many different types of sensors, accelerometers supply the most useful data
for activity recognition (Lester et al., 2006). However, if worn only at the hip, additional physical efforts resulting from activities with low body movement but high muscle tension are not detected. The use of only a heart rate monitor is less useful for activity recognition because HR has a delayed reaction to activity changes and lacks specificity for any particular activity. A combination of ACC and HR data may enhance precision in activity recognition. The advantage of ACC is its immediate response to body movements and its information on respective intensity. On the other hand, even if HR has a delayed reaction to activity changes, it is more accurate when describing activities with low body movement but high physical intensity than is ACC data.

However, body-fixed sensors must meet several demands in order to be applicable in military workday life. They need enough memory and battery lifetime to record data continuously over at least one week. Sensors have to be waterproof, shock resistant and wearable with all military equipment. In the present study, data provided by the sensors were used to recognize the most relevant physically demanding activity classes in the context of armed forces. Authors in previous studies defined activities as physically relevant in military service if they are a frequent part of physically demanding tasks during daily military routine (Jones and Knapik, 1999; Knapik et al., 2002; Rayson, 1998; Rayson et al., 2000; Sharp et al., 1998; Sharp et al., 2006). Walking, marching with backpack, lifting and lowering loads, lifting and carrying loads, digging and running were named most often in those publications and are therefore investigated in the present study.

The aim of this study was to recognize physically relevant, military-specific activities using easy-to-handle body-fixed sensors, thereby demonstrating that it is possible to objectively assess the duration and frequency of physically demanding, military-specific activities using this technology.
Methods

Study design

There were three steps in the data acquisition (Figure 6.1). At first, 15 volunteers performed six single, military-specific activities according to protocol. Their data were used to develop algorithms for activity recognition. Secondly, 18 volunteers performed the same six isolated activities to estimate the accuracy of the activity recognition system. In the third step, sensor-based activity recognition was compared to observation-based activity assessment on 24 volunteers during daily military routine.

Participants and anthropometric parameters

All participants were male recruits from the Swiss Army (Table 6.1). All were volunteers recruited from two selected military occupational specialties (rescue technician and infantry recruits). The volunteers received comprehensive information about the study and provided written informed consent for their participation as approved by the Cantonal Ethics Committee of Bern, Switzerland and from the Swiss Army Sports and Preventions Competence Center. Age, body weight and height data were collected in the first two...
weeks of military service. All volunteers were measured by the same examiner using a calibrated digital balance and a measuring tape.

**Instruments**

ActiGraph uniaxial accelerometers (GT1M, ActiGraph LLC, Fort Walton Beach, FL) were used to monitor volunteers' waist ACC in vertical direction and step frequency. A second ACC was mounted on the backpack to register backpack carrying. The inter-monitor variability between different GT1Ms is very small (< 1%; Rothney et al., 2008). The GT1M is lightweight (27 g), compact (3.8 cm x 3.7 cm x 1.8 cm) and splashproof. Its rechargeable battery is capable of providing power for over 14 days without recharging, and it has a memory capacity of one megabyte. More detailed specifications have been described elsewhere (Tryon and Williams, 1996). In the present study, accelerometers were wrapped in waterproof plastic and were placed in a belt pouch on the waist over the right anterior axillary line and on the side-strap of the personal backpack. The GT1Ms were programmed to record acceleration and step-count data in two-second intervals so that data could be gathered over six continuous days.

A Suunto monitor (Suunto Smartbelt, Suunto, Vantaa, Finland) was used to measure volunteers' HR. A Suunto Smartbelt is a lightweight (61 g) and waterproof standalone monitor worn on the chest. Its exchangeable battery is capable of providing power for over four weeks of continuous measurement. One million heartbeats can be stored in the internal memory, enough for more than one week of continuous measurement. A Suunto Smartbelt registers HR as long as the chest strap is worn. In the present study, data were transferred to a computer, after a five-day monitoring period, at two-second intervals using Suunto Training Manager Version 2.2.0.

**Data collection protocol**

In a laboratory setting, 33 volunteers (Table 6.1) completed six activities – walking, marching with backpack (10-15 kg), lifting and lowering loads (30 kg), lifting and carrying loads (30 kg, 10-40 m), digging and running – for seven minutes each with two minutes rest between activities. Apart from running, which was the last activity for all volunteers,
the order of the activities was random. Data from 15 randomly chosen volunteers were used to develop algorithms for activity recognition. Data from the remaining 18 volunteers were used to estimate the accuracy of the activity recognition system.

During daily military routine, 24 randomly chosen volunteers (Table 6.1) from two different troops were individually observed in situ over a 90-minute period in order to estimate the accuracy of the algorithm for activity recognition during daily military routine. An examiner observed the volunteers in various training sections and classified their activities in 20-second intervals. Observation was always done by the same examiner. In addition, two scientists joined the observation of 14 volunteers to investigate the inter-rater reliability of direct observation.

Analysis

Statistical comparisons of volunteer's anthropometric data between study groups of three parts of data acquisition were performed with SPSS for Windows (version 16.0, SPSS Inc., Chicago, IL) with an alpha level of 0.05 to indicate statistical significance. Therefore a one-way analysis of variance (ANOVA) with Tukey post-hoc analysis was conducted.

All ACC and HR data were synchronized for every volunteer by a self-programmed application using Matlab (Matlab 5.3, MathWorks, Natick, Massachusetts). Mean heart rate was calculated using a sliding window with a window size of two minutes on five days of continuous heart rate data for every two seconds. The lowest value found in such a window was used as resting heart rate. Implausible data, caused by short duration artifacts in the signal, were detected by visual inspection of the plotted raw data. This data were excluded from the identification process of resting heart rate.

**Development of the activity recognition algorithm**

The development of an activity recognition algorithm was focused on the most frequent military-specific, physically demanding activities such as 1) walking, 2) marching with backpack, 3) lifting and lowering loads, 4) lifting and carrying loads, 5) digging and 6) running (Jones and Knapik, 1999; Knapik et al., 2002; Rayson, 1998; Rayson et al., 2000;
Apart from these, daily military routine contains many other activities. The purpose of the present study is to recognize the six specific activities only and to assign all remaining activities to the "other activities" class. Box plots for hip acceleration (H-ACC), heart rate above resting heart rate (HRaR), step frequency (SF) and backpack acceleration (BP-ACC) were plotted for every activity class (Figure 6.2). The box plots were used to verify the discrimination of the activity classes by their inherent sensor data. The specific activity classes were determined as data within the 1.5 interquartile range of corresponding labelled data. Areas out of specific data ranges of the six activity classes represent the "other activities" class. Based on the activity classes-specific data ranges, the nodes of the decision tree were defined. Classifications were made in two-second intervals. In a post-processing step, first activity assignments (0.5 Hz) were filtered to reduce the number of short-duration misclassifications (Pärkkä et al., 2004). The used filter replaces short activities with the surrounding longer duration activity. Therefore, first-classified data were buffered in 60-second time segments. If at least 20 of the 30 decisions in a 60-second time segment were the same, the respective activity class was assigned.

**Testing of the activity recognition system**

The recognition rates of activities classified based on sensor data were calculated and presented in a confusion matrix (Kohavi and Provost, 1998). Each row of the confusion matrix represents the instances in an actual activity class, while each column represents the instances in a predicted activity class. The recognition rate is defined by the number of true positive-classified instances divided by the number of total instances of the respective actual activity class. The overall recognition rate is defined by the sum of all true positive-classified instances of all activity classes divided by the total number of investigated instances.
Results

Participants

Age, weight and height of the volunteers in three parts of data acquisition of this study did not differ (p = 0.702, 0.776, and 0.142, respectively, Table 6.1).

**Table 6.1**: Volunteers’ respective age, weight, height and military training school in three parts of data acquisition of the present study.

<table>
<thead>
<tr>
<th>Data acquisition</th>
<th>n</th>
<th>Swiss Army training school</th>
<th>Age [y]</th>
<th>Weight [kg]</th>
<th>Height [cm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development of the activity recognition algorithm</td>
<td>15</td>
<td>rescue technicians</td>
<td>20.7 ± 1.1</td>
<td>75.6 ± 10.3</td>
<td>179.3 ± 6.5</td>
</tr>
<tr>
<td>Testing of activity recognition system in isolated activities</td>
<td>18</td>
<td>rescue technicians</td>
<td>21.0 ± 0.8</td>
<td>75.4 ± 8.6</td>
<td>176.8 ± 8.0</td>
</tr>
<tr>
<td>Testing of activity recognition system during daily military routine</td>
<td>24</td>
<td>rescue technicians and infantry</td>
<td>21.3 ± 2.7</td>
<td>77.3 ± 9.5</td>
<td>180.9 ± 5.2</td>
</tr>
</tbody>
</table>

**Discrimination of military-specific activity classes by sensor data**

First, running can be separated from the other five military-specific activities using H-ACC data (Figure 6.2A). Furthermore, walking and marching can be separated from two materials-handling classes (lifting and lowering loads, digging) using SF (Figure 6.2B). Finally, walking and marching with backpack can be distinguished by BP-ACC (Figure 6.2D). The discrimination power of the registered data does not allow the separation of the three materials-handling classes. Discrimination of the six military-specific activity classes by HRaR was weak (Figure 6.2C). However, HRaR is relevant to distinguish between the six specific and other less physically demanding activities of the "other activity" class.
Activities were first classified in two-second time segments without considering temporal connections (Figure 6.3A). The problem of the three materials-handling classes not being able to be separated still remained; therefore, simple temporal logic was used in a second step. First-classified data were buffered in 60-second time segments. If at least 20 of the 30 decisions in a 60-second time segment were the same, the respective activity class was assigned (Figure 6.3B). If the assigned class was the cumulative class of materials handling, it was further separated into lifting and lowering loads (H-ACC < 42 c/2s) or digging (H-ACC ≥ 42 c/2s), depending on the mean H-ACC. For the class lifting and carrying loads, the 30 decisions in 60-second time segments were analyzed further. In this segment, short classifications as cumulative materials-handling and walking alternated cyclically. On average, 44% of the 30 two-second decisions in the 60-second time segment of lifting and carrying loads was assigned as materials-handling activity, 33% as walking and 23% as other activities. Based on that distribution, the last decision was made. If in a
60-second time segment at least 11 first decisions were materials handling and 8 were walking or marching, the lifting and carrying loads class was assigned. Otherwise, the segment was assigned to the heterogeneous "other activities" class (Figure 6.3).

Testing of the activity recognition system in isolated activities

The overall recognition rate of isolated activities assessed in a laboratory setting was 87.5% (walking: 95%, marching with backpack: 95%, running: 85% and materials-handling classes: 76%). Within the materials-handling classes, 60% of lifting and lowering loads was classified true positive and 22% was incorrectly classified as digging. Also, 60% of digging was classified true positive and 15% was incorrectly classified as lifting and lowering loads. Only 42% of the lifting and carrying loads class was classified true positive, while 33% was incorrectly classified as walking and 6% as lifting and lowering loads.

Comparison of the activity recognition system with observation during daily military routine

The overall recognition rate of activities classified by the sensor-based activity recognition system compared to observation-based activity classification during daily military routine was 85.5%. True positive recognition for the military-specific activity classes ranged from 48% (materials-handling classes) to 89% (marching with backpack; see confusion matrix in Table 6.2).
Figure 6.3A-B: Decision tree in two phases as an activity classifier. In phase A, sensor data sampled at 0.5 Hz is used to distinguish between walking (C1), marching with backpack (C2), materials-handling activities (C3-5), running (C6) and other activities (C0). In phase B, results of part A are buffered in 60-second sequences to filter short duration misclassifications and to distinguish between all six relevant, military-specific physically demanding activity classes and the "other activities" class.
Table 6.2: Confusion matrix with actual (observation based) vs predicted (sensor data based) activity classes assessed during daily military routine.

<table>
<thead>
<tr>
<th>sensor based activity classes</th>
<th>observation based activity classes</th>
<th>walking</th>
<th>marching w. backp.</th>
<th>lifting &amp; low. loads</th>
<th>lifting &amp; carr. loads</th>
<th>digging</th>
<th>running</th>
<th>other activities</th>
<th>sum (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>walking</td>
<td>walking</td>
<td>92 (66%)</td>
<td>0 (0%)</td>
<td>2 (1%)</td>
<td>8 (6%)</td>
<td>0 (0%)</td>
<td>1 (1%)</td>
<td>37 (26%)</td>
<td>140 (100%)</td>
</tr>
<tr>
<td>marching with backpack</td>
<td>0 (0%)</td>
<td>197 (89%)</td>
<td>1 (0%)</td>
<td>8 (4%)</td>
<td>1 (0%)</td>
<td>0 (0%)</td>
<td>15 (7%)</td>
<td>222 (100%)</td>
<td></td>
</tr>
<tr>
<td>lifting and lowering loads</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>46 (48%)</td>
<td>1 (1%)</td>
<td>3 (3%)</td>
<td>0 (0%)</td>
<td>45 (47%)</td>
<td>95 (100%)</td>
<td></td>
</tr>
<tr>
<td>lifting and carrying loads</td>
<td>7 (10%)</td>
<td>2 (3%)</td>
<td>9 (13%)</td>
<td>14 (20%)</td>
<td>1 (1%)</td>
<td>0 (0%)</td>
<td>37 (53%)</td>
<td>70 (100%)</td>
<td></td>
</tr>
<tr>
<td>digging</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>4 (18%)</td>
<td>1 (5%)</td>
<td><strong>11 (50%)</strong></td>
<td>0 (0%)</td>
<td>6 (27%)</td>
<td>22 (100%)</td>
<td></td>
</tr>
<tr>
<td>running</td>
<td>1 (4%)</td>
<td>1 (4%)</td>
<td>0 (0%)</td>
<td>1 (4%)</td>
<td>0 (0%)</td>
<td><strong>16 (70%)</strong></td>
<td>4 (17%)</td>
<td>23 (100%)</td>
<td></td>
</tr>
<tr>
<td>other activities</td>
<td>15 (1%)</td>
<td>2 (0%)</td>
<td>79 (5%)</td>
<td>24 (2%)</td>
<td>3 (0%)</td>
<td>0 (0%)</td>
<td><strong>1510 (92%)</strong></td>
<td>1633 (100%)</td>
<td></td>
</tr>
<tr>
<td>sum</td>
<td>115</td>
<td>202</td>
<td>141</td>
<td>57</td>
<td>19</td>
<td>17</td>
<td>1654</td>
<td>2205</td>
<td></td>
</tr>
</tbody>
</table>

Note: In 24 volunteers a total of 2205 minutes of daily military activities was measured. Unit: minutes (%).

Inter-rater reliability of direct observation

Total data from three examiners matched 91.8% to 92.6% of the pairwise-compared instances. The concordance of activity classification based on observation for the six activities ranged from 51% between examiners one and two in lifting and carrying loads to 100% between examiners two and three in running (see Table 6.3).

Table 6.3: Accordance of observation-based activity classification by three different examiners following, simultaneously, the same volunteer.

<table>
<thead>
<tr>
<th>activity class</th>
<th>examiner 1 vs. examiner 2</th>
<th>examiner 1 vs. examiner 3</th>
<th>examiner 2 vs. examiner 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>walking</td>
<td>81%</td>
<td>81%</td>
<td>87%</td>
</tr>
<tr>
<td>marching with backpack</td>
<td>99%</td>
<td>99%</td>
<td>98%</td>
</tr>
<tr>
<td>lifting and lowering loads</td>
<td>83%</td>
<td>57%</td>
<td>79%</td>
</tr>
<tr>
<td>lifting and carrying loads</td>
<td>51%</td>
<td>72%</td>
<td>54%</td>
</tr>
<tr>
<td>digging</td>
<td>63%</td>
<td>100%</td>
<td>83%</td>
</tr>
<tr>
<td>running</td>
<td>92%</td>
<td>85%</td>
<td>100%</td>
</tr>
<tr>
<td>other activities</td>
<td>96%</td>
<td>97%</td>
<td>94%</td>
</tr>
</tbody>
</table>

Note: Twelve volunteers were observed for 90 minutes each.
Discussion

Direct observation (Dybel and Seymour, 1997), energy expenditure estimations by doubly labelled water (DLW; Jones et al., 1993; Wilkinson et al., 2008), and self-report questionnaires (Lang et al., 2007; Pope et al., 1998) are the most common methods of assessing job requirements in armed forces. Unfortunately, direct observations are not feasible for large-scale studies; DLW does not differ between activity classes, and self-report questionnaires are of low objectivity. Established body-fixed sensors were used in the present study to objectively assess duration and frequency of military-specific physically demanding activities in larger groups. Algorithms have thus been developed to recognize six military-specific activities.

The overall recognition rate of the presented activity recognition system of 87.5% for isolated activities and 85.5% for activities observed during daily military routine is high and comparable with findings in other studies (Aminian et al., 1999; Pärkkä et al., 2004; Pober et al., 2006). However, there are important differences concerning the methods between the studies in terms of choice and number of assessed activities, in numbers of sensors used and their data density, and in validation methods. Pober et al. (2006) achieved a mean recognition rate of 80.8% by classifying four activities (walking, walking uphill, vacuuming and computer work) with one sensor signal (chest acceleration, 1Hz) in a laboratory setting. Pärkkä et al. (2004) found a mean recognition rate of 86% by classifying five activities and three posters (running, nordic walking, walking, rowing, cycling, sitting, standing and lying) with 22 different signals (synchronized by 1 Hz) in an out-of-laboratory environment. Aminian et al. (1999) showed a mean recognition rate of 89.3% by classifying three posters and two activities (sitting, standing, lying, dynamic activities and other activities) with two sensor signals (chest and thigh acceleration, 10 Hz) in a laboratory setting.

The recognition rate of specific activities investigated during daily military routine was lower than in isolated activities performed after protocol. This is due to a greater variability in activity durations and intensities during daily military routine. Especially, activities with short durations have a lower recognition rate because with every change of activity, a difference between sensor-based and observation-based activity classification is likely. If the activity changes often, direct observation is more difficult and subjective appraisal is more important, like in frequent changes between walking and standing, for
example. Misclassifications of activities with short duration explain why walking achieves a lower recognition rate than marching with backpack as an example of an activity with longer durations. All materials-handling classes show only moderate recognition rates (Table 6.2). However, with over 90 minutes of data collection per subject, on group level, false negative and false positive misclassifications cancel each other at random. Table 6.2 shows that the total sum of instances classified based on sensor data is similar to the total sum of observed instances for all activity classes.

Only the true positive classification of the lifting and carrying loads class was apparently low. However, parts of lifting and carrying loads activities were classified as either walking and marching (33% in laboratory setting and 13% during daily routine) or lifting and lowering loads (6% in laboratory setting and 13% during daily military routine, see Table 6.2). These classifications are not entirely wrong because the class carrying loads contains walking and, to a small extent, lifting. Apart from that, there are no systematic misclassifications between the six military specific activity classes.

Possibilities for enhancing output in activity recognition

The lifting and carrying loads class is important in the military setting because walking with heavy loads is much more physically demanding than simply walking. It is worthwhile to attempt to enhance the respective recognition rate of the used method. We suggest investigating temporal patterns in the data using Hidden Markov Models (Duda et al., 2001), for example. Temporal patterns are suggested to be useful in recognizing lifting and carrying loads because this activity class is composed by cyclic alternations of a small number of short activities. In order to test such an approach, a new dataset with higher resolution and precision of the label segments is needed. With the use of more complex sensors, higher data density and additional sensors placed elsewhere on the body, the accuracy of the activity recognition system may be increased. Unfortunately, continuously assessing physical activities over one week of military service puts very high restrictions on sensors and body positions. Therefore, it is important to maintain a balance between accuracy and feasibility, especially in this setting.
Limitations

Although direct observation was found to have good inter-observer reliability in general, it is unlikely that it is entirely precise. The comparison of observations of three examiners showed enhanced variances in short duration activities. The concordance for the lifting and lowering loads class between different examiners was only 51-72\%, for example (Table 6.3). The use of video observation would have been ideal, as it may provide a more accurate reference for sensor-based activity recognition. Video analysis allows for watching a specific sequence several times, using slow motion and other software functions that facilitate the task, allowing for definition of label segments with a higher resolution and precision. However, the use of video was not allowed in the Army’s daily routine and video analysis can be expensive.

In the presented study, volunteers were observed during randomly chosen 90 minutes of their daily military routine. The disadvantage of this approach is its unequal outcome in duration and frequency of different military-specific activities. Unfortunately, the dataset sampled during daily military routine contained only 22 minutes of digging and 23 minutes of running. Each of these activities represents 1\% of the registered activity time (Table 6.2). However, to counter this limitation, the military-specific activity recognition system was additionally compared to isolated activities performed after protocol containing the same duration for every activity class in every subject.

Strengths

The developed algorithms for military-specific activity recognition are validated not only in isolated activities in a laboratory setting but also during daily military routine. Therefore, the results are more meaningful for future applied studies.

The presented classification method is simple to use and comprehensible. Additionally, the algorithm of this study can be combined with algorithms from prior studies to estimate activity intensities (Brage et al., 2004; Swartz et al., 2000). Those algorithms for energy expenditure estimation are based on the same sensor signals (uniaxial accelerometry and HR monitors). However, such algorithms have to first be validated in a military setting.
Relevance for future applications

Body-fixed sensors have been applied successfully in recent studies to investigate job requirements in military occupations (Knapik et al., 2007; Wilkinson et al., 2008; Wixted et al., 2007). With the algorithm presented in the current study body-fixed sensors deliver not only previously used indexes of activity intensities (ACC and HR; Wilkinson et al., 2008; Wixted et al., 2007) and walking distances (SF; Knapik et al., 2007), but also information on type, duration and frequency of military-specific activities.

The advantages of the chosen body-fixed sensors for future investigations in military live action are the ability to collect and store data of many participants, without any technical support or recharging over one week. In contrast to prior observation studies (Bos et al., 2004; Pope et al., 1998), there is no need for a researcher to accompany the participants during military field exercises with the current approach. A limitation of this approach when applied in the field may be the reduced control of participants' commitment to wearing the sensors.

The presented algorithm was developed to provide scientific answers in the field of occupational medicine, injury prevention and physical training, in military settings. Additionally, physical activities and demands can be determined in order to develop job descriptions in military organizations. So far, a relation between general physical demands and injury incidence has been demonstrated (Almeida et al., 1999; Jones et al., 1994; Jordaan and Schwellnus, 1994). The present method can provide useful information to further specify physical demands-related injury risk factors. Therefore, progression, type, amount and frequency of physical demands during military basic training can be assessed and compared with occurrences of injuries, dismissals or changes in physical performances.
Conclusion

Established, easy-to-handle body-fixed sensors deliver data for specific and valid activity recognition in a military setting. With the discussed sensors and the developed algorithm, military-specific activities can be recognized in one-minute intervals over six continuous days. The presented method allows investigators to objectively assess type, occurrence, duration and frequency of military-specific physical activities.
7. Energy expenditure estimation during daily military routine with body-fixed sensors

Introduction

In physically demanding occupations such as those in a military setting, an optimal balance between physical job requirements and individual physical capability is crucial in terms of injury prevention, unit performance and morale (Gledhill and Jamnik, 1992; Rayson, 1998; Rosendal et al., 2003; Sharp et al., 1998). The first step in ensuring this balance is obtaining an accurate description of the physical job requirements (Gledhill and Jamnik, 1992; Rayson, 1998). Procedures commonly used to assess military- or firefighting-specific job requirements include self-report questionnaires, interviews, observations and physical measurements, such as heart rate, energy expenditure and oxygen consumption (Chahal et al., 1992; Gledhill and Jamnik, 1992; Rayson, 1998; Stevenson et al., 1992). Bos et al. (2002) as well as Almeida et al. (1999) recommended assessing the exposure to work demands objectively in terms of duration, frequency and intensity. Each of the methods listed above has some limitations. While self-report questionnaires and interviews have low objectivity, observations are not feasible in large study groups, and heretofore used physical measurements yielded no information on activities’ characteristics or on the distribution during the daily military routine. Therefore, in Chapter 6 (Wyss and Mäder, 2010) of this thesis a new method was adapted to the military setting, using body-fixed sensor data to identify and measure the duration and frequency of the six most common physically demanding, military-specific activities (walking, marching with backpack, lifting and lowering loads, lifting and carrying loads, digging and running; Jones and Knapik, 1999; Knapik et al., 2002; Rayson, 1998; Rayson et al., 2000; Sharp et al., 1998; Sharp et al., 2006). This method does not completely describe exposure to work demands since it fails to assess the intensity of activities. However, the same data signals (hip acceleration and heart rate) collected for this method of military-specific activity recognition have
previously been applied to estimate intensity of activities. Algorithms from these prior studies (Brage et al., 2004; Swartz et al., 2000) or newly developed, military-specific algorithm for energy expenditure estimation (EEE) require validation in a military setting before being routinely applied to obtain more complete descriptions of work-related physical demands.

The aim of the present study was to develop and validate activity-class-specific multiple linear regressions (AS-MLR) to estimate physical activity energy expenditure (PAEE) during the daily military routine, using easy-to-handle body-fixed sensors. Further, the accuracy of EEE with AS-MLR was compared to the algorithms presented in prior studies (Brage et al., 2004; Swartz et al., 2000), developed in a different setting. For AS-MLR, the activity class will be assigned first, based on Wyss and Mäder’s (2010) algorithm. Secondly, for every activity class, a specific multiple linear regression with heart rate (HR), hip acceleration (H-ACC) and body weight as the independent variables will be applied for EEE.

Methods

Participants

All participants were male recruits from the Swiss Army (Table 7.1). All were volunteers recruited from two selected military occupational specialties (reconnaissance and fusilier infantry training school). The volunteers received comprehensive information about the study and provided written informed consent for their participation as approved by the Cantonal Ethics Committee of Bern, Switzerland, and from the Swiss Army Sports and Preventions Competence Center.

Study design and protocol

In both parts of data acquisition volunteers were equipped with a portable spirometer and body-fixed sensors. Firstly, eight volunteers performed isolated, military-specific activities in a laboratory setting according to the protocol described below. Their data were used to
develop AS-MLR with PAEE as the dependent variable and HR, H-ACC, and body weight as the independent variables. Secondly, twelve volunteers were investigated during the daily military routine to validate the accuracy of AS-MLR for EEE (Figure 7.1).

In a laboratory setting, eight volunteers (Table 7.1) completed six military-specific, physically-demanding activities, namely, walking, marching with a backpack (10-15 kg), lifting and lowering loads (30 kg), lifting and carrying loads (30 kg, 10-40 m), digging and running. Further, other less demanding activities such as sitting, office work, cleaning dishes, mopping the floor, cleaning shoes or manipulating a weapon were investigated. The order and duration (3-7 minutes) of the activities were randomized. Data were used to develop AS-MLR for PAEE estimation based on sensor data.

During the daily military routine, the energy expenditure of twelve randomly chosen volunteers (Table 7.1) from two different occupational specialties was recorded with an indirect calorimetry (portable spirometer) over a 90-minute period for each volunteer. Data were used to assess the accuracy of the sensor-based PAEE estimation.

**Measurements and instruments**

Age, body weight and height data were collected during the first two weeks of military service. Therefore, a measuring tape and a calibrated balance (Seca model 877, Seca GmbH, Hamburg, Germany) were used.
ActiGraph uniaxial accelerometers (GT1M, ActiGraph LLC, Fort Walton Beach, FL) were used to monitor volunteers' hip acceleration (H-ACC) in the vertical direction and step frequency. A second accelerometer was mounted on the backpack to register backpack carrying. The inter-monitor variability between different GT1Ms is very small (< 1%; Rothney et al., 2008). The GT1M is lightweight (27 g), compact (3.8 cm x 3.7 cm x 1.8 cm) and splash-proof. Its rechargeable battery is capable of providing power for more than 14 days without recharging, and has a memory capacity of one megabyte. More detailed specifications have been described elsewhere (Tryon and Williams, 1996). In the present study, accelerometers were wrapped in waterproof plastic and were placed in a belt pouch on the waist over the right anterior axillary line and on the side strap of the personal backpack. The GT1Ms were programmed to record acceleration and step-count data in two-second intervals so that data could be gathered over six continuous days.

A Suunto monitor (Suunto Smartbelt, Suunto, Vantaa, Finland) was used to record HR. It is a lightweight (61 g) and waterproof standalone monitor worn on the chest. The monitor’s exchangeable battery is capable of providing power for more than four weeks of continuous measurement. One million heartbeats can be stored in the internal memory, enough for more than one week of continuous measurement. A Suunto Smartbelt registers the HR as long as the chest strap is worn. In the present study, data were transferred to a computer, after a five-day monitoring period, at two-second intervals using Suunto Training Manager Version 2.2.0.

Indirect calorimetry (METAMAX, Cortex Biophysik GmbH, Leipzig, Germany) measured the energy expenditure. The reliability and validity of this specific, mobile spirometric system have been shown to be very good (intra-class reliability of oxygen uptake, carbon dioxide output and minute ventilation: \( r > 0.973 \); Meyer et al., 2001; no significant output differences found compared to a stationary spirometric system OXYCONgamma by Mijnhardt, Netherlands; Schulz et al., 1997). The turbine flowmeter was calibrated with a three-liter calibration syringe, and the \( \text{O}_2 \) and \( \text{CO}_2 \) sensors were calibrated with room air and calibration gas (16% \( \text{O}_2 \), 5% \( \text{CO}_2 \), 79% \( \text{N}_2 \)) before each testing session.
Analysis

The assessed data were H-ACC, BP-ACC, HR, energy expenditure measured by indirect calorimetry, body weight, height and age. Data were synchronized for every volunteer by a self-programmed application using Matlab (Matlab 5.3, MathWorks, Natick, Massachusetts).

To assess the resting HR, first the mean HR in a sliding window with a duration of two minutes was calculated. This process was repeated for every two seconds on five days of continuous HR data. Finally, the lowest value found in such a window was used as the resting HR. Implausible data, caused by short duration artifacts in the signal, were detected by visual inspection of the plotted raw data. These data were excluded from the identification process of resting HR.

To estimate PAEE based on the comparative algorithms of Brage et al. (2004; without individual calibration) and Swartz et al. (2000; based on H-ACC data only), the GT1M acceleration monitor data were divided by a constant factor (0.91). This factor was introduced by Corder et al. (2007) to compensate for the differences in the data output of two generations of Actigraph accelerometers. In studies by Brage et al. (2004) and Swartz et al. (2000), a prior generation of Actigraph accelerometers (Model 7164) was used.

The measured PAEE was calculated as total energy expenditure (TEE) minus resting energy expenditure (REE) values. TEE was assessed with indirect calorimetry using Péronnet and Massicotte’s formula (Jeukendrup and Wallis, 2005; Peronnet and Massicotte, 1991). REE was calculated using anthropometric data and the formula for men by Mifflin et al. (1990). Military-specific activities were classified in one-minute sequences based on Wyss and Mäder’s algorithm (2010; presented in Chapter 6) using H-ACC, BP-ACC, step frequency and the HR above the resting heart rate (HRaR).
Statistical analysis

SPSS version 16.0 for Windows (SPSS Inc., Chicago, IL) was used for all statistical analyses, with an alpha level of 0.05 to indicate statistical significance.

Multiple linear regressions between PAEE as the outcome variable and HRaR, H-ACC and body weight as the independent variables were calculated separately for walking, marching with backpack, cumulative materials-handling class (lifting and lowering loads, lifting and carrying loads and digging), running and the “other activities” class. Lifting and lowering loads, lifting and carrying loads and digging were merged in one activity class, called materials-handling, because those activities can often hardly be separated. For the multiple linear regressions, HRaR, H-ACC and body weight data were included only when they had a significant relationship to PAEE.

A Shapiro-Wilk test (n < 50) and a Kolmogorov-Smirnov test (n > 50) were conducted to determine whether the data were normally distributed. A paired t-test was used to examine differences in normally distributed anthropometric data between the study groups in the two parts of data acquisition and between the volunteers’ estimated and measured mean PAEE values. Differences were presented as means ± standard deviations (SD). Errors of estimated mean PAEE values were quantified as the root mean sum of squared errors (RMSE) and 95% confidence intervals (CI-95%). Bland-Altman plots were used to visualize systematic errors in PAEE predictions (Bland and Altman, 1986). A Wilcoxon signed-rank test was used to examine the differences between not normally distributed data. In some single activity classes, PAEE values of all assessed one-minute sequences were not normally distributed. Therefore, the Wilcoxon signed-rank test was applied on all comparisons between the estimated and measured PAEE values within single activity classes. Differences were presented as the median and interquartile range (IQR) of differences.
Results

The age, weight and height of the volunteers in the two parts of data acquisition of this study did not differ (p = 0.323, 0.679, and 0.823, respectively, see Table 7.1).

Table 7.1: Volunteers’ respective age, weight, height and military training school in the two parts of data acquisition in the present study.

<table>
<thead>
<tr>
<th>Part of the study</th>
<th>n</th>
<th>Military training school</th>
<th>Age [y]</th>
<th>Weight [kg]</th>
<th>Height [cm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development of AS-MLR for EEE</td>
<td>8</td>
<td>Swiss Army reconnaissance infantry</td>
<td>21.0 ± 0.9</td>
<td>72.1 ± 7.0</td>
<td>178.8 ± 4.3</td>
</tr>
<tr>
<td>Testing of AS-MLR for EEE during daily</td>
<td>12</td>
<td>Swiss Army fusilier infantry</td>
<td>21.1 ± 3.0</td>
<td>72.6 ± 10.7</td>
<td>179.3 ± 4.5</td>
</tr>
<tr>
<td>military routine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**AS-MLR:** activity class-specific multiple linear regressions; EEE: energy expenditure estimation.

Development of AS-MLR for EEE

Investigated isolated activities collected for eight volunteers added up to a total time of 284 minutes of PAEE and sensor data (25 to 50 minutes per volunteer). Derived activity-class-specific, multiple linear regressions are presented in Table 7.2. In this group of volunteers, body weight had no significant relationship to PAEE.
Table 7.2: Multiple linear regressions, respective correlations (r) and root mean squared errors (RMSE) for physical activity energy expenditure (PAEE) as the dependent variable and the heart rate above resting heart rate (HRaR) and hip acceleration (H-ACC) as the independent variables for five different activity classes.

<table>
<thead>
<tr>
<th>activity class</th>
<th>linear regression</th>
<th>r</th>
<th>p</th>
<th>RMSE [kJ/min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>walking</td>
<td>PAEE [kJ/min] = HRaR [bpm] * 0.5859 - 5.1508</td>
<td>0.701</td>
<td>0.000</td>
<td>5.7</td>
</tr>
<tr>
<td>marching with bp</td>
<td>PAEE [kJ/min] = HRaR [bpm] * 0.7810 - 15.9337</td>
<td>0.848</td>
<td>0.000</td>
<td>6.6</td>
</tr>
<tr>
<td>materials handling</td>
<td>PAEE [kJ/min] = HRaR [bpm] * 0.4485 + H-ACC [cpm] * 0.0023 - 1.5166</td>
<td>0.705</td>
<td>0.000</td>
<td>6.4</td>
</tr>
<tr>
<td>running</td>
<td>PAEE [kJ/min] = HRaR [bpm] * 0.7542 - 8.7800</td>
<td>0.877</td>
<td>0.000</td>
<td>9.2</td>
</tr>
<tr>
<td>other activities</td>
<td>PAEE [kJ/min] = HRaR [bpm] * 0.4840 + H-ACC [cpm] * 0.0010 - 4.7964</td>
<td>0.800</td>
<td>0.000</td>
<td>6.1</td>
</tr>
</tbody>
</table>

*note* bp: back pack

Testing of AS-MLR for EEE during the daily military routine

For 12 volunteers, 757 minutes, of a total of 1080 investigated minutes, of energy expenditure data were collected. One to 64 minutes of the spirometer data for ten of the 12 volunteers were lost due to technical problems. The volunteers’ mean estimated PAEE (20.97 ± 8.31 kJ/min), based on AS-MLR, did not differ from the mean measured PAEE (21.08 ± 8.00 kJ/min, p = 0.898, RMSE = 2.21 kJ/min and CI-95% = -1.97 to 1.75 kJ/min). The correlation between the estimated and the measured mean PAEE values in the 12 volunteers was high (r = 0.936, p < 0.001). The correlation between the alternatively estimated PAEE values (with algorithms from Brage et al. (2004) and Swartz et al. (2000), respectively) and the measured PAEE values was large as well (r = 0.915, p < 0.001 and r = 0.705, p = 0.011). However, these algorithms resulted in a significant underestimation of PAEE (-25.12 and -32.47%, respectively; p = 0.002 and p < 0.001, RMSE = 5.29 and 6.84 kJ/min, and CI-95% = -7.42 to -3.16, and -10.45 to -3.23 kJ/min, respectively). The Bland-Altman plot (Figure 7.2A) shows that no systematic error was found with AS-MLR for PAEE estimation. On the other hand, using Brage et al.’s (2004) and Swartz et al.’s (2000) algorithms, volunteers’ mean PAEE was systematically underestimated (Figures 7.2B and 7.2C).
Figures 7.2A-C: Bland-Altman plots on the average physical activity energy expenditure (PAEE) of 12 volunteers measured with indirect calorimetry and estimated using body-fixed sensor data for (A) activity-class-specific multiple linear regressions (AS-MLR), (B) Brage et al.’s (2004) algorithm and (C) Swartz et al.’s (2000) algorithm. The full lines represent the mean and ± 2SD of the difference between the measured and estimated PAEE values; the dashed line represents the zero line.

Activity classification showed that not every volunteer participated in all military-specific activity classes during the randomly chosen period of daily military service. The median of the differences between estimated and measured PAEE, of a total of 43 minutes assessed of the walking activities of ten volunteers, was -4.419 kJ/min (p = 0.002, IQR = 11.85 kJ/min). The median of the differences between estimated and measured PAEE, of a total of 84 minutes assessed of marching with backpack sequences by six volunteers, was -0.436 kJ/min (p = 0.046, IQR = 9.74 kJ/min). Between estimated and measured PAEE, of a total of 90 minutes assessed of materials-handling activities investigated among 12 volunteers,
no significant difference was registered (median of differences = -0.652 kJ/min, \( p = 0.941, \) IQR = 16.40 kJ/min). Only four volunteers ran for a short period while wearing the portable spirometer (13 min total). The respective median of the differences between estimated and measured PAEE was 1.193 kJ/min (\( p = 0.917, \) IQR = 15.58 kJ/min). Between estimated and measured PAEE, during a total of 527 minutes of other activities investigated among 12 volunteers, no significant difference was registered (median of differences = -0.421 kJ/min, \( p = 0.270, \) IQR = 7.57 kJ/min).

**Discussion**

The AS-MLR developed in the present study for PAEE estimation have been shown to be accurate and without systematic error. Meanwhile, the estimations according to Brage et al. (2004) and Swartz et al. (2000) have been shown to underestimate PAEE during the military routine (see Figure 7.2A-C). However, in these researchers’ studies, other activities were investigated, and these algorithms are thus probably less adequate for the specific military setting. Many military-specific activities, especially those resulting in high HR but low H-ACC data such as materials-handling classes, are less frequent in other settings. Therefore, we conclude that it is reasonable to develop and use setting-specific algorithms for EEE. Further, we conclude that continuous information about the current activity class is very valuable for precise EEE. This conclusion is supported by previous studies showing that the rates of the linear relation between H-ACC or HR and PAEE may differ if assessed in different activity classes. Collins et al. (1991) studied the relationship between HR and oxygen uptake (\( \text{VO}_2 \)) during weight lifting. They concluded that the HR/\( \text{VO}_2 \) relationship during weight lifting exercise is linear but is different from that reported for dynamic low-resistance exercise such as running or cycling. Choi et al. (2005) found that PAEE increases as the intensity of activities (H-ACC) increases with different rates depending on the type of activity. The researchers suggested that new algorithms estimating PAEE based on H-ACC data will depend on the types of activity.
Body weight had no significant influence on PAEE in any of the activity classes. This might be due to the homogeneity within the group of volunteers randomly chosen for the development of the energy expenditure estimation algorithm (see Table 7.1). While Wyss and Mäder (2010) showed that H-ACC is more important for activity recognition than HR data, the present study demonstrated that for assigning intensity of activities in a military setting HR is more powerful than H-ACC data (see Table 7.2).

**Implication**

Direct observation (Dybel and Seymour, 1997), energy expenditure estimation by doubly labeled water (DLW; Jones et al., 1993; Wilkinson et al., 2008) and self-report questionnaires (Lang et al., 2007) are the most common methods of assessing physical job requirements in the armed forces. Unfortunately, direct observations are not feasible for large-scale studies, DLW does not give any information about the distribution of PAEE during the daily military routine and self-report questionnaires have low objectivity. In the present study, established body-wearable sensors were used, and multiple linear regressions have been developed to objectively assess the intensity of military-specific physical activities in one-minute intervals during the daily military routine. Combined with Wyss and Mäder’s (2010) activity-recognition algorithm, the duration, frequency and intensity of military-specific physical activities can be assessed, and therefore physical job requirements can be registered objectively, feasibly and with information on the distribution of physical demands during the daily military routine.

**Limitations**

Because of technical problems with the internal memory of the portable spirometer, only data transferred wirelessly to the examiners’ computer on-site were saved. This transition did not always work from the beginning of the 90 minutes of data acquisition. Therefore, the dataset for ten volunteers was of reduced duration.

Activity misclassifications by Wyss and Mäder’s (2010) algorithm can lead to reduced accuracy in the presented activity-class-specific PAEE estimation. However, misclassifications within the three materials-handling classes are not relevant for PAEE
estimation, since the same regression model was used for these three classes. The residual activity misclassifications did not lead to a significant deviation between the estimated and measured mean PAEE values. For PAEE estimation within single activity classes, activity misclassifications might have had a negative influence on the accuracy of the estimation.

In the present study, volunteers were investigated during randomly chosen 90 minutes of their daily military routine. The disadvantage of this approach is its unequal outcome in duration, frequency and intensity of different military-specific activities. Unfortunately, the dataset sampled during the daily military routine contained only four volunteers who did some running. Additionally, the respective periods were short (13 minutes in total, 1.3% of the registered activity time). However, the primary interests and the study design were configured to analyze accuracy of the estimation of cumulative energy expenditure over investigated time frames for every subject, and not to analyze the energy expenditure during specific single physical activity classes.

**Strengths**

The developed AS-MLR for PAEE estimation was validated during a daily military routine and not in a laboratory setting. Therefore, the results are more meaningful for future applied studies. Another strength of the present study is that the method was developed particularly for the military setting. This fact and its combination with activity recognition allow good accuracy in PAEE estimation during the daily military routine of recruits and soldiers.
Conclusion

Established, easy-to-handle body-fixed sensors deliver valid data for PAEE estimation in a military setting. With the discussed sensors and the developed multiple linear regressions, PAEE can be investigated objectively in one-minute intervals over six continuous days during the daily military routine. Additionally, activities’ duration and frequency can be estimated with the algorithm developed in Chapter 6. Therefore, the combination of both methods is ideal for quantifying physical job requirements in terms of the type, duration, frequency and intensity of military-specific physical activities. Further, we conclude that knowing the activity class may provide important information for high precision in sensor-based PAEE estimation.
8. Estimation of gait velocity based on continuous step count data

Introduction

The two previous chapters described the development of algorithms to recognize military-specific activities and to estimate energy expenditure during daily military routines using body-fixed sensors. One of the applied sensors was the uniaxial accelerometer GT1M (ActiGraph LLC, Fort Walton Beach, FL), which registered subjects’ hip accelerations and step frequency (SF). Based on the findings of prior studies, SF is linearly related to gait velocity (GV) over a large range of walking speeds (Chen and Bassett, 2005; Cho et al., 2002; Corder et al., 2007; Sekiya and Nagasaki, 1998; Terrier and Schutz, 2003). It may therefore be possible to derive additional information about daily physical activities from GT1M data by calculating subjects’ walking speed and distance covered on foot using continuously sampled SF.

Body height (BH) respectively leg-length may influence the ratio between GV and SF (Sekiya et al., 1996; Wixted et al., 2007). Additionally, at a slow walking speed, the validity of mechanical and digital step counters is poor (Esliger et al., 2007; Knapik et al., 2007; Melanson et al., 2004). However, GT1M is very accurate at reading step detection at normal walking and running speeds (Esliger et al., 2007).

The aim of this study was to describe the linear regression between BH and SF as an independent variable and GV as a dependent variable for young men and to quantify its accuracy to estimate walking speeds in one-second intervals. Additionally, the GV estimation, based on the new algorithm, was compared to the regression by Terrier and Schutz (2003).
Method

Study design and participants

Thirty, healthy male volunteers (age range of 20-30 years) were recruited from among students of the Swiss Federal Institute of Sports Magglingen. The volunteers had to cover six laps of 290 meters each on a level floor. Individuals chose their walking speed after instruction from the examiner. The examiner instructed the volunteer to conduct the first lap at a very slow walking speed, the second at a slow walking speed, the third at the individual’s preferred walking speed, the fourth at a brisk walking speed, and the fifth at the individual’s preferred aerobic running speed. For the last lap, volunteers were instructed to start the lap at a very slow walking pace and progressively increase their velocity to a running pace. Participants’ BH was measured prior to the walking assessments.

Fifteen, randomly chosen volunteers were allocated to the training-data group, the other 15 volunteers to the control-data group. Data from the training-data group were used to develop the linear regressions between BH and SF as an independent variable, and GV as a dependent variable. Data from the control-data group were used to verify the accuracy of the developed GV-estimation algorithm.

Instruments

The GT1M accelerometer and step counter was used to assess the subjects’ step counts. After programming the GT1M to record step-count data in one-second intervals (1 Hz), it was placed in a belt pouch worn on the waist over the right anterior axillary line. A 200-cm measuring tape was used to assess subjects’ BH to the nearest 0.5 cm. Time per lap was hand stopped by a chronometer (Suunto t3d, Suunto, Vantaa, Finland) and valued with the accuracy of 0.1 s.
**Data analysis**

Assessed data were BH, time per lap, and step-counts per lap. To derive GV and SF per lap, the distance (290 m) and the number of steps were divided by the time (s).

Statistical analyses were performed with SPSS for Windows (version 16.0, SPSS Inc., Chicago, IL) with an alpha level of 0.05 to indicate statistical significance. The data from the training-data group were used to calculate multiple linear regressions with BH and SF as independent variables and with GV as dependent variable.

Using the BH and SF of the control-data group, the GV was estimated for each lap. For that purpose, the regression model developed in the first part of this study, and concurrently the regression of Terrier and Schutz (2003), was applied. For approximate normally distributed data, a paired \( t \)-test was conducted to detect statistical differences between estimated and measured GV-values. Errors of estimated GV were quantified with 95% confidence interval of differences (CI-95%). Further, Bland-Altman plots (Bland and Altman, 1986) were used to visualize agreement between estimated and measured GV.

**Results**

Body height in both study groups did not differ (179.4 ± 6.1 cm and 178.8 ± 8.0 cm, \( p = 0.819 \)) and had no significant influence on GV in the regression model. Therefore, only a linear regression model was applied in the present study as an algorithm for GV-estimation: \( GV = 1.675 \cdot SF - 1.464 \).

The linear regression applied to the data from the control-data group resulted in a correlation of \( r = 0.917 \) between directly measured and estimated GV. However, the Bland-Altman plot in Figure 8.1A shows systematic underestimation of GV for walking speeds below 1 m/s. The misclassifications for low walking speeds resulted in an average underestimation in the total data set of -0.087 m/s (\( p = 0.020, \text{CI-95\%} = -0.159 \text{ to } -0.014 \text{ m/s} \)). In comparison, the underestimation of the GV was even stronger (-0.576 m/s, \( p = 0.000, \text{CI-95\%} = -0.676 \text{ to } -0.477 \text{ m/s} \)) when using the regression of Terrier and Schutz (2003). The Bland-Altman plot in Figure 8.1B shows that with Terrier and Schutz's
regression, GV-estimation is accurate for a slow walking speed. However, with increasing walking speed, a growing underestimation was observed. Thus, a branched model was expected to provide better results than both single methods. Therefore, the regression of Terrier and Schutz (2003) was used for an SF of up to 1.42 steps per second (sps), while the linear regression developed in the present study was used for higher SF. Average gait velocity estimated using the branched model (1.967 ± 0.746 m/s) did not differ from measured GV (1.997 ± 0.773 m/s, p = 0.281, CI-95% = -0.084 to 0.025 m/s) and showed no systematic misclassification in the Bland-Altman plot in Figure 8.1C.

**Figure 8.1A-C:** Bland-Altman plots on gait velocity (GV) of the control-data group (n=15) measured and estimated using A) the linear regression model developed in the present study, B) the regression from Terrier and Schutz (2003), and C) the branched model. The dashed lines represent the mean of the differences between measured and estimated GV values and the limits of agreement (± two standard deviations). The full line represents the zero line.
Discussion

A branched model, using the regression \( GV = 0.705 \cdot SF \) by Terrier and Schutz (2003) for SF up to 1.42 sps (1 m/s) and the regression \( GV = 1.675 \cdot SF - 1.464 \) for higher SF, showed better accuracy than both single methods for estimating GV. Limits of agreement of respective Bland-Altman plot were within \( \pm 0.5 \) m/s, and average estimated GV did not differ from average measured GV. The branched model benefits from the different strengths of the applied linear regressions.

For very slow walking speeds, the accuracy of step-counts assessed with GT1M is reduced (Esliger et al., 2007). This reduction in sensitivity to detect low step forces is most probably due to the manufacturer's filter (0.30 g acceleration threshold) put in place to help discrimination between actual stepping movements and nonambulatory oscillations (Esliger et al., 2007; Le Masurier, 2004). The reduced sensitivity to detect low step forces explains that only walking speeds of 1 m/s and faster showed a linear relation to SF in the present study (Figure 8.1A and B). The linear regression from Terrier and Schutz (2003) is due to a smaller slope (B-value) in the regression formula less sensitive to SF underestimation. However, very slow walking is of low relevance in quantifying physical demands during a daily military routine.

Limitations

The volunteers' body height varied similarly to that of the male group, but less than that of the mixed-gender group of subjects in Sekiya et al. (1996) with 30 male and 34 female subjects (171.2 ± 6.1 cm and 157.4 ± 4.5 cm, respectively). The homogeneity among male participants of the same age may be the reason for the contrast with prior studies (Sekiya et al., 1996; Wixted et al., 2007) showing that BH, or leg-length, influenced the walking ratio (GV/SF).

In this study only young men were investigated. Volunteers' body weight was not assessed. However, prior studies found that body weight and gender did not influence the walking ratio (GV/SF) while age and body height did influence the walking ratio (Sekiya...
and Nagasaki, 1998; Sekiya et al., 1996; Wixted et al., 2007). Consequently, for other age groups these investigations will have to be repeated before application.

Implications

Continuously estimated GV permits calculation of distance covered on foot if the sampling rate of the initial SF assessment is considered. Therefore, estimated GV is divided by the sampling rate of SF (1 Hz in this study) to calculate the distance covered over the specific time interval. Then, the sum of distances calculated for every time-interval results in the distance covered on foot during the investigated measurement period. This application allows the daily distances covered on foot to be assessed.

With the method for estimation of GV and distance covered on foot, the present study allows deriving additional information on daily physical activities and demands from body-wearable sensors chosen in Chapter 6 and 7 for activity recognition and energy expenditure estimation during military service. Especially in the military setting, the distance covered on foot is a relevant outcome to quantify soldiers’ physical activities and demands during their daily routines.

Conclusion

It is possible to accurately estimate GV and therefore calculate distance covered on foot based on SF assessed in one-second time intervals by a digital step counter like GT1M. The linear regression developed in the present study is more appropriate than the one developed by Terrier and Schutz (2003) for preferred walking, fast walking, and running speeds. However, the regression by Terrier and Schutz is less sensitive to GT1M’s SF-underestimation during very slow walking speeds. Therefore, we recommend using both linear regressions, combined in a branched model.
9. Physical activities and demands in boot camps of different Swiss Army occupational specialties

Introduction

During their daily routines, soldiers perform many physically demanding activities. It is important in terms of injury prevention, unit performance, and overall morale that the individual’s physical capabilities correspond to the demands (Bilzon et al., 2005; Gledhill and Jamnik, 1992; Rayson, 1998; Rosendal et al., 2003; Sharp et al., 1998). However, to balance demands and individual capacities, physical job requirements of the occupational specialties within a military organization must first be identified.

Prior studies attempted to identify the most frequently performed physically-demanding activities (Jaenen, 2009; Rayson, 1998). Rayson (1998) showed that 89% of physically-demanding tasks in British Army occupations involved either lifting or carrying loads. The following activities were identified as the key common tasks performed in recent and current North Atlantic Treaty Organization (NATO) missions: walking, marching with a backpack, running, and physically demanding materials-handling including lifting and lowering loads, lifting and carrying loads, and digging (Jaenen, 2009). Other studies quantified physical demands during daily military routines by investigating energy expenditure using the doubly labeled water method (DLW; Castellani et al., 2006; Forbes-Ewan, 2004; Tharion et al., 2004; 2005; Wilkinson et al., 2008). The total energy expenditure (TEE) assessed during daily military routine ranged from 14.1 MJ/d in US support soldiers to 17.2 MJ/d in US special forces, respectively (Tharion et al., 2004). In their metaanalysis, Tharion et al. (2005) summarized TEE values of 16.8 to 17.1 MJ/d in boot camps comparable in duration to the Swiss Army. However, to our knowledge, objectively determined values of duration and intensity of military-specific activities in Swiss Army occupational specialties have not yet been assessed.
In the previous three chapters of the presented thesis, a sensor-based system was developed in order to recognize military-specific, physically demanding activities, and to estimate energy expenditure, as well as distance covered on foot during daily military routine.

The aim of the present study was to quantify the physical demands in boot camps of five selected Swiss Army occupational specialties by applying the developed measurement methods. Physical demands shall be quantified in terms of activity type, frequency, intensity, and duration, using objective methods of data collection.

Method

Participants and measurement protocol

Four investigated Swiss Army training schools (rescue technicians, armored infantry, fusilier infantry, and reconnaissance infantry) were chosen by the criteria of being physically demanding, based on expert appraisal. Additionally, the investigated communications intelligence training school was chosen to represent less physically demanding occupational specialties. In every training school, 50 recruits were randomly chosen to participate. Additionally, 20 recruits were randomly chosen to replace those selected recruits who declined to volunteer or left the boot camp early. The volunteers received comprehensive information on the study. They provided written informed consent for their participation, as approved by the Cantonal Ethics Committee of Bern, Switzerland, and the Swiss Army Sports and Preventions Competence Centre.

Each participant's height, weight, and age were assessed at the start of boot camp. They were asked to wear body-fixed sensors during each day of weeks two, four, eight, and ten of their military service, from the time they woke up in the morning until going to bed in the evening. Since the majority of the trainees in a Swiss Army boot camp do not serve on the weekend, only Monday through Friday were investigated.
**Instruments**

Body height was measured to the nearest 0.1 cm using a stadiometer (Seca model 214, Seca GmbH, Hamburg, Germany), and body weight was measured to the nearest 0.1 kg on a calibrated digital balance (Seca model 877, Seca GmbH, Hamburg, Germany). A heart rate monitor (Suunto Smartbelt, Suunto, Vantaa, Finland), and a combined step and acceleration monitor (GT1M, ActiGraph LLC, Fort Walton Beach, FL) were used to register data of physical activity-related parameters. GT1Ms were wrapped in waterproof plastic and placed in belt pouches worn on the waist over the right anterior axillary line, and on the side strap of the personal backpack. The sensors were programmed to record heart rate, acceleration, and step-count data in two-second intervals so that the data could be gathered over six continuous days.

**Data processing and statistical analysis**

Data from volunteers who provided over 480 minutes of sensor signals per day were included in analyses. For each investigated day of boot camp, the median of volunteers’ distance covered on foot, minutes spent in the military-specific activity classes or in other activities intensity classes and physical activity energy expenditure (PAEE), were calculated.

Heart rate, hip-acceleration, backpack-acceleration, and step frequency data were synchronized for each volunteer by a self-programmed application. Next, they were processed utilizing the algorithms presented in the previous chapters, using Matlab (Matlab 5.3, MathWorks, Natick, Massachusetts).

Synchronized sensor data were first used for recognition of physically demanding, military-specific activities. Activity recognition was performed using a decision tree that was developed and validated in Chapter 6. The decision tree assigned data in one-minute intervals, either to six military-specific activity classes or to an "other activities" class. The assigned military-specific activity classes included: walking, marching with a backpack, lifting and lowering loads, lifting and carrying loads, digging, and running.

Secondly, based on activity recognition and sensor data, PAEE was estimated in one-minute intervals. Therefore activity-class-specific multiple linear regressions, with heart
rate and acceleration as the independent, and PAEE as the dependent variable, were used (see Chapter 7). TEE values were calculated as the sum of PAEE and resting energy expenditure (REE). The REE was determined using anthropometric data applied to the formula for males by Mifflin et al. (1990). To compare TEE with values published in the compendium of physical activities intensity of Ainsworth et al. (1993; 2000), results were additionally expressed as metabolic equivalent (MET) intensity levels. One MET is equivalent to 69.78 J/kg/min (Ainsworth et al., 2000).

Finally, step count data were used to estimate gait velocity and distance covered on foot in two-second intervals. Specifically, a branched algorithm was used to estimate gait velocity. For a step frequency of up to 1.42 steps per second, the regression by Terrier and Schutz (2003) was applied. For higher step frequencies a linear regression was utilized, with step frequency as the independent variable, and gait velocity as the dependent variable (see Chapter 8).

Statistical analyses were performed with SPSS for Windows (version 16.0, SPSS Inc., Chicago, IL) with an alpha level of 0.05 to indicate statistical significance. To compare anthropometric data, PAEE, distance covered on foot, and time spent in physically demanding activity classes (between week days, weeks, and study groups, respectively), one-way analysis of variances (ANOVA) with Tukey post-hoc analysis for multiple comparisons were conducted.

**Results**

On average, volunteers in all study groups were $20.59 \pm 1.15$ years old, $178.80 \pm 6.63$ cm tall, weighed $74.19 \pm 10.36$ kg, and had a body mass index of $23.19 \pm 2.84$ (Table 9.1). Age and anthropometric data of volunteers in the five study groups did not differ.
Table 9.1: Anthropometric data (mean ± standard deviation) of participants in the five study groups.

<table>
<thead>
<tr>
<th></th>
<th>rescue technicians</th>
<th>armored infantry</th>
<th>fusilier infantry</th>
<th>reconnaissance infantry</th>
<th>communications intelligence</th>
</tr>
</thead>
<tbody>
<tr>
<td>age [y]</td>
<td>20.42 ± 0.96</td>
<td>20.60 ± 0.95</td>
<td>20.73 ± 1.18</td>
<td>20.69 ± 1.36</td>
<td>20.37 ± 1.22</td>
</tr>
<tr>
<td>height [cm]</td>
<td>180.00 ± 6.78</td>
<td>176.60 ± 6.86</td>
<td>179.34 ± 6.50</td>
<td>179.49 ± 6.02</td>
<td>179.26 ± 6.81</td>
</tr>
<tr>
<td>weight [kg]</td>
<td>77.21 ± 11.04</td>
<td>73.05 ± 6.80</td>
<td>75.90 ± 13.28</td>
<td>73.98 ± 9.77</td>
<td>70.79 ± 8.32</td>
</tr>
<tr>
<td>BMI</td>
<td>23.76 ± 2.67</td>
<td>23.43 ± 1.92</td>
<td>23.60 ± 3.92</td>
<td>22.92 ± 2.29</td>
<td>22.02 ± 2.20</td>
</tr>
</tbody>
</table>

Note: BMI = body mass index.

For data analysis, 52.2% of volunteers' investigated days could be included. Other data were excluded, either because participants did not wear one or more sensors (61% of data loss), or due to technical problems and mechanical defects (39% of data loss).

Daily PAEE values and distances covered on foot on the weekdays Monday to Thursday were similar (p > 0.49). However, Fridays were less physically demanding than the other weekdays (p < 0.05). Therefore, in the following analyses, data investigated on Fridays were not included.

In the investigated Swiss Army boot camps, the average PAEE was 10.52 ± 2.43 MJ/day, and trainees covered 12.87 ± 3.32 km/day on foot. Recruits spent 61.04 ± 23.29 min/day in marching (thereof 30.42 ± 22.52 min/day with a backpack), 33.12 ± 19.49 min/day in performing physically demanding, materials-handling activities, 36.20 ± 25.23 min/day in running and sports activities, 114.38 ± 46.73 min/day in inactivity, 271.22 ± 70.16 min/day in other activities with a moderate, and 8.13 ± 5.16 min/day with a vigorous intensity (Table 9.2.).

Average estimated PAEE of the physically demanding activity classes were: 22.24 ± 3.30 kJ/min for walking (a), 29.42 ± 7.35 kJ/min for marching with a backpack (b), 26.22 ± 4.01 kJ/min for materials-handling activities (c), and 34.08 ± 13.81 kJ/min for running and sports (d). Thus, TEE intensity levels were: a) 5.30, b) 6.68, c) 6.06, and d) 7.58 MET, respectively.
Table 9.2: Average daily physical activities and demands in Swiss Army boot camps (mean ± standard deviation).

<table>
<thead>
<tr>
<th></th>
<th>rescue technicians</th>
<th>armored infantry</th>
<th>fusilier infantry</th>
<th>reconnaissance infantry</th>
<th>communications intelligence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAEE [MJ/day]</td>
<td>10.46 ± 1.11</td>
<td>13.23 ± 1.03 e</td>
<td>10.30 ± 2.06</td>
<td>10.51 ± 3.16</td>
<td>8.11 ± 1.81 b</td>
</tr>
<tr>
<td>REE [MJ/day]</td>
<td>7.54 ± 0.60</td>
<td>7.27 ± 0.41</td>
<td>7.46 ± 0.64</td>
<td>7.38 ± 0.53</td>
<td>7.25 ± 0.47</td>
</tr>
<tr>
<td>TEE [MJ/day]</td>
<td>18.00 ± 1.11</td>
<td>20.05 ± 1.03 e</td>
<td>17.76 ± 2.06</td>
<td>17.89 ± 3.16</td>
<td>15.36 ± 1.81 b</td>
</tr>
<tr>
<td>distance on foot [km/day]</td>
<td>12.33 ± 1.30</td>
<td>15.61 ± 1.85 e</td>
<td>13.93 ± 1.66 e</td>
<td>14.74 ± 2.77 e</td>
<td>7.72 ± 1.52 b,c,d</td>
</tr>
<tr>
<td>marching [min/day]</td>
<td>64.19 ± 14.54 e</td>
<td>65.50 ± 16.42 e</td>
<td>80.00 ± 8.29 e</td>
<td>69.25 ± 24.80 e</td>
<td>26.25 ± 6.99 a,b,c,d</td>
</tr>
<tr>
<td>running &amp; sports [min/day]</td>
<td>13.50 ± 15.59</td>
<td>36.75 ± 21.70</td>
<td>35.75 ± 28.83</td>
<td>46.75 ± 20.53</td>
<td>48.25 ± 32.18</td>
</tr>
<tr>
<td>PDMH [min/day]</td>
<td>32.85 ± 10.75</td>
<td>60.50 ± 14.27 t,e</td>
<td>22.75 ± 15.24 b</td>
<td>25.75 ± 18.30 b</td>
<td>23.75 ± 13.65 b</td>
</tr>
<tr>
<td>inactivity [min/day]</td>
<td>85.65 ± 13.38 e</td>
<td>71.50 ± 20.70 t,e</td>
<td>97.50 ± 21.93 e</td>
<td>147.00 ± 42.37 b</td>
<td>170.25 ± 37.83 a,b,c</td>
</tr>
<tr>
<td>moderate OA [min/day]</td>
<td>293.36 ± 40.06</td>
<td>346.25 ± 21.88</td>
<td>229.25 ± 76.36</td>
<td>262.00 ± 84.32</td>
<td>225.25 ± 52.09</td>
</tr>
<tr>
<td>vigorous OA [min/day]</td>
<td>6.38 ± 1.89 b</td>
<td>16.25 ± 3.59 t,e</td>
<td>5.5 ± 2.38 b</td>
<td>6.75 ± 4.50 b</td>
<td>5.75 ± 3.86 b</td>
</tr>
<tr>
<td>sleep [h/night]</td>
<td>7.76 ± 0.62</td>
<td>6.54 ± 1.53</td>
<td>8.00 ± 0.55</td>
<td>6.17 ± 1.09</td>
<td>6.84 ± 1.78</td>
</tr>
</tbody>
</table>

Note: PAEE: physical activity energy expenditure, REE: resting energy expenditure, TEE: total energy expenditure, PDMH: physically demanding materials-handling, OA: other activities. a, b, c, d, and e, respectively: significantly different (p < 0.05) from: rescue technicians (a), armored infantry (b), fusilier infantry (c), reconnaissance infantry (d), or communications intelligence (d), respectively.

In most investigated boot camps, physical demands decreased by trend from week two to week eight of basic training (Figure 9.1 and 9.2). Average PAEE values over all five of the study groups significantly decreased (approximately 25%) from week two to week eight, with a difference of -2.93 ± 3.01 MJ/day.
**Figure 9.1:** Physical activity energy expenditure (PAEE, mean ± standard deviation) during daily military routine in weeks 2, 4, 8, and 10 in 5 boot camps of the Swiss Army. Average PAEE values over all five of the study groups significantly decreased from week 2 to week 8 (p < 0.05).

**Figure 9.2:** Daily distance covered on foot (mean ± standard deviation) during military routine in weeks 2, 4, 8, and 10 of military service in 5 boot camps of the Swiss Army.
Discussion

These findings demonstrate that the physical demands in different boot camps of the Swiss Army are heterogeneous in terms of type, duration, and intensity. As expected, recruits in the communication intelligence training school perform less physically demanding activities during daily military routines than recruits in the other training schools included in our study (Table 9.2). This may be due to the focus on more technical and cognitive content during their formation. The most physically demanding training school was investigated in the armored infantry. While PAEE in the training school of rescue technicians, fusilier infantry, and reconnaissance infantry were similar, the rescue technicians spent by trend more time with physically demanding, materials-handling activities, but covered shorter distances on foot.

Energy expenditure

The total energy expenditure assessed in the US armed forces by DLW ranges from moderate to high values during daily military routines (14.1 MJ/d in support soldiers to 17.2 MJ/d in special forces; Tharion et al., 2004), to very high values during specific field training exercises (25.7 MJ/d in marines; Castellani et al., 2006). In their metaanalysis, Tharion et al. (2005) summarized TEE values of 16.8 to 17.1 MJ/d in 60-day boot camps, which is comparable to the Swiss Army. Similar data were found in training schools for Australian infantry and British parachute recruits (17 and 18.3 MJ/d, respectively; Forbes-Ewan, 2004; Wilkinson et al., 2008). Compared to these studies, the TEE for the Swiss Army rescue technicians (18.0 MJ/d), fusilier infantry (17.8 MJ/d), and reconnaissance infantry (17.8 MJ/d) training school were similar. The TEE in the Swiss Army armored infantry training school (20.5 MJ/d) was in the upper range while TEE in communications intelligence (15.4 MJ/d) was in the lower range of values found in comparative studies of other nations (Table 9.2).

During basic training in the Swiss Army, the TEE values are in the same range as those found in professional sportsmen of similar age and anthropometry engaged in their daily routines (Ebene et al., 2002; Vogt et al., 2005). Professional soccer players, for example,
have been found to exhibit an average TEE of 14.8 MJ/d (Ebine et al., 2002), and elite cyclists during pre-season training have an average TEE of 19.1 MJ/d (Vogt et al., 2005). This comparison clarifies that physical demands in boot camps are high. However, in contrast to professional athletes, recruits often do not have a specific physical preparation and adaptation to high physical demands.

Average TEE intensity levels that were estimated during specific activity classes in the present study correspond well to values published in the compendium of physical activities intensity of Ainsworth et al. (1993; 2000). This indicates strong validity of the method for energy expenditure estimation developed in Chapter 7.

**Distance covered on foot**

Recruits in the Swiss Army armored infantry boot camp covered higher distances on foot (15.61 km/d) during daily military routines than trainees in the US Army basic training (11.7 km/d, range among ten investigated battalions: 9.7 to 14.0 km/d; Knapik et al., 2007). The distance covered on foot by Swiss Army rescue technicians, fusilier, and reconnaissance infantry recruits (12.3 to 14.5 km/d) are in the upper range of the comparative data of trainees in the US Army basic training. Recruits in the communication intelligence training school covered shorter distances (7.7 km/d) than trainees in the US Army basic training (Table 9.2 in the present study and Table 5 in Knapik et al., 2007, p. 109).

Male, US citizens of similar ages (i.e., 19 years) perform an average of 11'660 steps per day (Tudor-Locke et al., 2010). Using an average step length assessed among young adults of 0.73 m (Menz et al., 2003), a total daily distance of 8.55 km/d results. Therefore, we conclude that trainees in most occupational specialties of the armed forces walk more than comparative civilians in their daily routines.

**Development of physical demands during basic training**

It is well established in both sports and exercise sciences that subjects shall gradually adapt to new kinds of physical demands. This is particularly true for recruits in the first
weeks of basic training. However, the present study observed a decrease, rather than a progression, in physical demands for the first eight weeks of basic training (Figure 9.1 and 9.2). The same pattern of declining physical demands for the first nine weeks of basic training in British Army Parachute recruits (Wilkinson et al., 2008) and in South African Army recruits (Jordaan and Schwellnus, 1994) was related to an enhanced number of injuries. Therefore, we must find ways to introduce a more progressive development of physical demands in Swiss Army training.

Limitations

Sensor defects and participants' failure to wear all three sensors reduced the data set by 47.8%. The number of complete data sets may be increased by any of the following: a smaller number of sensors per participant, a daily surveillance of participants' involvement, or an improvement of the sensors' mechanical stability. The investigated weeks of boot camp (weeks 2, 4, 8, and 10) represent only a part of the basic training in the Swiss Army training schools. It is possible that the weeks of basic training that were not investigated contained different patterns of physical demands. However, the content of school commanders' appraisals and weekly military training plans do not support this hypothesis.

Strengths

The present study objectively describes the physical demands involved during daily military routines in military training schools by quantifying type, intensity, and duration of physical activities. Furthermore, the method employed was unobtrusive, and did not restrain the daily military training.

Implications

During the Swiss Army recruitment process, physical, psychological, psychosocial, and cognitive capabilities, as well as job experiences and personal interests, are assessed. These characteristics are then compared to the respective requirements of different occupational
specialties. It is therefore crucial to have a differentiated description of job requirements. Today's physical job requirements are based solely on an expert's appraisal. Therefore, with this study, new and objective information was generated that can be applied in the description of physical job requirements.

This study provides novel reference data to quantify the physical demands of daily military routines. It completes already published values of energy expenditure during daily military routines in the armed forces of many nations. Further, the present study is one of the first to quantify distances covered on foot. Moreover, to our knowledge, it is the first to objectively quantify the time spent in military-specific activities.

Conclusion

The new, objective measurement system utilized in the present study yielded data comparable to prior studies that have applied alternative methods. Nevertheless, the sensor-based, objective measurements system provides more information on daily physical activities and demands than the traditional, single methods.

Physical demands in Swiss Army boot camps vary among occupational specialties. However, the average daily TEE values in all investigated occupational specialties are in the range of values found in other nations' armed forces, as well as for professional sportsmen. We demonstrated that physical demands in the Swiss Army decreased between the second and the eighth week of basic training. We conclude that the actual development of physical demands during first eight weeks of military training does not meet the training principle of progressive overload recognized in sport sciences.
10. General discussion

In the present thesis, the research fields of physical fitness, injuries, and physical demands in a military setting were investigated (see Figure 2.1 in Chapter 2). Prior studies showed that those fields are related to each other: In a military setting, a mismatch between physical fitness and physical job requirements can increase the risk of injury and jeopardize unit military performance and overall moral (Gledhill and Jamnik, 1992; Rayson, 1998; Rosendal et al., 2003; Sharp et al., 1998). Therefore, the importance of obtaining an accurate description of the job requirements in physically demanding military occupations cannot be overestimated (Gledhill and Jamnik, 1992; Rayson, 1998).

Commonly used procedures to assess military or fire fighting specific job requirements include self-report questionnaires (Pope et al., 1998; Rayson, 1998), observations (Bos et al., 2004; Dybel and Seymour, 1997; Jones et al., 1993; Rayson, 1998), interviews (Rayson, 1998), and measurements of physical demands such as energy expenditure with the doubly labeled water (DLW) method (Bilzon et al., 2005; DeLany et al., 1989; Tharion et al., 2004; Wilkinson et al., 2008) or step counts (Knapik et al., 2007). These methods all have inherent limitations. While self-report questionnaires are of low objectivity, observations are not feasible in large study groups, and DLW as well as step counts yield no information on activity characteristics or on their progression across the day. However, the ideal method to objectively assess physical demands during daily military routines is not yet available. Therefore, in the present thesis, new methods were developed and existing methods were adapted to the military setting. Those methods do meet the propositions of Almeida et al. (1999) and Bos et al. (2002), who claim that future efforts to assess job requirements should include a more precise quantification of physical demands in terms of activity type, frequency, intensity, and duration.

Further, in the context of the present work, a new, feasible, and reliable fitness-test battery was assembled and validated. The discriminative power of its parameters to predict overuse injuries during military service was demonstrated.
Physical demands on soldiers

The sensor-based activity recognizing system developed in Chapter 6 achieved an overall recognition rate similar to findings in prior studies conducted in other settings (Aminian et al., 1999; Pärkkä et al., 2004; Pober et al., 2006). Based on the activity recognizing system, activity class-specific multiple linear regressions with hip acceleration and heart rate as independent variables have been developed to estimate energy expenditure in one-minute intervals. While this approach showed to be appropriate, prior published algorithms developed by Brage et al. (2004) and Swartz et al. (2000) underestimated energy expenditure during daily military routines. This may be due to the fact that their algorithms were developed in different settings. However, we conclude that using setting-specific algorithms and including information on type of recent physical activity may increase precision in sensor-based energy expenditure estimation.

Average total energy expenditure (TEE) and distance covered on foot in Swiss Army boot camps are similar to values found in comparative boot camps of the armed forces of other nations (Forbes-Ewan, 2004; Knapik et al., 2007; Tharion et al., 2005; Wilkinson et al., 2008). Values are also in the range of TEE found during daily routines of professional sportsmen of similar age and anthropometry (Ebine et al., 2002; Vogt et al., 2005). This comparison clarifies that physical demands in boot camps are high. However, in contrast to professional athletes, recruits are often physically unprepared to high physical demands. That may increase the risk for overuse health complaints.

In the investigated occupational specialties, physical demands did not increase progressively with time in boot camps. TEE values even decreased significantly from the second to the eighth week of boot camp. The same pattern of declining physical demands in the first nine weeks of basic training was observed in British Army Parachute recruits (Wilkinson et al., 2008) and in South African Army recruits (Jordaan and Schwellnus, 1994) as well. The authors showed a relation between a declining development of physical demands and enhanced number of injuries.

Since the investigated boot camps and volunteers among whom injuries and physical demands were assessed were the same in the present thesis, outcomes of those studies can be related. Even if the number of investigated boot camps is too small for relevant
statistical analyses, a relation between the degree of decrement in physical demands from week two to week eight (Figure 7.1 and 7.2 in Chapter 7) and injury incidences (Table 4.3 in Chapter 4) is indicated. The strongest decrement in physical activity energy expenditure and distance covered on foot during the first eight weeks of basic training was investigated in reconnaissance infantry with the highest overuse injury incidence proportion of 43%, followed by fusilier infantry with an overuse injury incidence proportion of 40%. In armoured infantry with approximately constant physical demands in the first eight weeks of military service, an overuse injury incidence of 38% was investigated. In rescue technicians, a trend indicating an increase in daily distance covered on foot was investigated in the first weeks of basic training. Compared to the other three study groups, in rescue technicians, a significantly lower overuse injury incidence proportion of 23% was observed. Further studies are needed to investigate the relation between the development of physical demands in the first weeks of Swiss Army basic training and overuse injuries.

Physical fitness tests

Most test batteries of the military organizations of different nations contain a running test to assess aerobic endurance capacity and sit-ups to measure abdominal muscle fitness (NATO, 1997). Some contain further tests to assess muscle power, like push-ups, hand grip tests, standing long jumps, and others (NATO, 1997). The fitness-test battery assembled for the present thesis consists of similar fitness parameters and includes a test of balance ability. The following five tests are included in the battery: Progressive endurance run (PER), trunk muscle strength test (TMS), one-leg standing test (OLS), standing long jump (SLJ), and seated shot-put (SSP). To assess aerobic endurance and trunk muscle fitness, the test protocols chosen differed from those of the armed forces of other nations.

In 12-minute and distance running tests, inexperienced runners have difficulties finding the optimal speed and may therefore be underestimated (Jorgensen et al., 2009). The PER may thus be more appropriate to investigate aerobic endurance in heterogeneous groups of conscripts. The paced velocity allows individuals with low physical fitness or without experience in self-pacing to achieve their individual maximal aerobic endurance performance. However, further studies are needed to prove this assumption. Consistent
with previous studies, demonstrating a strong relation between aerobic endurance and injury risk during military service (Blacker et al., 2008; Gilchrist et al., 2000; Jones and Knapik, 1999), PER showed strong discriminative power to predict overuse injury risk.

Less consistent and less significant associations between other physical fitness measures such as push-ups or sit-ups and risk of injury were found in previous studies (Blacker et al., 2008; Jones et al., 1993; Jones et al., 1993; Knapik et al., 1993; Reynolds et al., 1994). Divergent to those studies, our data suggests that the TMS is of strong power to predict injury risk during military service. Conducting TMS in larger groups is as feasible as sit-ups and curl-ups. Its test-retest reliability expressed as a Pearson correlation coefficient was found to be $r = 0.77$ to $0.87$ (Bourban et al., 2001; Tschopp et al., 2001; Wyss et al., 2007). These values are in the upper range of results reported for sit-ups and curl-ups with $r < 0.50$ (Sparling et al., 1997; Suni et al., 1996) and $r = 0.72$ to $0.84$ (DiNucci et al., 1990; Erbaugh, 1990; Tsigilis et al., 2002). Based on the discussed strengths, we conclude that the TMS may outclass sit-ups and curl-ups as a test to assess the physical capacities of military occupational selections.

A negative association between balance ability and physical training-related injuries has been shown in college students (Hrysomallis, 2007). The present thesis confirms that finding for two out of four investigated Swiss Army occupational specialties.

SLJ and SSP have high feasibility, reliability, and concurrent-validity (Cronin and Owen, 2004; Markovic et al., 2004; Mayhew et al., 1991; Tsigilis et al., 2002; Wyss et al., 2007), but they showed only weak discriminative power to predict overuse injuries. However, our pre-test data (not shown) and results from Williams and Wilkinson (2007) indicate that the tests for muscle power may have strong discriminative power to predict military performance assessed by staff reports or task performances (box-lifting), respectively.

**Injuries in soldiers**

The injury incidence rates in three out of four investigated boot camps in the present thesis are in the upper range of values published in meta-analyses of previous review studies (Kaufman et al., 2000; National Research Council, 2006). Consistent with other
studies, the lower extremities are most often affected (Almeida et al., 1999; Heir and Glomsaker, 1996; Knapik et al., 1993). The knee is the anatomical site most often injured in Swiss Army military service and is in the upper range of data in comparative studies (Almeida et al., 1999; Jones et al., 1993; Knapik et al., 1993; Popovich et al., 2000). Numbers of reported injuries concerning the back and Achilles tendon were higher in Swiss Army recruits than in US Army trainees (Jones et al., 1993; Knapik et al., 1993; Popovich et al., 2000). Otherwise, injuries to the foot and ankle are less often reported in the Swiss Army than in comparative studies (Almeida et al., 1999; Jones et al., 1993; Knapik et al., 1993; Popovich et al., 2000).

For the investigated occupational specialties, different physical fitness parameters had descriptive power to predict overuse injuries during military service. For recruits of the rescue technician school, for example, trunk muscle strength and balance ability were proven relevant to injury prediction. Conscripts with performances at the recruitment of at least 130 seconds in TMS and 37 seconds in OLS had an overuse injury incidence of 11% during the first 18 weeks of military service. Those who performed below one or both of these minimum physical fitness standards had an overuse injury incidence of 28%.

However, the efficacy of such selection standards on injury prevention has to be quantified first in a further study with a new sample of subjects. Minimum physical fitness standards applied in the future need to be selective enough to reduce injury incidences, days off duty, and discharge numbers but still low enough to find sufficient conscripts fulfilling the requirements. Solutions have to be found for the conscripts of the compulsory Swiss Army who do not fulfill the minimal standard of any physically demanding occupational specialty. They could be assigned to jobs with low physical demands, or they could be preconditioned prior to military service. Pre-training of subjects with a low physical fitness level has been shown to be successful to reduce injury and discharge rates among US and Singapore Army recruits (Knapik et al., 2006; Lee et al., 1997).

Limitations

Due to limited resources, the investigation of military performance is missing in the present thesis (Figure 2.1 in Chapter 2). The descriptive power of physical fitness test
parameters to predict individuals’ military performance shall be investigated in the future. The challenge will be to adequately assess individuals’ military performance. We suggest using staff reports and questionnaires answered by the closest superior officers, as applied by Dyrstad et al. (2010). Further, peer evaluations, self-evaluations, simulated battlefield tasks (Harman et al., 2008; Popper, 1997; Stevenson et al., 1992), assessments of metabolic biomarkers (Nindl, 2009), and data of total unit performances might provide useful information to quantify individual military performance. Our pre-tests indicated that of all investigated physical fitness parameters, the explosive power of upper extremities may have the strongest relation to the individuals’ marks in staff reports.

**Implication**

We suggest investigating all relevant occupational specialties in the Swiss Army with the methods developed and adapted in the present thesis. Further, interventions shall be accomplished and evaluated to balance physical demands and trainees’ physical capabilities. The potential to reduce the total number of injuries in the Swiss Army exists. Further studies will show if morale and military performance can be increased as well when an optimal balance between job requirements and physical capabilities is achieved.

**Main aims**

The three main aims of this thesis were attained:

- With the study presented in Chapter 4, a reliable and feasible fitness-test battery to assess physical capabilities in larger groups of young men was assembled.
- The results in Chapter 5 demonstrated that the Swiss Army physical fitness-test battery is a valid instrument to predict risk of injuries in physically demanding military service. For the five investigated occupational specialties, different test parameters turned out to be relevant for injury prediction. TMS and PER were found to have the strongest discriminative power to predict overuse injuries.
In Chapters 6 to 8, new military-specific methods to objectively assess physical demands in terms of activity type, intensity, duration, and frequency were presented. The methods are based on body-fixed sensors. They were applied in Chapter 9 to five different occupational specialties to quantify their daily physical demands.

Conclusion

We conclude that the fitness-test battery developed in the present thesis is appropriate to detect conscripts with enhanced risk of overuse injuries in physically demanding military occupational specialties.

With the described sensors and the developed algorithms, military-specific activities can be recognized and energy expenditure and distance covered on foot can be estimated in one-minute intervals over six continuous days. Therefore, the presented methods allow investigators to objectively assess type, occurrence, duration, frequency, and intensity of military-specific physical activities.

Physical job requirements in Swiss Army boot camps differ between occupational specialties. Nevertheless, physical demands during daily routines in all investigated Swiss Army occupational specialties are in the range of values assessed in the basic training of other nations and of daily routines of professional sportsmen. An important difference between army recruits and elite sportsmen is recruits’ lack of adequate preparation for these high daily physical demands. We conclude that the decrease in daily energy expenditure during the first eight weeks of military training does not meet the training principles of sport sciences and may therefore lead to an inflated number of overuse health complaints.
11. References


Dybel GJ and Seymour CJ. Identifying the physical demands of Army Reserve personnel during deployable medical systems training. *Mil Med*. 1997; 162: 537-42.


## 12. Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-MRT</td>
<td>12-minutes running test</td>
</tr>
<tr>
<td>95%-CI</td>
<td>95% confidence interval</td>
</tr>
<tr>
<td>ACC</td>
<td>acceleration</td>
</tr>
<tr>
<td>AFEB</td>
<td>US Armed Forces Epidemiological Board</td>
</tr>
<tr>
<td>ANOVA</td>
<td>analysis of variances</td>
</tr>
<tr>
<td>AS-MLR</td>
<td>activity-class-specific multiple linear regressions</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>BP-ACC</td>
<td>backpack acceleration</td>
</tr>
<tr>
<td>CMJ</td>
<td>counter movement jump</td>
</tr>
<tr>
<td>DLW</td>
<td>doubly labeled water method</td>
</tr>
<tr>
<td>EEE</td>
<td>energy expenditure estimation</td>
</tr>
<tr>
<td>H-ACC</td>
<td>hip acceleration</td>
</tr>
<tr>
<td>HR</td>
<td>heart rate</td>
</tr>
<tr>
<td>HRaR</td>
<td>heart rate above resting heart rate</td>
</tr>
<tr>
<td>IMFIS</td>
<td>injury related minimum physical fitness standard</td>
</tr>
<tr>
<td>MST</td>
<td>multistage 20-m shuttle run test</td>
</tr>
<tr>
<td>NATO</td>
<td>North Atlantic Treaty Organization</td>
</tr>
<tr>
<td>OLS</td>
<td>one-leg standing test</td>
</tr>
<tr>
<td>OUI</td>
<td>overuse injury</td>
</tr>
<tr>
<td>PC</td>
<td>pole climbing</td>
</tr>
<tr>
<td>PER</td>
<td>progressive endurance run</td>
</tr>
<tr>
<td>PS</td>
<td>pendulum sprint</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic analysis</td>
</tr>
<tr>
<td>RR</td>
<td>relative risk</td>
</tr>
<tr>
<td>SF</td>
<td>step frequency</td>
</tr>
<tr>
<td>SJ</td>
<td>squat jump</td>
</tr>
<tr>
<td>SLJ</td>
<td>standing long jump</td>
</tr>
<tr>
<td>SSP</td>
<td>seated shot put</td>
</tr>
<tr>
<td>TMS</td>
<td>trunk muscle strength test</td>
</tr>
<tr>
<td>VO\textsubscript{max}</td>
<td>maximal oxygen consumption</td>
</tr>
<tr>
<td>WB</td>
<td>walking on the beam</td>
</tr>
</tbody>
</table>
13. Curriculum vitae

Thomas Wyss
Käppelistrasse 53, 4600 Olten, +41 62 296 46 17, wysst@gmx.net

16.08.1977; Kappel SO, Switzerland; married

Educations

2007 - ETH Zurich: PhD in "human movement and sport sciences".
2004 (Oct. 99 – June 04) ETH Zurich: MSc in "human movement and sport sciences".
2001 (Oct. 97 – Oct. 01) ETH Zurich: Studies in "physical education and sports"
(Eidgenössisches Turn- und Sportlehrerdiplom II).
1997 (February 1997) High School, Olten: Matura Typ C.

Professional experiences

2007 - (since April 2007) Institute of Human Movement Sciences and Sport, ETH Zurich: External doctoral thesis "Physical activities and demands in Swiss soldiers".
2004 - (since Juli 2004) Federal Office of Sport FOSPO, Swiss Federal Institute of Sports Magglingen SFISM, Physical Activity and Health Branch: Research assistant in the field of "physical activity and health". Lecturer in "mathematical and natural science", and since 2008 in "scientific working".
2003/2004 (20.10.03 – 30.06.04) University Clinic Balgrist, Zurich - Center for paraplegia: Internship as research assistant including master thesis (in the field of gait analysis).
2001–2003 (20.08.01 – 21.08.03) Secondary School I, Illnau-Effretikon: Teacher in physical education and sports (30%).
1998–2001 (01.11.98 – 28.02.01) Institute for Biomechanics, ETH Zurich: Junior Research Assistant.
1998–2001 Various internships and representations as teacher in physical education and sports.
1997 (03.09.97 - 19.10.97) SEGA, banking sector, Olten: Controlling in the context of anonymous funds.
**Publications**

<table>
<thead>
<tr>
<th>Year</th>
<th>Type</th>
<th>Title</th>
</tr>
</thead>
</table>