Establishing the Test-Retest Reliability and Concurrent Validity of the Repeat Ice Skating Test in Adolescent Boys Ice Hockey Players

Allan Power

A Thesis Submitted to:
The School of Kinesiology
Lakehead University
Thunder Bay, Ontario

August, 2009

TABLE OF CONTENTS

ABSTRACT

This study examined the test-retest reliability and concurrent validity of the Repeat Ice Skating Test (RIST). This was an on-ice field anaerobic test that measured average peak power. It was validated with 3 anaerobic lab tests: Vertical Jump, Margaria-Kalamen Stair test, and the Wingate cycle ergometer test. The study's subjects were 14 males aged 11-12 on a Peewee "A" level ice hockey team in the Thunder Bay area. Results revealed that the RIST was a reliable test at measuring average peak power in watts (r = .99) and watts per kilogram (r = .98). The RIST was also a valid test when correlated with the 3 anaerobic lab tests for measuring peak power in watts: Vertical Jump (r = .86), Margaria-Kalamen Stair test (r = .71) and Wingate cycle ergometer test (r = .85). These correlations were much lower for watts per kilogram (w/kg): Vertical Jump (r = .50), Margaria-Kalamen Stair test (r = .40) and Wingate cycle ergometer test (r = .18). The RIST will be useful to coaches and teams for various reasons, including monitoring anaerobic power and velocity of players over a season, determining strengths and weaknesses of players and if training programs are required. The test may help as an evaluation tool for picking players on a team, may determine which players get more ice time, test the players work effort over various shifts of time, be used as a comparison measure for other anaerobic tests, and the continued use of the test will also help develop normative data. In summary, it is of the author's opinion that this test will save time, money, will not require any specific equipment, and is a sport-specific on-ice test that can be administered during a regular hockey practice.

ACKNOWLEDGEMENTS

The completion of this study has been made possible with the help and support of many people. I wish to express my gratitude and thanks to:

Dr. William Montelpare, my advisor, who was always available, and whose guidance helped me achieve this goal.

My committee members, Dr. Brent Faught, and Eryk Prysucha, whose attention to detail and advice helped me tremendously.

The professors and staff at the School of Kinesiology, with their knowledge, helped me through this process at Lakehead University.

The Thunder Bay "AAA" Bantam Kings and "A" Peewee Hawks ice hockey teams who were great to work with and gave up their time to participate in this study.

My parents, for all their love and support.

Jillian Morgan, for her patience and encouragement throughout this process.

LIST OF TABLES

LIST OF FIGURES

DEFINITIONS

Anaerobic Energy – The short-term energy systems lasting from less than 10 seconds up to 2 minutes with ATP-PC (adenosine triphosphate - phosphocreatine) and Glycolysis providing energy (Powers and Howley, 2004).

Anaerobic Average Peak Power – A term designed specifically for this study involving the RIST. It is the measure of peak power over the length of time of one sprint during the RIST. Since the actual "peak power" cannot be pinpointed during one sprint, it will be calculated as the average peak power over the time of one whole sprint.

Anaerobic Capacity – Total energy output during maximum effort lasting anywhere from 10 to 60 seconds (Bar-Or, 1996).

Anaerobic Peak Power – Total energy output during maximum effort lasting anywhere from 1 to 10 seconds (Bar-Or, 1996).

Concurrent Validity – Relationship between a test and a criterion test when the two tests are measured at the same time (Morrow, Jackson, Disch, & Mood, 2005).

Intra-class Correlation Coefficient (ICC) – Is the correlation or consistency between two or more measures based on analysis of variance (ANOVA) (Morrow et al, 2005). The reliability of a test or test score is shown by the size of the coefficient that ranges from 0 (no correlation) to +1 (perfect correlation) (Looney, 2000).

Pearson product-moment Correlation Coefficient (PPM) – Is the measure of the co-variation between 2 variables indicating the magnitude (amount) and direction (- / +) of the relationship. The coefficient ranges from –1 (perfect negative correlation) to 0 (no correlation) to +1 (perfect positive correlation) (Morrow et al, 2005). This method will be used to examine concurrent validity.

Random Error – An error in research that may randomly affect the outcome of one measurement in a sample such as motivation or injury but does not affect the entire sample.

Reliability – The repeatability or consistency of a set of results under similar conditions or times: that is, measuring to see if one person can score relatively the same each time on the same test (Lyman, 1998).

Stability Reliability – Is the consistency of measures across time (days, weeks) by giving a test on two occasions to the same group and seeing if the two test scores are relatively the same (Morrow et al, 2005).

Systematic Error – Factors that systematically affect the measurement of a variable across a sample. Sources of systematic error include, equipment, environment, and observation and can affect all scores in the sample.

True Score Theory – It is a psychometric theory of measurement. It assumes that every score is the sum of two components: true scores and random error and that each person would obtain a true score if no measurement errors occurred.

X (observed score) = T (true score) + E (error of measurement)

Validity – The extent of a test's ability to accurately measure what it has been designed to measure (Lyman, 1998).

INTRODUCTION

The sport of ice hockey requires high-intensity skating, frequent body contact, and sudden changes in speed and skating direction (Montgomery, 1988). Ice hockey involves the use of both the aerobic and anaerobic energy systems. The anaerobic system is needed for the high intensity bursts or sprints during each shift while the aerobic system is needed to recover quickly from each shift (Montgomery, 1988). A study by Green, Bishop, Houston, Mckillop, Norman, and Stothart (1976) assessed performance changes in the University of Waterloo Men's hockey team. The results showed the average ice hockey shift ranged from 60-80 seconds with 2-3 stoppages. This makes the average time of a continuous shift roughly 39.7 seconds which is in the anaerobic energy systems time range of zero to sixty seconds.

Another study by Green (1978) found lower lactate levels in players who performed shorter shifts (mean range = 2.9 to 5.5 mmol/L). One reason for this is that during a shift, there are 2 or 3 stoppages in play. These stoppages provide enough time for 60 to 65% of the phosphocreatine (PC) to be re-synthesized for the player's next shift (Green, 1979). This implies that shorter shifts are better for the hockey player since this reduces the lactate build-up by allowing triphosphate-phospohocreatine (ATP-PC) to restore in muscle (Green, 1979). During this recovery period, myoglobin stores are reloaded and PC is re-synthesized. This leads to a larger amount of phosphocreatine and oxidative phosphorylation (OP) to ATP turnover, reducing the amount of aerobic glycolysis and conserving the glycogen reserves (Montgomery, 1988).

There have been numerous laboratory and field tests to estimate aerobic and anaerobic measures in ice hockey players. Laboratory tests that have been completed in past studies to measure anaerobic fitness include the: *Margaria-Kalamen Stair test* (Houston and Green, 1976; Reed, Hansen, Cotton, Gauthier, Jétte, Thoden, and Wenger 1979; Vainikka, Rahkila, and Rusko, 1982), *Treadmill Run* (Houston et al, 1976; Rhodes, Cox, and Quinney, 1986), *Vertical Jump* (Bracko and Fellingham, 1997; Burr, Jamnik, Dogra, and Gledhill, 2007; Diakoumis and Bracko, 1998; Geithner, Lee, and Bracko, 2006; Mascaro, Seaver, and Swanson, 1992; Vescovi, Murray, and Vanheest, 2006a), *Intermittent cycle test* (Gamble and Montgomery, 1986) and the *Wingate cycle ergometer test* (Burr, Jamnik, Baker, Macpherson, Gledhill, and Mcguire, 2008; Quinney, Dewart, Game, Snydmiller, Warburton, and Bell 2008; Rhodes et al, 1986; Smith, Quinney, and Steadward, 1982; Twist and Rhodes, 1993; Vainikka et al, 1982; Vescovi, Murray, Fiala, and Vanheest 2006b).

Nevertheless, for anaerobic fitness assessments to be sport specific for ice hockey, tests must be designed for completion on-ice. The Repeat Ice Skating Test (RIST) was developed by the author for use on-ice to test anaerobic average peak power in ice hockey players. The test consists of 6 sprints in total over a distance of 49 meters per sprint with the fastest time in seconds used to calculate average peak power in watts. The intent of this study was to establish the test-retest reliability and concurrent validity of the RIST. The reliability was completed by measuring repeated trials of the RIST while the concurrent validity was calculated by comparing the RIST to three reliable and valid "gold standard" lab tests: Vertical Jump, Margaria-Kalamen Stair, and the Wingate cycle ergometer test.

The reason these three tests were used is that they are the most widely used anaerobic tests in the literature for ice hockey as well as being reliable and valid tests for measuring anaerobic power. All three tests are related to the RIST physiologically in that they all test anaerobic power of the larger leg muscles. The Vertical Jump tests the large thigh muscles in a squatting jump motion and has been proven in the literature to be significant in estimating anaerobic power in ice hockey players. The MK Stair test requires running up stairs which involve the same muscles as the RIST. This test is similar to the RIST in terms of a sprinting component since there is a running sprint before jumping up the set of stairs. However, the connection of this test to the RIST might be lesser since leg length or any number of factors could affect scores on this test. Lastly, the Wingate requires cycling against resistance that uses the same muscles for skating in the RIST. Since the participant is sitting during the Wingate and are cycling against resistance, this test might not correlate as well as the other two.

The literature currently identifies three validated field anaerobic tests that have been performed on-ice. These include the: *Reed Repeat Sprint Skate (RRSS)* (Bracko, 2001; Reed et al, 1979; Rhodes, Mosher, and Potts, 1985; Smith et al, 1982) *Sargeant Anaerobic Skate (SAS)* (Watson and Sargeant, 1986), and the *18.3-Meter Sprint test* (Lariviere, Lavallee, and Shepard, 1976). From these tests, the RRSS is the most reliable and valid test but is long in duration and does not fully complement the average duration of a hockey player's regular shift. The RIST was intended to replicate the length and intensity of a regular shift in an ice hockey game, more so than the RRSS. The RIST was also designed to assess anaerobic average peak power in an efficient and practical manner in hockey players.

PURPOSE

The purpose of this study was to establish the test-retest reliability and concurrent validity of the Repeat Ice Skating Test (RIST) and 3 anaerobic lab tests in estimating avaerage peak power and peak power in adolescent (11-12 years old) boys ice hockey players.

HYPOTHESES

For this study, it was expected that the RIST and the 3 anaerobic lab tests would be reliable tests that produce stable measures of average peak power and peak power over time and to have a test-retest reliability coefficient of (r = .70) or higher. Furthermore, it was expected that players who score high on the RIST would also score high on the 3 validated lab tests measuring peak power in watts and w/kg. Specifically, the Vertical Jump should correlate the strongest with the RIST, followed by the MK Stair and then the Wingate.

REVIEW OF LITERATURE

Off-Ice / Laboratory Measures for Ice Hockey

Vertical Jump

The first lab test used was the Vertical Jump. In 1921, Sargent came up with the first test to assess muscle power. This was called the Sargent Jump and was performed by crouching down, followed by a forward arm swing to jump as high as possible with arms and fingers extended. The jump height would be compared to the standing height which is basically the subject standing with arms and fingers extended and is measured by having the subject holding a piece of chalk or tape in their fingers

(Bar-Or, 1996). Some new advances for the vertical jump test include a vertical axis containing a series of rotating plastic sticks when touched. This eliminates other errors of dropping chalk or tape not sticking to the wall when measuring jump height (Chandler and Brown, 2008). The participant usually performs 3 jumps and the best or average score from the 3 jumps is taken as the test measure. There are many different protocols, most of which differ in terms of allowing a swinging arm movement before the jump or not. The vertical jump can be considered either a field (jumping to reach) or lab test (jumping on a force platform) depending on the equipment used (Vandervalle, Peres, and Monod, 1987).

Davies and Young (1984) found a high correlation (r = .92) between jump height and peak power calculated from a force platform. Other studies have presented the test-retest reliability coefficients of certain vertical jump protocols. Gray, Start, and Glencross (1962) achieved an r-value greater than (.98) for their protocol while Glencross (1966) had an r-value greater than (.92) for his protocol. Bosco, Luhtanen, & Komi (1983a) had a high test-retest reliability of (.95) for the standard vertical jump test. Vandewalle et al (1987) suggest the test-retest reliability for the vertical jump test would be even higher if more than 3 trials were completed on untrained subjects.

The vertical jump test has been used in several ice hockey related studies to measure anaerobic power. The first of which by Mascaro et al (1992) examined the best off-ice predictors of sprint skating speed in nine pro ice hockey players aged 23 years old. Off-ice tests used were the 40-yard dash, the vertical and long jump, and isokinetic testing of the quads and hamstrings while the on-ice test was a 54.9 meter sprint. The results of the study showed that skating speed was significantly positively correlated with the vertical jump and this was the best predictor of skating speed of the measures used.

Bracko et al (1997) wanted to identify off-ice variables that contribute to high performance skating acceleration, speed and agility in 30 youth hockey players aged 10-14. The off-ice tests included the sit and reach, vertical and standing long jump, sit-ups, push-ups, and hip abduction flexibility. The on-ice tests included: 6.10-meter acceleration skating test, 47.85 meter speed skating test, 16.3 meter full speed skating test, and the agility cornering S test. Regression analysis showed that vertical jump height and push-ups were the strongest predictors of variables tested for skating acceleration and speed in youth hockey players aged 10-14. Diakoumis et al (1998) followed up that study by comparing off-ice variables that could be used to predict high performance skating in 50 deaf hockey players aged 11-19. Players completed the various tests: vertical jump with and without arm swing, 40-yard dash, 20-feet acceleration skating test, and 147-feet speed skating test. Results from the study showed that the 40-yard dash and both vertical jump tests were reasonable predictors of skating acceleration and speed in that sample.

Studies testing NHL entry draft players have also used the vertical jump as one of the components of the NHL Combine test battery (Burr et al, 2008; Burr et al, 2007; Vescovi et al, 2006a). One such study (Burr et al, 2007) examined prospective draft players at the 2005 NHL entry draft combine to determine the best jumping protocol to test leg power as a predictor for hockey playing ability and draft selection order. The study tested the top 95 players entering the draft on two different jump protocols: the squat vertical jump (SVJ) and the countermovement vertical jump (CVJ). Of the two jump protocols, the squat vertical jump had the highest correlation between leg power ranking and draft selection order. Sayers, Harackiewicz, Harman, Frykman, & Rosenstein (1999) also support the squat vertical jump protocol to be more accurate in predicting peak power.

Margaria-Kalamen Stair Test

Another commonly used test to measure anaerobic power is the Margaria-Kalamen Stair Test. In 1966, Margaria, Aghemo, and Rovelli developed a test for maximum anaerobic power involving stair running. The test used an electronic clock to record the maximum speed of running up stairs (each stair with a height of 17.5 cm) with power measured in watts or kg-m/sec. The test was relatively short with a duration of 4-5 seconds per sprint and involved running up two stairs at a time. The time of the vertical displacement was recorded to the nearest (.01) second with the measurement of power recorded during the time from the 4^{th} to the 6^{th} jump (70 cm height). Usual times are about (.40) to (.50) seconds. To reach maximal speed, participants are allowed a 2-meter running start. The test is not physically demanding since it has a short duration time and can be repeated after only a few minutes of rest (Margaria et al, 1966).

The original protocol by Margaria et al (1966) was modified by Kalamen in 1968. This protocol has a 6-meter running start and requires the participant to run up 3 stairs at a time (Vandewalle et al, 1987). The vertical distance of the stairs in the Margaria-Kalamen protocol is 1.02 meters and the number of stairs is 9 with the time measured from the 3^{rd} step to the 9^{th} step. The participant would perform this test 3 times with the best score or average being taken as the test measure. This protocol has also produced the highest power output (Mayhew, Piper, Etheridge, Schwegler, Beckenholdt, and Thomas, 1990). This test is also very dependent on weight meaning the total power output will be higher in heavier subjects so results sometimes need to be interpreted with caution (Powers et al, 2004).

Margaria et al (1966) states that data from their protocol is reproducible since the variability of power measured on two subjects 9-10 times was less than 2 percent during a 5-week period. Previous studies have tested reliability and produced high test-retest correlation coefficients of (.85) (Ayalon, Inbar, and Bar-Or, 1974) and (.90) (Sawka, Tahamont, Fitzgerald, Miles, and Knowlton, 1980).

Some of the advantages of the Margaria-Kalamen stair test are: it measures power output, requires no particular knowledge of the test beforehand, no expensive equipment is needed (just stairs and stopwatches), the test is not physically exhausting, and it tests large muscle masses in the body. Disadvantages of the test are: it may not be applicable to people not able to perform maximal work or handicapped people (Margaria et al, 1966). Another disadvantage of the test is that various muscle groups outside of the legs contribute to step running, such as the upper limbs and trunk (Bar-Or, 1996).

The Margaria-Kalamen stair test has been used in ice hockey studies but is not as common in the literature as the Vertical Jump or Wingate bike tests. Reed et al (1979) used the MK stair test, along with other tests to validate the Reed Repeat Sprint Skate (RRSS) on-ice anaerobic test. Houston et al (1976) used the MK stair test in their study involving forty-eight players from two Major Junior A hockey teams and one University hockey team in Canada aged 18 to 23. The study investigated physiological and anthropometric measures such as: body fat percentage, pulmonary function, VO_2max (Treadmill), and lactate levels. From the results, the MK Stair anaerobic power mean was 131 kg-m/sec for junior players and 125 kg-m/sec for university players. In a study by

Vainikka et al (1982) collecting performance measures of the Finnish National Men's ice hockey team, the MK stair test was used, as well as, body fat percentage, VO_2max and anaerobic capacity on a cycle ergometer. The mean anaerobic power was 235 watts which was similar to other national hockey teams at the time (Canada, Sweden, Germany and Czechoslovakia), however no recent studies have been completed using this test to measure anaerobic fitness in ice hockey players.

Wingate Cycle Ergometer Test

The most frequently used laboratory test of anaerobic peak power and capacity in ice hockey players is the Wingate test performed on a cycle ergometer. The test was designed to be: simple to conduct, require no specifically skilled personnel, non-invasive, measure muscle performance of anaerobic peak power and capacity in the arms and legs, and be applicable to a wide population (Bar-Or, 1987). It is considered a good test since it requires work from the major muscle groups of the lower limbs to measure peak power, and capacity over a 30-second time period. One disadvantage of the Wingate test is that it requires a cycle ergometer which is expensive and usually only used for research purposes (Chandler et al, 2008).

The test requires arm or leg pedalling against a resistance for 30 seconds. There are three main measures of the test: peak power (highest point of power during the test), capacity (average power sustained over full 30 seconds) and the rate of fatigue (degree of drop-off in the test). Basically, peak power is the maximal anaerobic power in short duration (5 seconds) and the capacity is a measure of the endurance (30 seconds) of the muscles to sustain power (Bar-Or, 1987). The peak power shows the maximum rate of the ATP-PC energy system to make ATP and usually appears during the first few seconds of the test. The mean power output during the test is a measure of anaerobic endurance and shows the maximum capacity of the ATP-PC and glycolysis energy systems to make ATP. The decrease in power is the percent decline and is calculated from the difference between the highest and lowest power output (Powers et al, 2004). The force or resistance recommended for the test varies by different protocols but the Wingate group suggests 7.5% body mass. This was based on a study for children and has proved to be too low for adults. The resistance for adults should be around 9.0% for non-athletes and 10% for athletes. The introduction of toe stirrups on the bike pedals has increased scores of peak and mean power by 5 to 12 percent (Bar-Or, 1987).

Test-retest reliability of anaerobic peak power measured by the Wingate showed correlation coefficients greater than (.90) (Bar-Or, Dotan, and Inbar, 1977). Correlation coefficients for the Wingate measuring anaerobic peak power and anaerobic capacity under standard conditions range from (.89) to (.99) but usually are higher than (.94) (Inbar, Bar-Or, and Skinner, 1996). The reliability of the Wingate has been tested further in different environments, different age groups, and with special populations not involving able-bodied people. One study by Tirosh, Rosenbaum, and Bar-Or (1987) tested the Wingate on children with spastic cerebral palsy and muscular dystrophy for two trials and produced a correlation coefficient of (.96). An unpublished study by Berman and Bar-Or (1992) tested adults aged 54 to 84 years old with chronic obstructive lung disease on the Wingate (15 second protocol) two times one hour apart and produced a score of (.89). Dotan and Bar-Or (1980) tested the Wingate on boys and girls in three different climates (neutral, warm-dry and hot-humid) and produced scores of (.89) to (.93).

To determine the validity of the Wingate, it has been correlated with other anaerobic tests involving sprinting (300-m), swimming (25-m), ice-skating (SAS 40) and the Vertical Jump. The correlation coefficient for most of these tests was (.75) or higher while the weakest was the SAS 40 (.32). The relation between the Wingate and other anaerobic tests is high but not enough to use the Wingate as a predictor of success in other specific tasks (Inbar et al, 1996). In two other studies comparing the Wingate to the MK stair test, the correlation was high with measures of (.79) (Ayalon et al, 1974) and (.83) (Maud and Schultz, 1986).

Mastrangelo, Chaloupka, Kang, Lacke, Angelucci, Martz, and Biren (2004) conducted a study on sixty boys aged 11 to 13 comparing the Wingate to a 30-meter dash sprint test to develop a power equation. For the Wingate protocol, 7.5% body mass was used as a resistance and measurements taken were peak and mean power. From the study, the 30-meter dash was a significant predictor for mean power in the Wingate. It also has been noted that children recover faster from the Wingate test than adults. Young men needed a 10-minute recovery to repeat their performance while prepubescent boys only needed 2-minutes to recover (Bar-Or, 1996).

Granier, Mercier, Anselme, and Prefaut (1995) tested fourteen short and middle distance runners aged 20 years old to evaluate the aerobic and anerobic components of the Wingate test. The results showed that the sprinters had higher anaerobic contributions to the test. Numerous studies have tried to determine the percent of aerobic contribution to the test with Stevens and Wilson (1986) showing 44.3 percent and Bar-Or (1987) showing 27 percent.

There have been numerous ice hockey studies that use the Wingate to measure anaerobic power and capacity. Smith et al (1982) measured physiological variables on the 1980 Canadian Men's Olympic hockey team. The study's intent was to gather baseline measures on VO_2max (cycle), muscular power (handgrip), anaerobic fitness off-ice (Wingate) and on-ice (Reed Repeat Sprint Skate). The lab performance of the team was lower than other elite players at the time but the on-ice performance scores were higher than pro and junior players in other studies. These results again emphasize that sport specific on-ice tests are needed to measure anaerobic fitness in ice hockey players. Twist and Rhodes (1993) examined anaerobic fitness on the Wingate in thirty-one NHL players on the Vancouver Canucks along with other measures such as VO_2max, and maximum heart rate to determine any positional differences. The scores on the Wingate were similar for both forwards and defensemen showing no significant difference among positions.

During the 1985-86 NHL season, Rhodes et al (1986) monitored seventy-five players on physiological measures such as VO_2max, anaerobic power and capacity (Wingate), body composition, grip strength and abdominal endurance. Results of the NHL players showed anaerobic peak power values of 1064 watts for defensemen, 1044 watts for forwards, and 906 watts for goalies. Anaerobic capacity was 802 watts for defensemen, 794 watts for forwards, and 683 for goalies showing that NHL defensemen had higher anaerobic power and capacity when compared to forwards and goalies. These results were also very similar in another study by Rhodes et al (1987) when comparing eighty NHL players on various physiological parameters.

Other studies have used the Wingate as part of the NHL Combine test battery for NHL entry draft players. Vescovi et al (2006a) examined positional profiling for physiological measures for players at the 2001, 2002, and 2003 entry draft combines. The combine test battery includes: height, weight, body fat percentage, bench press, push ups, curl ups, vertical and long jumps, medicine ball throw, sit and reach, Wingate, and VO_2max on a bike. The study reports that defensemen were taller, heavier, and stronger while goalies had the best flexibility but no positional differences were found for the jumps, curl ups, Wingate, and VO_2max. However the NHL combine protocol has various limitations. The first is that the entire protocol must be performed within 2-hours which will certainly lead to fatigue and affect the outcome on performance measures. Another limitation listed by the author is that there are no sport specific tests or on-ice tests for these hockey players whose draft status and hockey career could improve or decline based on lab tests. Vescovi et al (2006b) again looked at NHL entry draft players for the 2001-2003 drafts to see whether off-ice performance tests could predict draft order. The NHL combine test battery was again used and the intent was that performance would indicate or predict certain draft order of players. This was not the case as none of the 12 tests could predict hockey playing potential or distinguish between rounds for the 3 draft years. It has been suggested by the authors that on-ice testing be investigated to develop this relationship.

A similar study involving NHL entry draft players was conducted by Burr et al (2008) to predict whether fitness measures predict playing potential by draft order. This study tested 853 entry draft players from 1998 to 2006 at the NHL combine. The same 12 tests were used as described earlier by Vescovi et al (2006a). From the study, defence were found to have higher anaerobic peak power while the forwards had better aerobic power and lower body fat percentage. This is different from the findings found by Vescovi et al (2006b) and states that anaerobic peak power is important to high draft round selections in all positions. Variables such as body fat percentage and anaerobic power had an influence on draft success while there was no relation to aerobic power and draft success.

Some studies have also used variations of the Wingate protocol to test ice hockey players. Watson & Sargeant (1986) used a 40-second Wingate protocol to correlate with the Sargeant Anaerobic Skate (SAS) test. Quinney et al (2008) reported that the Wingate used to test the Edmonton Oilers NHL team over a span of 26 years involved four 5-second sprints instead of the normal 30-second all out sprint. An anaerobic cycling test for ice hockey players was developed by Gamble and Montgomery (1986) which consists of six 15-second repetitions. Reliability of the test with university and junior players was strong with an intraclass correlation coefficient of (.93). The test was correlated with an on-ice fitness test (RRSS) for validity producing scores of ($r = -.87$) and ($r = -.78$). The test was concluded to be a reliable and valid measure of anaerobic endurance of ice hockey players.

On-Ice / Field Measures for Ice Hockey

In 1994, Gledhill & Jamnik developed a test battery for NHL players to be featured yearly at the NHL entry draft for prospective players. Studies have been done on entry draft players using the test battery such as Vescovi et al (2006a) and Burr et al (2008). However, this test battery has one major flaw. All tests included are performed off-ice rather than on-ice which eliminates a task and sport specific method of measuring values for ice hockey players. Other pro leagues, such as the National Football League (NFL) have entry draft test batteries that are performed in sport specific situations and involve sport specific drills. They have a physical test battery and a positional test battery requiring specific drills to be performed (www.nfl.com/combine/workouts). This is not the case for ice hockey or the National Hockey League (NHL) or any other hockey league. From the literature, there are some aerobic and anaerobic on-ice hockey tests, however only a few exist with the majority of these testing aerobic power. This is re-enforced by Twist & Rhodes (1993) who have also stated that very few on-ice tests exist.

The *Faught Aerobic Skating Test (FAST)* developed by Faught, Nystrom, and Montelpare (2003) is a continuous skating protocol using a CD audio beep to pace the player around a 160 feet course. Increases in workload occur at measured time intervals while players continue to skate until they voluntarily stop due to exhaustion or received 2 violations which are: failure to remain behind pylons before beep was signalled or failure to reach pylons in time before beep was signalled. The FAST had an excellent test-retest correlation of (.94) while Petrella, Montelpare, Nystrom, Plyley, and Faught (2007) validated the FAST using 406 participants aged 9-25 from the atom to university skill level against a VO_2max on a treadmill using the modified Bruce protocol. During the FAST, players wore skates, helmets, gloves, and stick and the test could measure up to 8 players at a time. The males mean number of lengths (43) and mean VO_2max (55.71 ml/kg/min) were higher compared to female mean number of lengths (38) and mean VO_2max (42.73 ml/kg/min). In short the FAST is a quick, inexpensive, sport-specific method of assessing aerobic fitness on-ice.

Another test to measure aerobic fitness on-ice was developed by Leone, Leger, Lariviere, and Comtois (2007). Thirty elite hockey players aged 14.5 years performed the *Skating Multistage Aerobic Test (SMAT)* to predict VO_2max. Once the test was validated, 112 male and 31 female elite players performed the SMAT and the 20-metre Multistage

Shuttle Run Test (Beep) to predict VO_2max. The test procedure consisted of skating a distance of 45-metres back and forth paced by an audible signal. The initial velocity is 3.5 m/s with increments of 0.2 m/s every stage. The players were given 30 seconds rest after each sprint stage (1-minute stages). The test was stopped when the player could not keep pace and full hockey equipment was worn. The correlation between the SMAT and Beep test was modest, however they cannot be deemed equivalent tests. The SMAT is reliable, valid, sport-specific, and up to 20 players can perform at the same time. One limitation of the test is that it should not be used with inexperienced players who cannot stop abruptly.

Due to a recent surge in sport technology innovations, a skating treadmill has been created which simulates ice skating similar to a running treadmill (Dregor, 1997). In 1999, Dregor and Quinney developed a protocol to measure VO_2max on a *skating treadmill (ST).* The ST protocol consisted of skating at a constant speed (14.4-16 km/hr) with the grade or elevation increasing by 2% every 2 minutes. The test was discontinuous with 2-minute stages followed by 2-minute breaks. Subjects wore skates, gloves, stick, shorts. The test was stopped when the subject was exhausted and could not skate anymore. Comparing the ST protocol to a bike VO_2max test, results showed no significant difference for VO_2max (ST = 60.4 ml/kg/min vs. Bike = 59 ml/kg/min). This is a good way to measure aerobic fitness from a sport specific standpoint, but realistically most players or teams do not have access to a skating treadmill which makes it much harder to perform this type of testing.

For the majority of the on-ice anaerobic fitness tests in the literature, most have been developed in the 1970's and 1980's, so are not recent. The first by Lariviere et al

(1976) involved skating as many lengths possible over a 100-foot course during a 5-minute period. The purpose of the study was to develop norms for their skating test. The mean distance skated during the test was between 3417 to 4219 feet and the norms are presented for skill levels ranging from atom to junior age. Another anaerobic test developed by Lariviere & Godbout (1976) involves players skating a distance of 18.3 metres at maximum speed 6 times backwards and 6 times forwards for a total of 12 sprints. The time to complete the total sprints is used to calculate anaerobic capacity.

The most widely used and reliable on-ice anaerobic test in the literature is the *Reed Repeat Sprint Skate (RRSS)* developed by Reed et al (1979). It consists of 6 repeated sprints at maximum speed over a distance of 91.4 metres or 300 feet with a 30-second break between each sprint. In the study, data from 277 players at the junior, university and NHL level was collected. The test was validated with such other tests as: VO_2max, 1-mile run time, peak lactate, and MK stair test. Tests-retest correlations for these measures yielded scores of (.78), (.68), and (.74). The range of times for the test were 12.5 to 21.7 seconds per sprint with a mean sprint time of 14.7 seconds and a mean post-test heart rate of 184 beats/minute. This test has been repeatedly used in the literature for testing anaerobic fitness on-ice. It is perhaps the best test for this purpose. One limitation is that the test is physically exhausting to younger and non-elite players. This can rule out quite a number of players and teams playing hockey. Another limitation is that the test becomes long and starts to enter into the aerobic energy range, therefore not being a true anaerobic test. A better test is needed to test players of all ages and skill levels that is easy to use and does not cause severe fatigue after performance. This is one reason why the author hopes the RIST will be better than the RRSS to test anaerobic performance in ice hockey players.

A study by Watson and Sargeant (1986) examined the suitability of the Wingate test as a measure of anaerobic fitness against the RRSS and the *Sargeant Anaerobic Skate (SAS).* The SAS consisted of 9 pylons placed 6.1 metres apart along the length of the ice. The player would start at the goal line (pylon 1), sprint to pylon 9, then stop, reverse direction, and sprint back to pylon 1. This is repeated for 40 seconds with no break. Forty seconds was selected since 30 seconds was not long enough to factor out maximal glycolytic power according to Katch, Weltman, Martin, and Gray (1977). Twenty-four junior A and university players performed all 3 tests over a 7-day period. The Wingate anaerobic power and capacity scores were lower than the RRSS and SAS scores with both skate tests producing great test-retest reliability (RRSS = .96, SAS = .97) for anaerobic capacity but only fair (RRSS = .73, SAS = .65) for anaerobic power. The Wingate did not show a high relationship with the skate tests, however this might have been since a 40-second Wingate protocol was used instead of the 30-second protocol.

Rhodes et al (1985) performed a similar method by measuring anaerobic measures on-ice (RRSS) and in the lab (Wingate 40 & Anaerobic Treadmill Test) in twenty-four A-level players aged 10 years old. Unlike the Watson and Sargeant (1986) study, there was a significant relationship between the RRSS and Wingate anaerobic capacity (.71) and anaerobic power (.83) implying a high correlation between on-ice and lab anaerobic tests.

Other tests have been used in the research but have not been identified as reliable or valid in the literature as well as their origin. An example of this is a study by Bracko (2001) examining on-ice skating performance and fitness of elite and non-elite womens ice hockey players. The study tested twenty-three players consisting of seven members of the 1996-97 Canadian National Women's team, one member of the Finnish National Women's team and fifteen non-elite players. On-ice tests used were: Cornering S Agility test, 6.10 M Acceleration test, 47.84 M Speed, 15.20 M Full Speed and the RRSS. The results of the study showed that the elite players were faster skaters, and had greater anaerobic capacity than non-elite players. These tests were used in a study by Geithner et al (2006) to compare physical fitness and skating performance in elite women's ice hockey players. The study tested 112 players of the University of Alberta Women's hockey team aged 21.4 years old from the 1999 to 2004 seasons. This team was the best university team in Canada for the testing period winning 4 National CIS titles. Off-ice performance tests completed were: vertical jump, 40-yard dash, and Léger Beep test while on-ice performance tests were: Modified 3-Repeat sprint skate (MRSS), 6.10 metre Acceleration test and the Cornering S-Turn Agility test. From the results, forwards had greater anaerobic and aerobic power and were the fastest skaters on-ice.

METHODS

Participants

The participants in this study were fourteen boys aged 11-12 years old from the Peewee "A" Neebing Hawks ice hockey team in Thunder Bay.

Protocols

The participant's body mass and resting heart rate were taken before each on-ice session while body mass, height, resting heart rate, and blood pressure were taken before each lab session. These values were necessary since they were used to calculate anaerobic average peak power and peak power. The participants completed four different tests that evaluated their anaerobic/short term fitness levels. The four tests are:

1) Repeat Ice Skating test (RIST)
2) Margaria-Kalamen Stair test
3) Standard Vertical Jump test
4) Wingate Cycle Ergometer test

A brief description of all measures taken is explained below:

1) Standing Height

Each participant stood in an upright position with his arms by sides, shoes off with feet together and heels touching the wall. The participant's back was against the wall next to measuring tape. While the participant stood erect, the researcher placed a box against the subject's hairline and measured at a mark. The distance from the floor to the mark was recorded as the participant's standing height. This distance is recorded in meters (m) (Canadian Physical Activity, Fitness & Lifestyle Appraisal, 2003).

2) Body Mass

Each participant stood with his shoes off in an upright position and with light clothing on. The participant then stepped on a weight scale and body mass was measured to the nearest kilogram (kg) (Canadian Physical Activity, Fitness & Lifestyle Appraisal, 2003).

3) Resting Blood Pressure

Before the researcher began to take blood pressure, the participant sat down for 5-minutes to relax. Once measurement began, the subject extended one arm (left) with the hand supinated (palm facing up) and rested the arm on a table. The participant was asked to open and close his fist ten times to increase blood flow in vessels of the arm and the arm was placed at heart level. The blood pressure cuff was then applied directly over the brachial artery with the diaphragm of the stethoscope under the cuff. The cuff was inflated as fast as possible to 140 mmHg. If Korotkoff sounds were heard then the cuff was inflated another 20 mmHg to 160 mmHg or until no Korotkoff sounds were heard. The cuff was deflated at a rate of 5mmHg/sec until the first Korotkoff sounds were heard. Systolic pressure was read to the nearest 5mmHg and diastolic pressure was reported once no more Korotkoff sounds were heard to the nearest 5mmHg. Blood pressure was measured in millimetres of mercury (mmHg) (Canadian Physical Activity, Fitness & Lifestyle Appraisal, 2003).

4) Resting Heart Rate

Each participant sat down and relaxed for a couple of minutes before taking measurement. Once ready to begin, the researcher located the participant's pulse by pressing the middle and index fingers of one hand on the carotid artery (neck). The researcher then counted the pulse for 15 seconds and multiplied this number by 4 to account for 60 seconds. The number of beats counted in one minute was the resting heart rate.

5) Repeat Ice Skating Test (RIST)

Two participants completed the test at the same time. The participants started with a 5-minute warm-up of light skating and 5-minutes of stretching before beginning the test. When the test began, two participants started on the centre ice line at the edge of the centre ice circle. On the whistle, the participants sprinted from the centre ice line toward the net, skated around the net and returned to the centre ice line at the other side of the centre ice circle. The total distance covered was 160 feet or 49 meters. The participant's time was recorded in seconds by a stopwatch when the player's first skate hit the red line.

A 10-second break was taken. Following the 10-second break, the participants repeated the sprint distance as before and another 10-second break was taken. A third sprint was performed and the participant's time was recorded for all three trials. Immediately after the third sprint, the participant's heart rate was recorded using the carotid pulse. The participants then took a minimum 10-minute break. After the 10-minute break, the participants repeated the same sprint procedure for a total of 6 sprints. The participant's six sprint times were taken and the best sprint time was used to estimate average peak power. These values were calculated using the formula presented by Watson & Sargeant (1986):

Average Peak Power (watts) = Body Mass (kg) x Distance (49 m) x 9.81m/s / Time (s)

6) Margaria-Kalamen Stair Test

Each participant warmed up for 5-minutes on a bike and stretched. The researcher placed tape on the 3rd and 9th step on a series of stairs at least 17 cm tall each as well as tape on a line 6 meters away from first step. There was one timer on the 3rd step and one timer on the 9th step with stopwatches. Once instructed, the participant ran from the 6-meter mark as fast as they could up the stairs taking three at a time (3rd, 6th, 9th). The timers started recording when the participant hit the 3rd step and stopped recording at the 9th step. The average time was taken from the two timers for each trial. The participant did 3 trials and the best time was used. The participant had a 15 to 20-second break between trials. The anaerobic power was measured in watts and was the product of force (weight of participant) multiplied by distance (height of 9 stairs) and divided by time (seconds) (Margaria et al, 1966). An equation for calculating anaerobic power in watts was provided by Mayhew et al (1990):

Anaerobic Power (watts) = Body Mass (kg) x Distance (1.08m) x 9.81 m/s / Time (s)

7) Standard Vertical Jump Test

Each participant stood up facing sideways to a wall on which a measuring tape had been attached. Standing erect with feet flat on the floor, the participant reached as high as possible on the tape with one arm and fingers fully extended and the palm on the wall. This was recorded as the stand and reach height to the nearest 0.5 cm. Next, the participant moved a safe distance away from the wall (with the hand on the hip, the elbow should barely reach the wall). No run-up, step-up or pre-jump was permitted. The participant brought his arms downward and backward while bending the knees to a balanced, semi-squat position. The individual paused momentarily in this position (to minimize the possibility of a pre-jump) and jumped as high as possible with the arms moving forward and upward, touching the tape at the peak height of the jump with the arm and fingers fully extended. Three trials were completed with a rest period of 10-seconds between trials. Subtracting the stand and reach measurement from the peak height of the best of the three trials is the maximum difference or jump height (Canadian Physical Activity, Fitness and Lifestyle Appraisal, 2003). The peak power was calculated from the following equation provided by Sayers et al.(1999):

Peak Power (watts) = {60.7 x Jump Height (cm)} + {45.3 x Body Mass (kg)} – 2055

8) Wingate Cycle Ergometer Test

This test was completed on a Monark bicycle ergometer. The cycle ergometer was hooked up to a computer station that recorded pedal revolutions and all measurements needed. Each participant's body mass was recorded and a resistance mass (5% of subject's body mass) was calculated for the test phase. This resistance was then put on a lever that was lowered to create resistance during the sprint trials. The test had five major time periods. The first was a 5-minute warm-up that included cycling at low intensity interspersed with three 10-second sprints at the resistance mass. This was done to familiarize the participant with the test and the resistance mass. The participant then had a 5-minute rest. The next phase was the acceleration phase which was a 10-second countdown before lowering the resistance. This gave the participant time to accelerate the pedal revolutions and brace for the resistance. After this phase was the 30-second all out sprint with the resistance mass. The last phase was the cool down period where the participant cycled at low intensity for one or two minutes until they felt fine (Mastrangelo et al, 2004). Anaerobic peak power was measured from this test in watts and w/kg.

Procedure

All participants read and filled out: a consent form (Appendix A), physical activity readiness form (Appendix B), and a maximal exercise assessment form (Appendix C). These forms were used to determine any health problems that would affect the outcome of the study or prevent the players from participating. There was also a cover letter (Appendix D) attached that provided details about the study and had to be read and signed before filling out the previously stated forms. All laboratory tests (Vertical Jump, MK Stair, Wingate) were conducted in the Human Performance Lab at the CJ Sanders Fieldhouse in the School of Kinesiology at Lakehead University. One field anaerobic test (RIST) was conducted on-ice at the Current River Arena in Thunder Bay. Participants were asked to refrain from physical activity 24 hours before testing and also to refrain from eating two hours before testing.

The first day of testing consisted of on-ice testing. Players wore full ice hockey equipment even though it was not necessary. Before entering the ice, the player's resting heart rate, and body mass were recorded. Once on the ice, the players warmed up with 5-minutes of light skating followed by 5-minutes of stretching. The RIST was then completed.

The second day of testing consisted of all lab tests at Lakehead University. Each participant wore shorts, t-shirts, socks and sneakers for testing in the lab. Before testing began, each participant's height, body mass, blood pressure and resting heart rate was recorded. The participants warmed up on a bike for 5-minutes followed by 5-minutes of stretching before starting the tests. The lab tests were completed in the following order: Vertical Jump, MK Stair, and Wingate Bike. The participants received a 2 to 3-minute break before starting each test and this concluded the first phase of testing. The participants were then retested three days later in the exact same order.

Independent and Dependent Variables

The independent variable in this study is time. The dependent variables are: anaerobic peak power and average peak power measured using 4 different tests (RIST, Vertical Jump, Margaria-Kalamen Stair, Wingate).

Statistical Analysis

Descriptive statistics (means, standard deviations) were calculated for the following variables: body mass, blood pressure, resting heart rate, maximum heart rate, RIST sprint times, Vertical Jump height, Stair running sprint times, anaerobic peak power and average peak power in watts and w/kg for 4 tests.

Test-Retest Reliability

Test-Retest is a process of determining if a test is reliable by giving the same test on two occasions to the same population and seeing if the correlation between the two testing periods is high (Morrow et al, 2005).

For all tests (RIST, Vertical Jump, MK Stair, Wingate) 4 separate intra-class correlation coefficients (r) were computed for test-retest reliability as outlined by Safrit (1990). This was done to make sure that each test is reliable since each test used different protocols to calculate anaerobic peak power and average peak power. For instance, the protocols for the Wingate and Vertical Jump are specific for adolescents which is important since this is the population being investigated.

Concurrent Validity

After establishing reliability of the RIST, the next step was to validate the test. The type of validity the author was interested in was concurrent validity which is the relationship between a test and criterion measures when they are taken close together in time. It is based on a correlation coefficient between the criterion (gold standard) and the test (Morrow et al, 2005).

To establish validity, a Pearson product-moment correlation was used. The RIST was correlated with each lab test so three separate correlations were completed. Using the Pearson Product-moment correlation, the author wanted to show that a player who scored high on the RIST would also score high on the lab tests while a person who scored low on the RIST would also score low on the lab tests.

Limitations

1) Performance on this test depends on motivation to perform well. If the players are not motivated, then this will affect the consistency and accuracy of scores for the RIST.

2) Hockey training and cross-training activities that the team might take part in

before or around the same time as testing may fatigue the players and affect performance on the various tests.

3) The data collected and analyzed will only be useful to a specific population (adolescent or Peewee age, "A" skill level, ice hockey players).

4) Age was a limitation since the players did not perform the RRSS and the RIST during on-ice testing since it would have been too exhausting physically for them.

5) The ability of recruiting certain teams of different age and skill levels was difficult since testing took place at the end of the hockey season.

RESULTS

The physical fitness measures of the 14 Peewee A Hawks players are listed in Tables 1 to 5. The team consisted of 2 goalies, 5 defensemen, and 7 forwards. As evident in Table 1 the values did not really change from Trial 1 to Trial 2. The mean maximum heart rates of 169 bpm in trial 1 and 171 bpm in trial 2 showed that the players were working in 81% of their age predicted maximum heart rate (220 – age).

Table 1
Body Mass, Blood Pressure and Heart Rates of Peewee Hawks players in Trials 1 and 2

Subject	Body Mass (kg)		Blood Pressure (mmHg)		RHR	MHR (bpm)	
	Trial 1	Trial 2	Trial 1	Trial 2		Trial 1	Trial 2
1	58.0	58.5	122/78	112/64	72	174	174
2	43.0	43.0	122/82	122/82	90	180	169
3	42.0	42.0	100/58	102/58	102	162	174
4	41.0	41.0	100/66	108/54	66	180	168
5	41.0	41.0	110/73	110/73	76	150	150
6	43.0	42.0	100/58	100/60	84	150	168
7	39.0	38.0	104/72	100/70	90	186	180
8	39.0	39.0	108/58	108/54	90	144	174
9	40.0	40.0	109/64	112/60	72	180	174
10	37.0	36.0	98/64	100/60	65	168	168
11	37.0	37.5	104/68	102/60	60	180	174
12	35.0	35.0	112/68	90/66	78	168	174
13	42.0	41.5	102/64	104/58	90	162	168
14	38.0	39.0	112/64	110/80	96	180	180
Mean	41.1	40.9	107/67	106/64	81	169	171
SD	5.4	5.6			12.8	13.5	7.3

Note. RHR = Resting Heart Rate, MHR = Maximum Heart Rate during RIST and SD = Standard deviation

The results of the Repeat Ice Skating Test (RIST) are shown in Table 2 and Figure 1. The mean RIST sprint times and the average peak power in watts and w/kg did not change from Trial 1 to Trial 2. However, the range of scores between players individually from lowest to highest was quite large for average peak power in watts and quite small for w/kg. Upon viewing Figure 1 and Table 2, you can see that subject 1's average peak power in watts for the RIST is much higher than the rest of the team.

Table 2

Repeat Ice Skating Test results for Peewee Hawks in Trials 1 and 2

Subject	Time (s)		Average Peak Power (watts)		Average Peak Power (w/kg)	
	Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2
1	9.5	9.7	2935	2877	50.6	50.6
2	9.4	9.4	2197	2190	51.0	50.9
3	9.6	9.4	2103	2146	50.1	51.1
4	9.4	9.6	2094	2055	51.0	50.1
5	10.0	9.9	1971	1977	48.1	48.0
6	10.5	10.1	1969	2055	45.8	47.8
7	9.5	9.8	1967	1917	50.0	49.2
8	9.7	9.8	1923	1903	49.0	48.8
9	10.1	10.1	1911	1904	47.8	47.6
10	9.7	9.9	1832	1797	49.5	48.6
11	9.8	9.7	1819	1832	49.2	49.5
12	9.8	10.0	1713	1677	48.9	47.9
13	12.5	12.8	1615	1576	38.5	37.5
14	11.8	11.7	1543	1561	40.6	41.1
Mean	10.1	10.1	1971	1962	47.9	47.7
SD	0.93	0.95	332.6	326.4	3.8	3.8

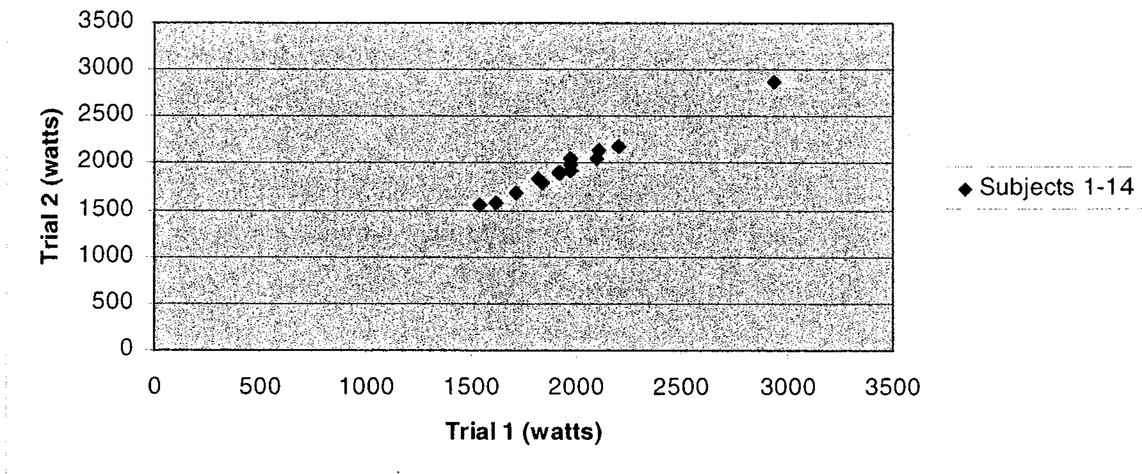

Figure 1. Scatter plot of Peewee Hawks RIST average peak power scores in Trials 1 and 2.

The results of the Wingate test are shown in Table 3 and Figure 2. The resistance used for the team was 5% of their body mass. Resistances used for other populations include 7.5% of body mass as the norm for a sedentary population while 9 or 10% of body mass for adults and athletes (Bar-Or, 1987). The mean resistance mass used was 2 kg for both trials with a range of 1.75 to 2.9 kg. The team's peak power in watts and w/kg did not change much between trials with only a slight decrease in peak power in watts. The range of the individual players scores for watts in trial 1 are from 191 to 416 which is large. When the scores are in w/kg in trial 1, the range is very small measuring from 5.5 to 7.2. Once again, subject 1's score on the Wingate in watts was much larger than the rest of the team's scores in both trials.

Table 3
Wingate Cycle Ergometer results for Peewee Hawks in Trials 1 and 2

Subject	Peak Power (watts) Trial 1	Peak Power (watts) Trial 2	Peak Power (w/kg) Trial 1	Peak Power (w/kg) Trial 2
1	416	439	7.2	7.5
2	272	259	6.3	6.4
3	254	262	6.0	6.2
4	256	233	6.2	5.7
5	283	259	6.9	6.3
6	264	284	6.1	6.9
7	266	228	6.8	6.0
8	255	245	6.5	6.3
9	234	271	5.9	6.8
10	234	217	6.3	6.0
11	266	283	7.2	7.5
12	191	196	5.5	5.6
13	243	222	5.8	5.3
14	258	228	6.8	5.8
Mean	263.7	259	6.4	6.3
SD	49.1	62.9	0.53	0.72

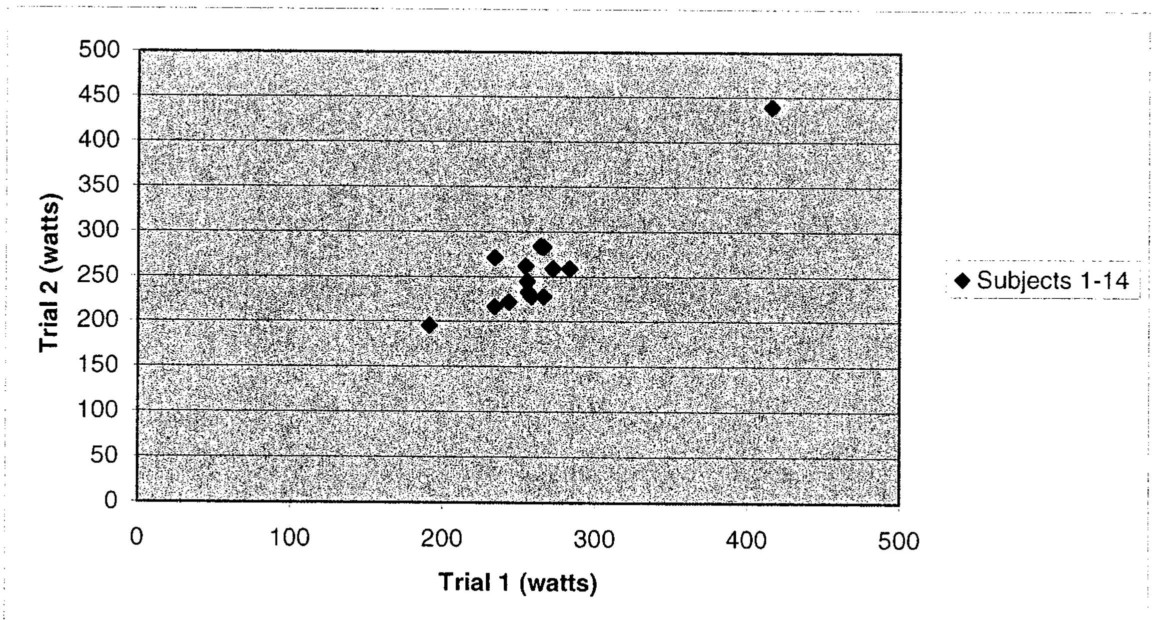

Figure 2. Scatter plot of Peewee Hawks Wingate peak power scores in Trials 1 and 2.

The values of the Margaria-Kalamen Stair test are listed in Table 4 and Figure 3.

The mean stair running times did not change between trials nor did mean peak power in w/kg. The mean peak power in watts did increase from 562 watts in Trial 1 to 587.2 watts in Trial 2. This was the only test in this study in which the team's scores improved from trial 1 to trial 2.

Table 4
Margaria-Kalamen Stair test results for Peewee Hawks in Trials 1 and 2

Subject	Time (s)		Peak Power (watts)		Peak Power (w/kg)	
	Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2
1	0.77	0.71	803	879	13.8	15.0
2	0.81	0.78	562	587	13.1	13.1
3	0.72	0.74	622	605	14.8	14.4
4	0.60	0.55	730	797	17.8	19.4
5	0.90	0.90	483	587	11.8	14.3
6	0.80	0.91	573	492	13.0	11.7
7	0.99	0.94	419	431	10.8	11.3
8	0.85	0.80	486	517	12.5	13.2
9	0.77	0.75	553	565	13.8	14.1
10	0.71	0.79	556	483	15.0	13.4
11	0.64	0.74	617	541	16.7	14.4
12	0.78	0.71	473	526	13.6	15.0
13	0.83	0.74	445	598	10.6	14.4
14	0.74	0.68	547	612	14.0	15.7
Mean	0.78	0.77	562	587.2	13.7	14.3
SD	0.10	0.10	106.5	129.9	2.0	2.1

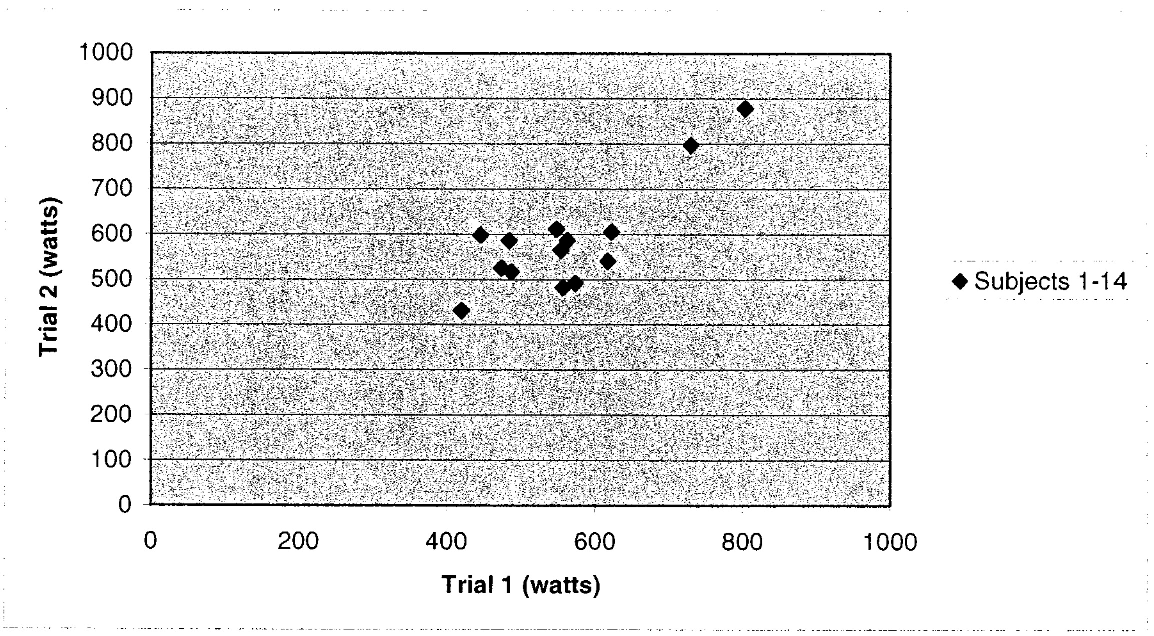

Figure 3. Scatter plot of Peewee Hawks MK Stair peak power scores in Trials 1 and 2.

The results of the Vertical Jump test are shown in Table 5 and Figure 4. The mean jump heights and peak power in watts and w/kg did not change from Trials 1 to 2.

Table 5

Vertical Jump results for Peewee Hawks in Trials 1 and 2

Subject	Jump Height (cm)		Peak Power (watts)		Peak Power (w/kg)	
	Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2
1	31.0	29.0	2454	2355	42.3	40.3
2	30.0	30.0	1714	1525	39.9	37.9
3	28.0	26.0	1547	1426	36.8	33.9
4	32.0	31.0	1745	1684	42.6	41.1
5	30.0	30.0	1623	1525	39.6	37.3
6	23.0	30.5	1289	1699	29.9	40.5
7	34.0	32.0	1776	1608	45.5	42.3
8	25.0	27.5	1229	1381	31.5	35.4
9	33.0	26.0	1760	1335	44.0	33.4
10	29.0	27.5	1381	1245	37.3	34.6
11	28.0	32.0	1321	1586	35.7	42.3
12	29.0	30.0	1291	1352	36.9	38.6
13	25.0	25.0	1365	1343	32.5	32.3
14	28.0	26.0	1366	1290	35.9	33.1
Mean	29.0	28.8	1562	1525	37.9	37.3
SD	3.1	2.4	324.4	304.3	4.7	3.9

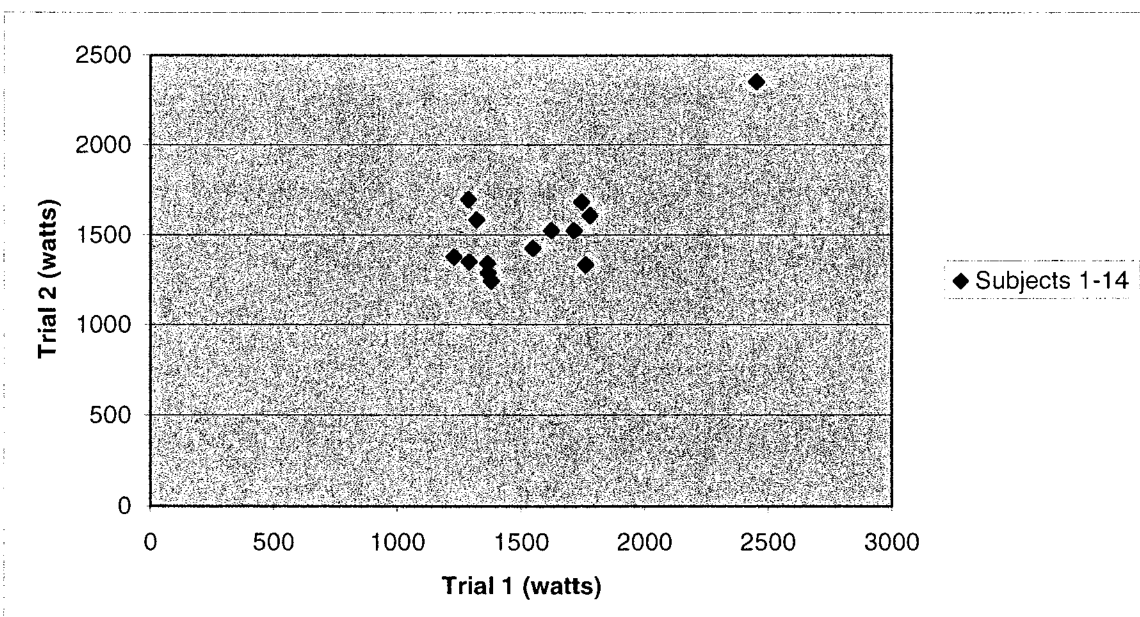

Figure 4. Scatter plot of Peewee Hawks Vertical Jump peak power scores in Trials 1, 2.

To examine the test-retest reliability of all anaerobic tests in this study, 4 Intra-class correlation coefficients were calculated to measure the stability of the test scores measuring peak power and average peak power in watts and w/kg across two trials. The r-values for all 4 tests are shown in Table 6. All r-values in watts were above (.87) showing a strong correlation of the test scores between trials. The r-values in w/kg were above (.45) showing a moderate to strong correlation of the test scores between trials.

Table 6
Intra-class Correlation Coeffecients (r) for Peewee Hawks tests measuring average peak power and peak power

Test	(watts)	(w/kg)
RIST	0.99	0.98
Wingate	0.95	0.72
MK Stair	0.89	0.76
Vertical Jump	0.87	0.45

To measure the concurrent validity of the RIST, it was correlated with 3 anaerobic lab tests producing measures of peak power in watts and w/kg using the Pearson product-moment correlation procedure. These values are shown in Tables 7 and 8. Scatter plots of these values are also shown in Figures 5, 6, and 7 in Appendix G. The r-values in watts for each correlation were above (.71). The strongest correlation to the RIST was the Vertical Jump, followed by the Wingate, and lastly the Margaria-Kalamen Stair test. The r-values in w/kg were considerably lower but above (.18). In these units, the strongest correlation to the RIST was the Vertical Jump, followed by the Margaria-Kalamen Stair test, and lastly the Wingate.

Table 7
Pearson product-moment Correlation results for tests measuring average peak power and peak power in watts

Test Correlations	Correlation Values
RIST vs. Vertical Jump	0.86
RIST vs. MK Stair	0.71
RIST vs. Wingate	0.85

Table 8
Pearson product-moment Correlation results for testing measuring average peak power and peak power in w/kg

Test Correlations	Correlation Values
RIST vs. Vertical Jump	0.50
RIST vs. MK Stair	0.40
RIST vs. Wingate	0.18

DISCUSSION

The purpose of this study was to establish the test-retest reliability and concurrent validity of the Repeat Ice Skating Test (RIST) in estimating anaerobic average peak power in male Peewee "A" (11-12 years old) ice hockey players. The hypothesis was that the RIST would produce stable measures of average peak power over time with an r-value of (.70) or higher. It was also hypothesized that the RIST would be a valid test that would produce high correlations with the Vertical Jump test, Margaria-Kalamen Stair test, and lastly the Wingate Bike test.

The Hawks r-values for test-retest reliability listed in Table 6 were all above (.87) for watts with the RIST scoring a value of (.99). These are considered to be "good to high" correlations. Previous literature has determined r-values that are deemed acceptable and are listed in Table 9. From those articles listed in Table 9, it is shown that r-values of (.80) or higher are "good" with those down to (0.70) deemed "questionable".

Table 9
Reported R-Values for Repeat Reliability in previous literature

Reference	Acceptable Test-Retest Reliability R-Values
Vincent (1994)	0.70 (questionable) 0.80 (good) 0.90 (high)
Atkinson & Nevill (1998)	>0.80 using Pearson Product-moment Correlation
Hopkins (2000)	>0.81 using ICC
Baumgartner & Chung (2001)	>0.70 using ICC
Morrow et al (2005)	>0.80 using ICC

The test-retest reliability r-values for average peak power and peak power in w/kg were above (.72) for all tests with the exception of the vertical jump (.45). Prior to this study, a pilot study was conducted with a male Bantam "AAA" ice hockey team in the Thunder Bay area. The pilot study consisted of the same tests and methods as the present study but the testing phases took place at the beginning and end of the team's season. The Bantam Kings test-retest reliability results are shown in Table 12 in Appendix E. The Kings r-values for all tests were above (.81). These two teams were selected for this study since they both played in the Thunder Bay area and were at different age and skill levels. Another reason is that the majority of ice hockey research involves teams at the elite level and players who are of adult age. This study wanted to investigate anaerobic power in ice hockey players in a younger population and at different skill levels. The age of the players might also affect scores since some players were entering the stage of puberty in adolescence. This stage of development is characterized by rapid growth where 10 inches in height and 40 pounds in body mass is added over a span of 4-5 years (Berk, 2001). For boys, this stage means lower arm and leg fat mass, increases in muscle mass, lung capacity, and red blood cells (Beunen and Malina, 1996). Due to this, some players might be developing faster physically than others meaning more body mass and leaner muscle mass is added. This could increase speed, endurance and strength and affect scores for these players. Testing this age group may be more variable since they are not fully developed physically compared to an adult population who are fully grown and developed.

Specifically, the reliability r-values in our study were lower in w/kg than watts for all 3 lab anaerobic tests while the RIST r-values were very similar for both units. This change in reliability for each unit of measure could be due to the wide range of body masses of the players that were from 35 to 58 kilograms. Since body mass is required for all 4 tests in this study to calculate peak power, the range of body masses might contribute to the variability of the scores. When these body masses are converted to w/kg, the range becomes restricted and there is less variability in the sample. Given the age of the players and the stage of growth they are experiencing during puberty, the size of the players physically is rather hard to control. Some players could go through this stage rather fast while others would develop slower. This could cause great differences in the range of scores on the tests. The time between the two testing phases was 3-days which was relatively short so there was no training effect. This is shown by the fact the players peak power scores decreased slightly between the two testing phases but not enough to be significant. With the exception of the vertical jump r-value of (.45) for peak power in w/kg all other correlations met this study's hypothesis of being above (.70).

Other previously reported reliability r-values for anaerobic tests are listed in Table 10. When comparing r-values for anaerobic tests specifically designed for hockey players, such as on-ice tests like the Reed Repeat Sprint Skate (Reed et al, 1979) and the Sargeant Anaerobic Skate (Watson & Sargeant, 1986) these r-values were lower than the those for the RIST. Compared to the other anaerobic tests, such as the Wingate, Vertical Jump and Margaria-Kalamen Stair test, the RIST r-values were on par with or higher than these tests r-values showing high test-retest reliability to measure stable scores of average peak power in watts. The reasons why the RIST r-value is so high could be due to the RIST being designed specifically for ice hockey players to perform on-ice and that is what the study sample consisted of. It would be expected for these players to do well on the test during both phases considering the test is sport specific. The mean sprint times on the RIST for the Peewee Hawks were the same for both trials showing how stable the scores were for the team. Another reason possibly contributing to why the RIST r-values may be higher is that the test is much shorter that the RRSS and SAS which allows the players an adequate rest and does not exhaust the players which would affect their scores. Since the RRSS and SAS are long tests, they start to leave the anaerobic energy time range and begin to enter the aerobic energy time range. The sample in Watson & Sargeant's (1986) study were Junior aged players (18-21 years old) while this study had Peewee aged players (11-12 years old) which might have affected the scores due to size, strength and skill level since the Junior players are at a more elite level. The test-retest scores were measured 7 days apart in the Watson & Sargeant (1986) study while this study was retested in 3 days. This length of time could have lead to injuries, fatigue, or a training effect for the players, affecting their retest scores. In terms of the other reliability r-values from Table 10, most of these studies have different study samples of different ages, conditions, and skill levels for those involving hockey players. The r-values shown in Table 10 are similar to this study's r-values involving the same anaerobic tests.

Another issue in this study was the use of two different types of peak power being measured. The 3 anaerobic lab tests were used to measure peak power while the RIST measured average peak power. The term "average peak power" was used since we cannot pinpoint the exact peak power during one sprint of the RIST. It could also be debated the peak is before the turn around the net or after the turn around the net in the RIST. Therefore, the whole sprint time of one sprint was taken to calculate peak power. This was effective as correlations for the reliability and validity of the RIST were still strong. This could be due to the average sprint time being roughly 9-10 seconds which is still in the anaerobic peak power range but is different from the other 3 lab tests which calculate peak power in 5 seconds or less. Since the tests calculate peak power using different variables, formulas and for different age populations, there will be more variability in players scores for each test. However, since the players were ranked on each test, and maintained similar ranks between testing trials, the correlations were strong.

Table 10
Reported Test-Retest Reliability r-values of anaerobic tests measuring peak power

Sample	Test	R-Value	Reference
38 children aged 6-20 with neuro-musclar disease	Wingate	0.96	Tirosh, Rosenbaum, & Bar-Or (1987)
28 children aged 10-12	Wingate	0.89-0.93	Dotan & Bar-Or (1983)
children and adults	Wingate	0.95-0.97	Bar-Or et al (1977)
17 varsity and 14 junior ice hockey players	Modified Bike Test	0.93-0.99	Gamble & Montgomery (1986)
	MK Stair	0.85	Ayalon et al (1974)
	MK Stair	0.90	Sawka et al (1980)
	Vertical Jump	>0.98	Gray et al (1962)
	Vertical Jump	0.92	Glencross (1962)
	Vertical Jump	0.95	Bosco et al (1983a)
24 junior A ice hockey players	SAS	0.65	Watson & Sargeant (1986)
24 junior A ice hockey players	RRSS	0.73	Watson & Sargeant (1986)

The Hawks r-values from the Pearson product-moment correlation coefficient for concurrent validity are listed in Table 7. The values were above (.71) for all correlations measuring average peak power and peak power in watts. The r-values for the pilot study involving the Bantam Kings team measuring peak power and average peak power in watts are listed in Table 13 in Appendix F. These values are above (.85) with the exception of the correlation between the RIST and Vertical Jump (.49).

The r-values in the present study involving the Peewee Hawks are considered to be "strong, positive" correlations according to Morrow et al (2005). This means that the players who scored high on the RIST also scored high on the lab anaerobic tests and the players who scored low on the RIST also scored low on the lab anaerobic tests. Basically, the players scored similar on all tests meaning they held the same ranking order for their team between testing trials. This shows that the RIST is similar to the other anaerobic lab tests for estimating anaerobic average peak power.

The tests that correlated the best with the RIST in watts were the Vertical Jump (.86), the Wingate cycle ergometer test (.85), and the Margaria-Kalamen Stair test (.71). This was not the right order of correlations hypothesized as the RIST was expected to correlate best with Vertical Jump, followed by the Margaria-Kalamen Stair test, and the Wingate cycle ergometer test. However, the current data does make sense since the vertical jump and Wingate are the more widely used anaerobic tests used in the literature and specifically used for studies in ice hockey. The vertical jump has been proven in the literature to be significantly correlated with anaerobic power in ice hockey. Studies by Bracko et al (1997), Mascaro et al (1992), and Diakoumis et al (1998) all found the Vertical Jump to be the best predictor of skating acceleration for ice hockey players. A study by Burr et al (2007) found that Vertical Jump had the highest correlation to draft order for NHL entry draft players. This could be due to the same muscles being used for each activity since the players are crouching for the vertical jump and also crouching when skating. These same muscles are being used in the Wingate cycle ergometer test but in a seated position. The Margaria-Kalamen Stair test is not measured in as many ice hockey studies so there are less correlations to report. The test is somewhat different from all other anaerobic tests since it relies on running up stairs. The players body mass or motivation might affect the run-up speed to the stairs as well as their height or leg length which might help or hinder certain players times when running up the stairs on the test.

The correlations in w/kg shown in Table 8 were positive but not strong correlations according to Morrow et al (2005). However, this study was concerned with peak power being measured in watts and not w/kg. This is appropriate since the study is measuring one team completing two trials, making it a within-group comparison. If the Peewee Hawks were compared to another team or the general population, w/kg would be used as it would be standard way to view scores across different populations. The correlations were much higher for peak power in watts than w/kg for the Peewee Hawks. This could be due to the wide range of body masses of the players (35 to 58 kg) on the team. This large range leads to a great deal of variability and would affect scores since body mass is used to calculate peak power for every test in this study. Every player on the team also has different amounts of muscle mass and body fat included in their weight and this could lead to variability among the subjects and their scores. Lastly, the lack of variability in w/kg could change the ranks of the players on all tests which would affect the validity of measuring peak power. There may not be enough variability in this study's data since only one team was measured causing a restricted range in the scores. The RIST has proven to be valid when measuring average peak power in watts but the low correlations in w/kg does lower the test's validity for this unit of measurement. However, w/kg would more likely be used as a comparison measure with two teams or a comparing a team of hockey players with non-hockey players. Perhaps including more teams, thereby increasing the number of participants, would produce a greater range of scores and increase the variability of the data. Also including more elite and older players would help increase the validity of the RIST in w/kg.

A list of previous literature reporting correlations between the Wingate and other anaerobic tests measuring peak power is shown in Table 11.

Table 11
Reported r-values for Validity correlations involving anaerobic tests and the Wingate

Sample	Test Correlations	R-Values	Reference
48 adolescent boys	Vertical Jump vs. Wingate	0.77	Davies & Young (1984)
56 male children aged 10-15	Vertical Jump vs. Wingate	0.70	Tharp et al (1985)
15 people	MK Stair vs. Wingate	0.79	Ayalon et al (1974)
112 males & 78 females aged 18-28	MK Stair vs. Wingate	0.83	Maud & Schultz (1989)
24 male ice hockey players aged 10	RRSS vs. Wingate	0.83	Rhodes et al (1985)
24 male Junior ice hockey players	RRSS vs. Wingate	0.57	Watson & Sargeant (1986)
24 male Junior ice hockey players	SAS vs. Wingate	0.32	Watson & Sargeant (1986)
14 Peewee A ice hockey players	RIST vs. Wingate	0.85	Present Study
15 Bantam AAA ice hockey players	RIST vs. Wingate	0.85	Present Study

The strongest correlations to the Wingate were the RIST (.85) followed by the RRSS (.83) and MK Stair (.83). The correlation between the RIST and the Wingate (.85) in this study was similar to one study by Rhodes (1985) correlating the RRSS and the

Wingate (.83). Another study by Watson & Sargeant (1986) correlating the same tests had a low r-value (.57) while the SAS and the Wingate (.32) was also a positive low correlation. This shows that the RIST is a valid test that can produce stable scores of average peak power in watts by having a "strong, positive" correlation with an anaerobic test that is considered a "gold standard". It also means that a person who can score high on the RIST should be able to score high on other anaerobic tests used in this study and a person who scores low on the RIST should score low on the other anaerobic tests. The level of concurrent validity for the RIST in this study is met with strong, positive correlations to anaerobic tests deemed to be "gold standards". The RIST however, can be compared to a certain level of construct validity in which no "gold standard" test exists. This is the case for anaerobic on-ice tests but the RIST is stable in its ability to measure average peak power. The validity of the RIST needs to be further reinforced and strengthened with testing on a greater spectrum of players at varying age, gender and skill levels.

SUMMARY AND RECOMMENDATIONS

Summary

This study examined the test-retest reliability and concurrent validity of the RIST in male adolescent ice hockey players. The RIST is an on-ice anaerobic fitness test that measures average peak power. The RIST was tested for reliability by a test-retest method over one week. It was validated with 3 anaerobic lab tests over one week of testing. These 3 tests were: vertical jump, Margaria-Kalamen Stair test, and the Wingate cycle ergometer test. All 4 tests were used to measure anaerobic peak power and average peak power in watts and w/kg.

The first hypothesis predicted that the RIST and the 3 lab tests are reliable tests that measure stable readings of anaerobic peak power and average peak power with a test-retest reliability coefficient (r) > .70 for all tests. Upon testing the players, each test had an r-value above (.87) for peak power in watts. When the results were expressed relative to body weight in w/kg, the tests had r-values above (.72) with the exception of vertical jump (.45).

The second hypothesis stated that players who scored high on the RIST would also score high on the other 3 lab tests and players who scored low on the RIST would score low on the other 3 lab tests, proving the RIST to be a valid test.

Recommendations

This study was crucial to the development of the Repeat Ice Skating Test by establishing the test-retest reliability and concurrent validity of the test in estimating average peak power in watts. Some observations during the on-ice testing of the RIST was that since the test measured 2 players at a time, it simulated a race situation that facilitated maintaining player motivation to achieve optimal sprint time. The advantage of performing two sets of 3 sprints is that it lets the players get used to the test, learn the best way to approach the turn around the net (by crossovers rather than gliding) and allows the players to see if their times were stable across both skating sets. One observation that was noted is that the players glided after their sprint for the 10-second break while waiting for the second sprint to begin. This will be recommended from now on since the players are constantly moving during gliding rather than stopping after each sprint which might make it harder for the players to get moving for their next sprint. Another recommendation is that players should skate through the side face-off circle to approach the turn around the net at proper angle to allow for crossovers to occur which in turn helps improve the sprint time after circling the net. Players were encouraged throughout the test by verbal instructions from the fellow players and test administrators. Before each sprint, the test administrator will verbally deliver the instructions "3,2,1" and then blow the whistle to begin the sprint.

More research needs to be completed on the RIST on different ages, genders and skill levels to further establish the reliability and validity. The hope is that using this test will save coaches and players time testing anaerobic fitness rather then completing this in a lab setting which is sometimes long, complicated and not available to every population. This test seems to be a more reliable and valid on-ice test than others reported in the literature which require more calculations, are physically exhausting for the players, and take longer to complete which eventually enter the aerobic energy range, proving that these are not true anaerobic tests. The results of this study has shown that the RIST is a reliable, and valid tool to measure average peak power in ice hockey players on-ice in a sport specific setting.

REFERENCES

Atkinson, G., & Nevill, A.M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine, 26*(4), 217-238.

Ayalon, A., Inbar, O., & Bar-Or, O. (1974). Relationships among measurements of explosive strength and anaerobic power. In Nelson, R.C., & Morehouse, C.A., *International series on Sport Sciences*, vol. 1, Biomechanics 4, (pp.572-577). Baltimore: University Press.

Bar-Or, O. (1996). Anaerobic Performance. In Docherty, D., *Measurement in Pediatric Exercise Science* (pp. 161-182). Champaign, IL: Canadian Society for Exercise Physiology.

Bar-Or, O. (1987). The Wingate Anaerobic Test: An update on Methodology, Reliability and Validity. *Sports Medicine, 4*, 381-394.

Bar-Or, O., Dotan, R., & Inbar, O. (1977). A 30 second all-out ergometric test – it's reliability and validity for anaerobic capacity. *Irasel Journal of Medical Sciences, 13*, 126.

Baumgartner, T.A., & Chung, H. (2001). Confidence limits for intraclass reliability coefficients. *Measurement in Physical Education and Exercise Science, 5*(3), 179-188.

Berk, L.E. (2nd). (2001). *Development through the lifespan.* Boston, MA: Allyn and Bacon.

Beunen, G., & Malina, R.M. (1996). Growth and biological maturation: relevance to athletic performance. In O. Bar-Or (Ed.), The child and adolescent athlete (pp. 2-

24). Oxford, England: Blackwell Science.

Bosco, C., Luhtanen, P., & Komi, P.V. (1983a). A simple method for measurement of mechanical power in jumping. *European Journal of Applied Physiology, 50*, 273-282.

Bracko, M.R. (2001). On-ice performance characteristics of elite and non-elite women's ice hockey players. *Journal of Strength and Conditioning Research, 15*(1), 42-47.

Bracko, M.R., & Fellingham, G.W. (1997). Prediction of ice skating performance with off-ice testing in youth hockey players. *Medicine and Science in Sport and Exercise, 29*(5), S172.

Burr, J.F., Jamnik, R.K., Baker, J., Macpherson, A., Gledhill, N., & Mcguire, E.J. (2008). Relationship of physical fitness test results and hockey playing potential in elite-level ice hockey players. *Journal of Strength and Conditioning Research, 22*(5), 1535-1543.

Burr, J.F., Jamnik, V.K., Dogra, S., & Gledhill, N. (2007). Evaluation of jump protocols to assess leg power and predict hockey playing potential. *Journal of Strength and Conditioning Research, 21*(4), 1139-1145.

Canadian Society for Exercise Physiology. (2003). *The Canadian physical activity, fitness & lifestyle appraisal.* Ottawa: Canadian Society for Exercise Physiology.

Chandler, T.J., & Brown, L.E. (2008). *Conditioning for Strength and Human Performance.* Philadelphia, PA: Lippincott, Williams & Wilkins.

Cox, M.H., Miles, D.S., Verde, T.J., & Rhodes, E.C. (1995). Applied physiology of ice hockey. *Sports Medicine, 19*(3), 184-201.

Davies, C.T.M., & Young, K. (1984). Effects of external loading on short term power output in children and young male adults. *European Journal of Applied Physiology, 52*, 351-354.

Diakoumis, K., & Bracko, M.R. (1998). Prediction of skating performance with off-ice testing in deaf ice hockey players. *Medicine and Science in Sports and Exercise, 30,* S272.

Dotan, R., & Bar-Or, O. (1983). Load optimization for the Wingate Anaerobic Test. *European Journal of Applied Physiology, 51,* 409-417.

Dotan, R., & Bar-Or, O. (1980). Climatic heat stress and performance in the Wingate Anaerobic Test. *European Journal of Applied Physiology, 44,* 237-243.

Dreger, R.W., & Quinney, H.A. (1999). Development of a hockey-specific, skate treadmill VO2max protocol. *Canadian Journal of Applied Physiology, 24*(6), 559-569.

Dregor, R.W. (1997). Using skate-treadmills to train hockey players for speed. *Strength and Conditioning, 19*(6), 33-35.

Faught, B.E., Nystrom, M., & Montelpare, W.J. (2003). Precision and accuracy of an on-ice skating test to predict maximal oxygen capacity. *Canadian Journal of Applied Physiology, 28,* S52.

Gamble, F., & Montgomery, M.L. (1986). A cycling test of anaerobic endurance for ice hockey players. *Canadian Journal of Applied Sport and Science, 11*(3), 14P.

Geithner, C.A., Lee, A.M., & Bracko, M.R. (2006). Physical and performance differences among forwards, defensemen, and goalies in elite women's ice hockey. *Journal of Strength and Conditioning Research, 20*(3), 500-505.

Gledhill, N., & Jamnik, V. (1^{st}). (1994). *Detailed fitness and medical assessment* *protocols for NHL entry draft players*. Toronto: York University.

Glencross, D.J. (1966). The nature of the vertical jump test and the standing broad jump. *Research Quarterly, 37*, 353-359.

Granier, P., Mercier, B., Mercier, J., Anselme, F., & Prefaut, C. (1995). Aerobic and Anaerobic contribution to Wingate test performance in sprint and middle-distance runners. *European Journal of Applied Physiology, 70,* 58-65.

Gray, R.K., Start, K.B., & Glencross, D.J. (1962). A test of leg power. *Research Quarterly, 33*, 44-50.

Green, H.J. (1979). Metabolic aspects of intermittent work with specific regard to ice hockey. *Canadian Journal of Applied Sport Science, 4*(1), 29-34.

Green, H.J. (1978). Glycogen depletion patterns during continuous and intermittent ice skating. *Medicine and Science in Sports, 10*(3), 183-187.

Green, H.J., Bishop, P., Houston, M., Mckillop, R., Norman, R., & Stothart, P. (1976). Time-motion and physiological assessments of ice hockey performance. *Journal of Applied Physiology, 40*(2), 159-163.

Hopkins, W.G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine, 30*(1), 1-15.

Houston, M.E., & Green, H.J. (1976). Physiological and anthropometric characteristics of elite Canadian ice hockey players. *Journal of Sports Medicine, 16,* 123-128.

Inbar, O., Bar-Or, O., & Skinner, J.S. (1996). *The Wingate Anaerobic Test.* Champaign, Illnois: Human Kinetics.

Kalamen, J.L. (1968). Measurement of maximum power in man. Unpublished Ph.D. thesis, Ohio State Univeristy, Columbus, OH.

Katch, V.L., Weltman, A., Martin, R., & Gray, L. (1977). Optimal test characteristics for maximal anaerobic work on the bicycle ergometer. *Research Quarterly, 48*(2), 319-327.

Lariviere, G., & Godbout, P. (1976). Measuring physical fitness and technical efficiency of ice hockey players: reference values for various age categories. Quebec: Editions du Pelican.

Lariviere, G., Lavallee, H., & Shepard, R.J. (1976). A simple skating test for ice hockey players. *Canadian Journal of Applied Sport Science, 1,* 223-228.

Leone, M., Léger, L., Lariviere, G., & Comtois, A.S. (2007). An on-ice aerobic maximal multistage shuttle skate test for elite adolescent hockey players. *International Journal of Sports Medicine, 28,* 823-828.

Looney, M.A. (2000). When is the intraclass correlation coefficient misleading? *Measurement in Physical Education & Exercise Science, 4*(2), 73-78.

Lyman, H.B. (6^{th}). (1998). *Test Scores and what they mean.* Boston, Ma: Allyn and Bacon.

Margaria, R., Aghemo, P., & Rovelli, E. (1966). Measurement of muscular power (anaerobic) in man. *Journal of Applied Physiology, 21*(5), 1662-1664.

Mascaro, T., Seaver, B.L., & Swanson, L. (1992). Prediction of skating speed with off-ice testing in professional hockey players. *Journal of Orthopaedics and Sports Physical Therapy, 15*(2), 92-97.

Mastrangelo, M.A., Chaloupka, E.C., Kang, J., Lacke, C.J., Angelucci, J., Martz, W.P., & Biren, G.B. (2004). Predicting anaerobic capabilities in 11-13 year old boys. *Journal of Strength and Conditioning Research, 18*(1), 72-76.

Maud, P.J., & Schultz, B.B. (1989). Norms for the Wingate anaerobic test with comparisons in another similar test. *Research Quarterly of Exercise and Sport, 60*(2), 144 –151.

Maud, P.J., & Schultz, B.B. (1986). Gender comparisons in anaerobic power and capacity tests. *British Journal of Sports Medicine, 20*, 51-54.

Mayhew, J.L., Piper, F.C., Etheridge, G.L., Schwegler, T.M, Beckenholdt, S.E., & Thomas, M.A. (1990). The Margaria-Kalamen Anaerobic Power Test: Norms and Correlations. *Journal of Human Movement Studies, 18,* 141-150.

Montgomery, D.L. (1988). Physiology of ice hockey. *Sports Medicine, 5*, 99-126.

Morrow, J.R., Jackson, A.W., Disch, J.G., & Mood, D.P. (3rd). (2005). *Measurement and evaluation in human performance.* Champaign, Il: Human Kinetics.

Petrella, N.J., Montelpare, W.J., Nystrom, M., Plyley, M., & Faught, B.E. (2007). Validation of the FAST skating protocol to predict aerobic power in ice hockey players. *Applied Physiology, Nutrition and Metabolism, 32*, 693-700.

Powers, S.K., & Howley, E.T. (5th). (2004). *Exercise Physiology: Theory and Application to Fitness and Performance.* New York, NY: McGraw-Hill.

Quinney, H.A., Dewart, R., Game, A., Snydmiller, G., Warburton, D., & Bell, G. (2008). A 26-year physiological description of a National Hockey League team. *Applied Physiology, Nutrition and Metabolism, 33*, 753-760.

Reed, A., Hansen, H., Cotton, C., Gauthier, R., Jétte, M., Thoden, J., & Wenger, H. (1979). Development and validation of an on-ice hockey fitness test. *Canadian Journal of Applied Sport Science, 4*(4), 245.

Rhodes, E.C., Cox, M.H., & Quinney, H.A. (1987). Comparison of National Hockey

League players on selected physiological parameters. *Medicine and Science in Sport and Exercise, 19*(2), S46.

Rhodes, E.C., Cox, M.H., & Quinney, H.A. (1986). Physiological monitoring of National Hockey League regulars during the 1985-86 Season. *Canadian Journal of Applied Sport Science, 11*(3), 36P.

Rhodes, E.C., Mosher, R.E., & Potts, J.E. (1985). Anaerobic capacity of elite pre pubertal ice hockey players. *Medicine and Science in Sport and Exercise, 17,* 265.

Safrit, M.J. (2nd). (1990). *Introduction to measurement in physical education and exercise science.* St. Louis, Missouri: Times Mirror / Mosby College Publishing.

Sawka, N.S., Tahamont, M.V., Fitzgerald, P.I., Miles, D.S., & Knowlton, R.G. (1980). Alactic capacity and power: reliability and interpretation. European *Journal of Applied Physiology, 45*, 109-116.

Sayers, S.P., Haracjiewicz, D.V., Harman, E.A., Frykman, P.N. & Rosenstein, M.T. (1999). Cross-validation of three jump power equations. *Medicine and Science in Sports and Exercise, 31*(4), 572-577.

Smith, D.J., Quinney, H.A., & Steadward, R.D. (1982). Physiological profiles of the Canadian Olympic Hockey Team (1980). *Canadian Journal of Applied Sport Science, 7*(2), 142-146.

Stevens, G.H.J., & Wilson, B.W. (1986). Aerobic contribution to the Wingate test. *Medicine and Science in Sport and Exercise, 18*, S2.

Tharp, G.D., Newhouse, R.K., Uffelman, L., Thorland, W.G., & Johnson, G.O. (1985). Comparison of sprint and run times with performance on the Wingate Anaerobic Test. *Research Quarterly of Exercise and Sport, 56*, 73-76.

Tirosh, E., Rosenbaum, P., Bar-Or, O. (1987). The reliability and feasibility of an anaerobic muscular evaluation in physically disabled children. *Developmental Medicine and Child Neurology*, In Press.

Twist, P., & Rhodes, E.C. (1993). A physiological analysis of ice hockey positions. *National Strength and Conditioning Association Journal, 15*(6), 44-46.

Vainikka, M., Rahkila, P., & Rusko, H. (1982). Physical performance characteristics of the Finnish National Ice Hockey Team. In Komi, P.V., *Exercise and Biology* (pp. 158-165). Champaign, Illnois: Human Kinetics.

Vandewalle, H., Peres, G., & Monod, H. (1987). Standard Anaerobic Exercise Tests. *Sports Medicine, 4*, 268-289.

Vescovi, J.D., Murray, T.M., & Vanheest, J.L. (2006a). Positional profiling of elite ice hockey players. *International Journal of Sports Physiology and Performance, 1,* 84-94.

Vescovi, J.D., Murray, T.M., Fiala, K.A., & Vanheest, J.L. (2006b). Off-ice performance and draft status of elite ice hockey players. *International Journal of Sports Physiology and Performance, 1,* 207-221.

Vincent, J. (1994). *Statistics in Kinesiology.* Champaign (IL): Human Kinetics Books.

Watson, R.C., & Sargeant, T.L. (1986). Laboratory and On-Ice Test Comparisons of Anaerobic Power of Ice Hockey Players. *Canadian Journal of Applied Sport Science, 11*(4), 218-224.

APPENDICES

Appendix A: Consent Form

"The reliability and validity of an anaerobic field test in ice hockey players."

In signing this consent form, I agree to participate in this study which will test various aerobic and anaerobic fitness measures of hockey players. I have been clearly informed that the risks involve being physically exhausted during some of the tests while the benefits include determining my fitness levels and identifying strengths and weaknesses. I understand the nature of the study and have understood the potential benefits and risks associated with participation in the study.

I acknowledge the information that I provided on this form and other health forms is accurate. I understand that any information collected about me will be kept confidential and I will not be identified if the results are published in the future. The results of the study are available to me upon request after the study is completed. I also acknowledge that the results will be made available to the coach upon request and that the results will be stored at Lakehead University for a period of 5 years as per policy.

I also understand that since I am under the age of 18, I am therefore not of legal consent and signature of legal guardian/parent is required to participate in this study. I understand that part of the on-ice test component may be videotaped and that I can chose to be or not to be filmed. I will not be identified in any way from the filming as the video will be made available only the researchers. The video will be shot from afar as to not specifically focus on any one person performing the test.

As a volunteer, I understand that I have the right to withdraw from the study at any given time even after signing this form.

Participant's Signature　　　　　　　　　　　　　　　　　　　Date

Parent / Legal Guardian's Signature　　　　　　　　　　　　　　Date

As a researcher, I have explained the nature of this study and provided the necessary information for the participant to chose their participation in this study or not.

Researcher's Signature　　　　　　　　　　　　　　　　　　　Date

Appendix B: PAR-Q Participation Form

The PAR-Q Physical activity readiness questionnaire

Surname: ______ First Name ______

Are you currently experiencing a condition that prevents you from performing activities of daily living?

☐ No ☐ Yes

In your lifetime, have you been injured to an extent that you could not perform activities of daily living?

☐ No ☐ Yes

☐ Yes ☐ No Has a doctor ever said that you have heart trouble?

☐ Yes ☐ No Have you experienced pains in your heart or chest?

☐ Yes ☐ No Do these pains in your heart or chest occur frequently?

☐ Yes ☐ No Have you felt faint or have you had spells of dizziness?

☐ Yes ☐ No Has a doctor ever said that your blood pressure was too high?

☐ Yes ☐ No Has a doctor ever said that you you have a bone or joint problem such as arthritis?

☐ Yes ☐ No Is this problem typically aggravated by exercise?

☐ Yes ☐ No Is there a reason why you could not be involved in a regular exercise program?

Explain the reason: ______

______
Signature of Participant

______
Date

______
Signature of Witness

______
Date

Appendix C: Maximal Exercise Test Form

Measuring maximal anaerobic capacity using the Wingate Bicycle ergometer.

I, ______________________________ consent to participate in this exercise which will require the measurement of maximal anaerobic capacity using the Wingate Bicycle ergometer.

I understand the following:

a) That there is very little risk of injury associated with testing for healthy individuals. There will be trained "spotters" assisting the Wingate bicycle test.
b) That there may be some physical discomfort due to exercising to the point of temporary exhaustion.
c) That I am obligated to immediately inform a member of the research staff (research assistants or supervisor) of any unusual pain, discomfort, fatigue or any other symptom(s) that I incur during or after the testing.
d) That I can withdraw from the testing at any time.
e) Individual data will be kept confidential. Publication of results will not reveal my subject identity since subjects will be referenced by number.

_________________________________ ____________________

Signature of Participant Date

_________________________________ ____________________

Signature of Parent/Guardian Date
(If Participant is under the age of 18 years)

_________________________________ ____________________

Signature of Witness Date

2. I have explained the nature of the testing to the athlete and believe he/she has understood it.

_________________________________ ____________________

Signature of Researcher Date

Appendix D: Cover Letter

Letter of Introduction

Dear Player and Parent / Guardian,

Thank you for your consideration to participate in the study "the reliability and validity of an anaerobic field test in ice hockey". This study will examine anaerobic (peak or sprint) fitness measures in the lab and on-ice. The purpose of the study is to validate an anaerobic on-ice field test with 3 lab tests: Vertical Jump, Stair Running Test and the Wingate Cycle Test. The total number of tests completed will be 4. These tests will give you a good idea of your level of short-term fitness.

By participating in this study, you will be required to fill out various forms which will ask you about your current level of health and physical fitness. These forms are:
-Consent form
-Physical Activity Readiness Questionnaire (PAR-Q)
-Maximal Exercise Test Form
Once ready to participate, you will be perform the following:
-Four fitness tests twice in total. The 4 tests will be completed in one week on 2 separate days and then will be retested 4-5 days later in the same fashion.

All fitness measurements taken will be in the supervision of trained researchers and supervisors from Lakehead University. All data collected during this study will be strictly confidential and will be made available only to the researchers and head coach if requested. At some point during the on-ice testing, certain segments of the fitness tests may be videotaped for analysis. All participants will sign off in the consent form about appearing on film or not. This videotape along with the measured data will not be released to anyone besides the main researchers and stored in Lakehead University for a 5-year period as per policy.

There are few risks physically or psychologically with this study. Performing some of these fitness tests will be physically demanding but not enough to cause harm to the participants. The benefits of this study will be: measure and evaluate your teams fitness levels which will identify strengths and weaknesses and help design off-season training programs.

Should any additional information or questions about the study be required, please contact the Lakehead University School of Kinesiology. We appreciate your willingness and co-operation to participate in this study.

Sincerely,

Dr. William Montelpare (Phone: 807-343-8481) Email: wmontelp@lakeheadu.ca
Allan Power (Phone: 807-344-9718) Email: acpower@lakeheadu.ca

Appendix E: Table 12

Table 12
Intra-class Correlation Coeffecients (r) for Bantam Kings tests measuring Peak Power

Test	Peak Power (watts)
RIST	0.97
Wingate	0.84
MK Stair	0.81
Vertical Jump	0.91

Appendix F: Table 13

Table 13
Pearson Correlation results for Bantam Kings tests measuring peak power in watts

Test Correlations	Correlation Values
RIST vs. Vertical Jump	0.49
RIST vs. MK Stair	0.86
RIST vs. Wingate	0.85

Appendix G: Figures 5, 6, 7

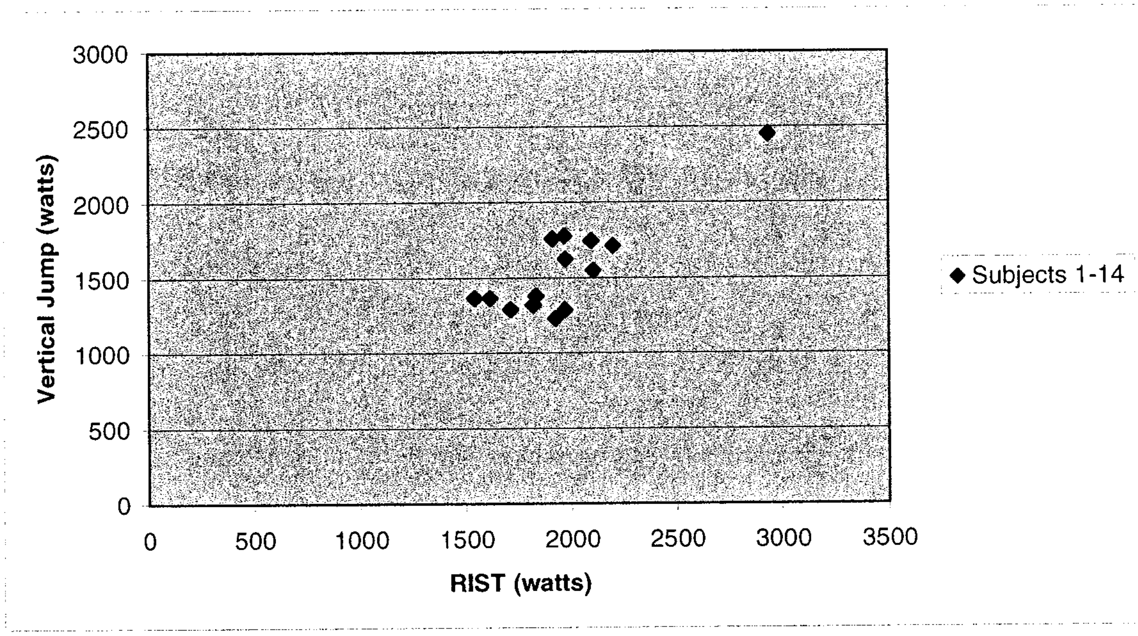

Figure 5. Scatter plot of Pearson product-moment correlation for RIST and Vertical Jump measuring average peak power and peak power in watts.

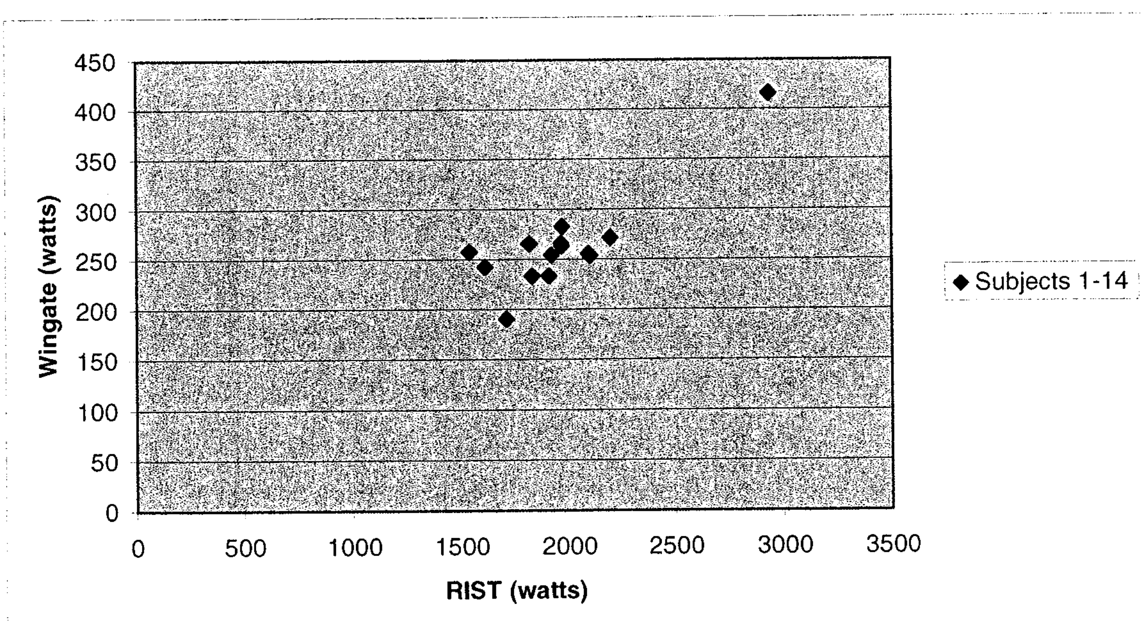

Figure 6. Scatter plot of Pearson product-moment correlation for RIST and MK Stair measuring average peak power and peak power in watts.

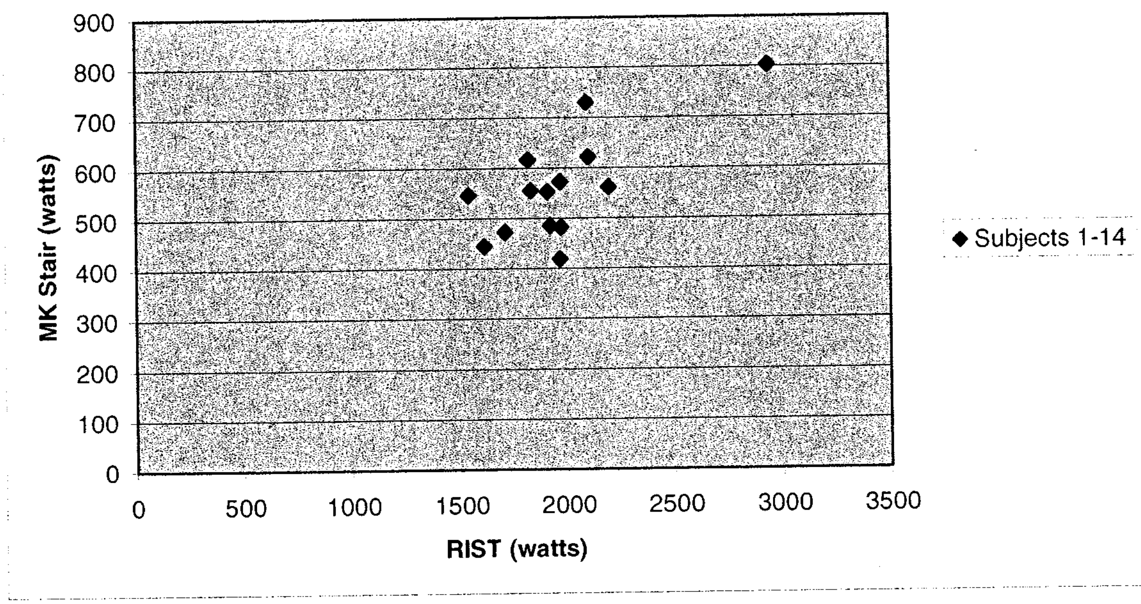

Figure 7. Scatter plot of Pearson product-moment correlation for RIST and Wingate measuring average peak power and peak power in watts.