

# Lack of Predictive Power in Commonly Used Tests for Performance in Alpine Skiing



## Authors

Robert Nilsson<sup>1</sup>, Apostolos Theos<sup>1</sup>, Ann-Sofie Lindberg<sup>1, 2</sup>, Richard A. Ferguson<sup>3</sup>, Christer Malm<sup>1</sup>

## Affiliations

- 1 Department of Community Medicine and Rehabilitation, Section of Sports Medicine, Umeå University, Sweden
- 2 Winternet, Winternet, Boden, Sweden
- 3 School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, United Kingdom of Great Britain and Northern Ireland

## Key words

exercise testing, athletic performance, FIS points, multivariate statistics,  $VO_2\max$

received 02.03.2019

revised 26.11.2019

accepted 27.11.2019

## Bibliography

Sports Medicine International Open 2021; 5: E28–E36

DOI 10.1055/a-1078-1441

ISSN 2367-1890

2021. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14,  
70469 Stuttgart, Germany

## Correspondence

Robert Nilsson  
Department of Community Medicine and Rehabilitation,  
Sports Medicine Unit Umea University  
Gösta Skoglunds väg 33  
90187 Umeå  
Sweden  
Tel.: +4690 786 50 00, Fax : +4690 13 56 92  
[robert.nilsson@umu.se](mailto:robert.nilsson@umu.se)

## ABSTRACT

Competitive alpine skiing is a complex sport that requires high physical and technical competence. Testing the physical status of athletes may be important to increase their ability to achieve elite sport-specific performance. This study aimed to investigate the predictive power of the national test battery of the Swedish Olympic Committee (Fysprofilen) and anthropometric variables in the prediction of competitive performance of elite alpine skiers, indicated by Fédération Internationale de Ski points. Data from fourteen Swedish elite female alpine skiers were analyzed using bivariate and multivariate statistical methods. Physiological test results and anthropometric data could not generate significant bivariate or multivariate models for prediction of competitive performance. Multivariate regression ( $R^2$ ) and prediction ( $Q^2$ ) models for Fédération Internationale de Ski Slalom and Giant Slalom rank reached  $R^2 = 0.27$  to  $0.43$ ,  $Q^2 = -0.8$  to  $-0.17$ , indicating no valid models. The overall interpretation of these and previous findings are that future test batteries must be validated before implemented, and that test results should be treated with caution when it comes to prediction of future competitive results. Applying tests that are not validated against competitive performance risk misleading coaches and training advisors who aim to increase the sports-specific performance of the individual athlete.

## Introduction

Competitive alpine skiing is a complex and challenging outdoor sport, requiring a high level of physical and technical proficiency [1, 2]. Alpine ski racing consists primarily of four different disciplines – slalom (SL), giant slalom (GS), super giant slalom (SG), and

downhill (DH) – which all differ from each other regarding speed, race duration, turning radius, and vertical offset between the gates [3, 4]. The technical disciplines (SL and GS) are often performed on relatively steep terrain with frequent and short turns [5–8], resulting in low to moderate speeds ( $20\text{--}90\text{ km}\cdot\text{h}^{-1}$ ) and race durations

of 45–90 s. The speed disciplines (SG and DH) usually are conducted on longer slopes that result in higher speeds (70–160 km · h<sup>-1</sup>) and race durations of 90–150 s [4, 9, 10]. Regardless of discipline, competitive skiing places the athletes under immense physical strains. In GS, for example, skiers must withstand snow reaction ≥ 2000 N during a regular turn [11, 12]. Others who have monitored muscle activity in both leg and lumbopelvic muscle groups have indicated that muscular workloads ranging between 50–280% of maximum voluntary isometric contraction (MVC) can be reached during large parts of the turn in the examined discipline [13]. In addition to isometric and eccentric muscle strength, physiological qualities such as maximum oxygen uptake [14–16], explosiveness [3, 17], and coordination [18] have been suggested as critical physiological qualities for elite skiing performance.

To be successful in this sport (and given the fact that the difference between first and fifth/sixth placement is often only a few hundredths of a second [19]), the ability to maintain high levels of physical performance is important [20]. Although it has been suggested that performance in alpine skiing is not limited by individual physiological qualities [15, 18], a quantification of variables that can predict future performance can be of great importance. Numerous test protocols have been used over the years [10, 14, 21–30], many of which attempt to evaluate the skier's physical status before and after the racing season. One of these protocols is the so-called Fysprofilen (Swedish for “Physical Profile”), a test battery initially developed by the Swedish Olympic Committee (SOC). The Fysprofilen involves tests aimed at evaluating muscle strength, power, and aerobic and anaerobic performance. The intent is to provide comparable values between different sports and to provide individual athletes with an indication of their physical status in relation to the physical demands of their sport [31]. According to the SOC, the highest achievable score in Fysprofilen (10 on a scale 1–10) is also a strong indication of what is optimal for being on an international elite level in a particular sport [31]. The results of each test converts into points and are intended to show how athletes differ between sports and the physiological requirements necessary to be successful in a competitive context. Thus, each sport, including alpine skiing, has its specific “profile,” however, the predictive power of Fysprofilen for competitive performance, particularly in relation to alpine skiing, has not yet been tested.

To be of real benefit, a physiological test battery must be reliable and valid to the intended measurement [32]. Commonly used physiological test batteries have been criticized for lack of correlation between test results and sport-specific performance [10]. Published studies including true athletic performance, such as race results or ranking [10, 14, 15, 22, 33, 34], often use inappropriate statistical methods and have not been validated. In addition, none has used multivariate statistical methods to predict competitive race performance based on a combination of test results. Consequently, the predictive power based on such test results, or test battery, is low. Therefore, the aim of this study was two-fold: first, using multivariate data analysis (MVDA) to investigate if physiological test results from repeated testing on elite female alpine skiers can be used to predict future performance; and second, to examine the combined point score from the test battery Fysprofilen as a predictor of race performance (Fédération Internationale de Ski (FIS) ranking) over time.

► **Table 1** Anthropometric data and physiological test results

Variable	(n = 14)
Stature (cm)	171 ± 5.4
Weight (kg)	67.4 ± 3.6
Squats (kg)	111.5 ± 10.7
Bench press (kg)	65.5 ± 6.3
Handgrip strength L (kg)	46.3 ± 4.2
Handgrip strength R (kg)	49.3 ± 5.3
Cleans (kg)	80.2 ± 5.9
Harre 's test (s)	10.3 ± 0.5
SJ (cm)	34.3 ± 4
CMJ (cm)	36.2 ± 3.9
CMJa (cm)	41.2 ± 4.6
20 m (s)	3.2 ± 0.1
Pull-ups (n)	10.1 ± 4.7
Brutal bench (n)	29.6 ± 6
$\dot{V}O_{2peak}$ mL/(kg · min)	49.6 ± 2.7

SJ, squat jump; CMJ, countermovement jump; CMJa, countermovement jump with arm swing;  $\dot{V}O_{2peak}$ , peak volume of oxygen uptake. Data are mean ± standard deviation.

## Material and Methods

### Participants

A total of fourteen senior elite female alpine skiers (n = 14), aged 18–30 years, volunteered to participate in the study (► **Table 1**). All skiers in this study were, at the time of testing, members of the Swedish national alpine ski team and the majority of them with a top-ranking world status in the investigated disciplines (SL and GS). All participants underwent a brief medical examination and completed a questionnaire regarding general health status. The exclusion criteria for participation were any disease, injury, or use of medication that could affect their health or performance. Participants were fully informed of the risks and discomforts associated with all experimental trials before providing written, informed consent.

### Ethical statement

The local ethics committee for the university granted ethical permission 2016-260-31M, and the study was conducted in accordance with the World Medical Association Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects (2008) and the ethical standards of the International Journal of Sports Medicine [35].

### Experimental protocol

Participants attended the testing facilities at the Section of Sports Medicine on at least five occasions over the course of two days. Visits were conducted in close proximity to the previous or upcoming race season, corresponding to the periods from April to May and August to October. For 24 h prior to each visit, participants were asked to avoid all strenuous physical activity and to follow the same routine regarding sleep and nutritional intake.

For alpine skiing, the Fysprofilen test battery is composed of 13 different physiological tests that are divided into four main categories: strength, power, aerobic, and anaerobic. The strength category includes tests of one repetition maximum (1RM) in squats, bench press, handgrip strength, and the maximum number of repetitions (MR) in pull-ups. The power category includes 1RM in cleans, change of direction (COD) ability in Harre's test and maximum jump height in squat jump (SJ), countermovement jump (CMJ) and countermovement jump with arm swing (CMJa). The aerobic category includes measurement of peak oxygen uptake ( $\dot{V}O_{2\text{peak}}$ ) and 3000-m running capacity. The anaerobic category includes MR in brutal bench test and sprint performance over a distance of 20 m. Each test result in all four main categories yields a certain score, which in turn are added together to provide a test point value for that part of the test. Finally, these values are counted together and thus provide a generalized point value for the entire test battery, scored 0–10 (Fysprofilen).

No detailed information on how the physiological test points or the Fysprofilen Index are calculated has been reported to the authors of this study and nor, to the best of our knowledge, has it been publicly disclosed elsewhere. To calculate the score points and Fysprofilen Index, results must be calculated using the Fysprofilen official website ([www.fysprofilen.se](http://www.fysprofilen.se)) or be submitted to a Fysprofilen representative. The calculated Fysprofilen results are then reported back to the users. For details about Fysprofilen, we refer readers to the official homepage at <http://www.fysprofilen.se/sv/default.aspx>.

## Testing procedures

All physical tests were conducted in accordance with the guidelines for Fysprofilen and are categorized according to the SOC definition of physical capacity tested [36]. Prior to any physiological testing, participants conducted a ~15-minute standardized warm-up including easy jogging or cycling as well as dynamic stretching. Furthermore, before the start of each specific test, a self-regulated, exercise-specific (increasing load or alternative, light exercise) warm-up was conducted. During each test, participants were provided with verbal encouragement to ensure exercises were performed with maximal effort (e. g., 1RM), or to volitional exhaustion (e. g.,  $\dot{V}O_{2\text{peak}}$ ).

## Anthropometric tests

Body mass was measured to the nearest 0.1 kg using a standard weight scale (Soehnle weighing scale; Leifheit AG, Nassau, Germany), and body stature was measured to the nearest 0.1 cm using a wall-mounted scale (Fosamax stadiometer; Merck & Co. Inc., Kenilworth, NJ, USA).

## Strength tests

All 1RM strength and power tests were conducted using calibrated competition weights and barbells (Eleiko Sport, Halmstad, Sweden). When testing 1RM, the load was incrementally increased following a successful attempt (suggested increase in load was as follows (set  $\times$  reps/percent of 1RM): 1  $\times$  8/50%, 1  $\times$  6/60%, 1  $\times$  5/70%, 1  $\times$  3/80%, 1  $\times$  2/90%, 1  $\times$  1/100, 1  $\times$  1/102%) until failure or voluntary discontinuation of more attempts at higher load. The recovery

period between each test was  $\geq 5$  min. The highest 1RM for each conducted test was registered.

Squat 1RM (kg) was performed from an upright standing position with the test being valid when the participant managed to descend to full depth, with the upper thigh being parallel with the floor, and back up again. Bench press 1RM (kg) was performed in a supine position on a bench with the test being valid when the participant managed to descend the barbell from fully extended arms, down to the chest and up again without bouncing on the chest or by raising the back of the seat.

The maximum number of pull-up repetitions was assessed with a grip slightly wider than shoulder-width and hands facing away from the participant. The test started in a hanging position with straight arms, and a repetition was approved when the participant successfully pulled herself until her chin was above the bar, without jerks or modifying the grip.

Handgrip strength (kg) test was assessed using a calibrated isometric grip dynamometer (T.K.K. 5401 GRIP D; Takei Scientific Instruments Co., Ltd., Niigata, Japan) in an upright standing position, with the arm slightly bent at the practitioner's side, not touching the body. The participant was instructed to squeeze the handgrip dynamometer with as much force as possible during 3 s. The best performance out of three on each hand was registered.

## Power tests

The clean 1RM (kg) test was performed with the barbell either on the floor or lifted at the height just below the knees (hang cleans). The test was valid when the participant managed to lift the barbell in an explosive motion from the starting position up to the chest and catch it on the chest in a stable standing position.

Harre's test was performed as previously described by Hoyek, Champely, Collet, et al. [37]. Briefly, the test combines a variety of coordinative properties with speed, explosiveness, and agility. The tests include a somersault, acceleration, direction changes, and rapid passages over and under obstacles over a course deployed as a cross. The best performance time (s) out of three trials was registered with the use of validated infrared timing systems: the PF MuscleLab MA4020e (Ergotest Innovation AS, Porsgrunn, Norway) [38], and the IVAR Jump and Sprint system (Spin Test, Tallinn, Estonia) [39, 40].

The squat jump (SJ, cm) was performed from a standing position with knees bent at 90° flexion, hips flexed, and hands placed on the hips. The test was approved when the participant successfully, on the instructor's command and without remission of the knee angle, jumped as high as possible and landed with normal knee flexion in the same location as the take-off. The best maximal jumping height (cm) out of three trials was registered with the use of validated infrared-mat-based systems (PF MuscleLab MA4020e and IVAR Jump and Sprint system).

Countermovement jumps (cm) were performed from a standing position. For CMJ, the hands were placed on the hips during the whole test, and for CMJa the arms could be used freely. Both the CMJ and the CMJa tests were approved when the participant jumped as high as possible and landed with normal knee flexion in the same location as the take-off. The best maximal jumping height (cm) out of three trials was registered using the same equipment described for the SJ test.

## Aerobic tests

$\dot{V}O_{2peak}$  was determined on either a cycle ergometer (Monark 839E, Varberg, Sweden) or a treadmill (Rodby Innovation, Vänge, Sweden). Two different testing protocols were used because some participants experienced back pain when running. These participants were allowed to carry out the test using the cycle ergometer. For the cycle ergometer test participants began cycling at a starting load of 100W and a constant pedal cadence of 60 rpm, after which power increased 30 W every 1 min until volitional exhaustion or when cadence fell 10% below the required cadence for more than 5 sec, despite strong verbal encouragement. For the treadmill running test, participants began running at a fixed treadmill speed of 12–13.5 km · h<sup>-1</sup>. During the first 3 min, the treadmill incline increased by 1° every minute after which the slope increased by 0.5° every minute until volitional exhaustion, despite strong verbal encouragement.

Pulmonary gas exchange was continuously measured throughout the test with an Oxycon Pro Jaeger System (CareFusion, Hoechberg, Germany) set on mixing chamber mode and calibrated according to the manufacturer's instructions. The highest mean for  $\dot{V}O_{2peak}$  during a 20-s recording was considered maximum and registered as  $\dot{V}O_{2peak}$ .

## Anaerobic tests

The brutal bench is a test resembling sit-ups [41]. Briefly, the participants hang upside down in a vertical position with their backs on a backrest and their legs at a 90-degree knee angle. A repetition was approved when the practitioner was able to get up from the starting position, touched their knees with their elbows without bouncing against the backrest at the starting position. The maximum number of brutal benches performed was registered.

The 20-m running sprint test was performed as fast as possible, from a standing start. The same equipment as for Harre's test was used. Performance time (s) was registered.

## Fédération Internationale de Ski points

Slalom and GS FIS points used in this study were collected twice, in December (6<sup>th</sup> list) and April (11<sup>th</sup> list) and correlated with the pre-season testing in April to May and September to October. Lists were selected because they represented the most significant change in FIS ranking over a season. Fédération Internationale de Ski ranking was used as a measurement of competitive performance (Y-variables) in the statistical analysis. For further reading on the FIS points system, we refer the reader to the alpine FIS points rules found on the official FIS homepage for alpine skiing [42] and [43].

## Statistical analysis

In sports research, validity and reliability studies usually apply bivariate and multivariate linear regression methods [44]. Due to the complex nature of alpine skiing [1, 10, 20] and the non-linear relationship between test results and performance, traditional statistical procedures are not optimal for accurately predicting future performance. Therefore, this study used multivariate statistical modeling or projection analysis, which has proven useful in other similar contexts [43, 45]. This statistical method can be especially useful to increase the statistical power and facilitate the subsequent interpretation of the statistical outcome when collected data is composed of a small

number of subjects, or there is a high correlation between two or more independent variables [46]. Furthermore, with the use of this method, data can easily be cross-validated by permutation to estimate the possible prediction error, thus providing additional information about the reliability of the models.

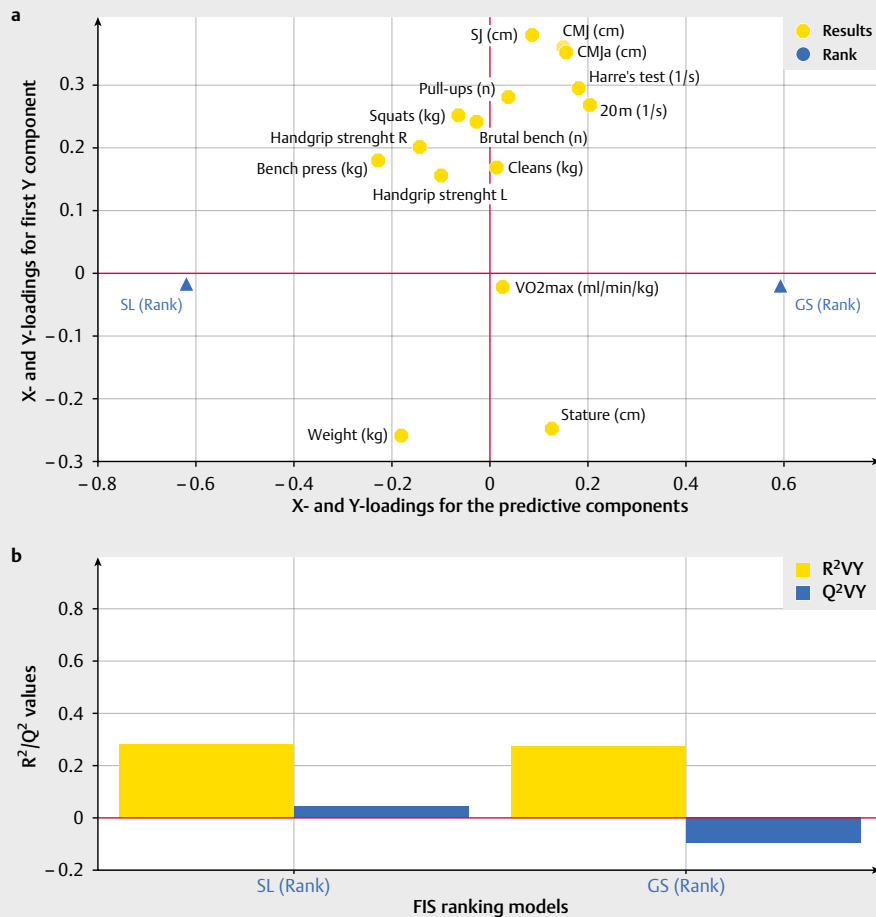
Orthogonal projections to latent structures (OPLS) and principal component analysis (PCA) were used to investigate regression and prediction of ranking from all anthropometric and physiological test results. Data used in the statistical analysis are compiled of Fysprofilen and anthropometric variables (X-variables) and FIS rankings (Y-variables) between the years 2012 and 2014. Data were analyzed using SIMCA 15.1 (Sartorius Stedim Data Analytics AB, Umeå, Sweden), as in previous publications by our group [43, 45]. A Mann-Whitney two-sample rank-sum test was conducted to examine whether there were significant differences in the Fysprofilen Index between the levels of the lowest and highest SL ranking and the lowest and highest GS ranking, respectively. Shapiro-Wilk tests were conducted to determine whether the distribution of included physiological test variables was significantly different from a normal distribution. Data were considered normally distributed if  $p > 0.05$ ; however, the data for pull-ups had distributions that differed significantly from normality: ( $W = 0.94$ ,  $p = 0.004$ ) and were, therefore,  $\log_{10}$ -transformed to normality. The Mann-Whitney two-sample rank-sum test, tests for normality, and  $\log_{10}$  transformation were conducted using JMP 14.0 (SAS Institute Inc., Cary, NC, USA). Running capacity for 3000 m was excluded from the statistical analysis due to a large amount of missing data (e. g., due to sickness or injury).

## Results

Anthropometric and physiological test results are presented in ► **Table 1**. No multivariate model could accurately predict (i. e., low  $Q^2$ ) competitive performance in slalom (SL) or giant slalom (GS) of senior elite female alpine skiers based on included physiological test results (► **Fig. 1a**). The model ( $R^2$ ) for FIS ranking with physiological test results by OPLS (► **Fig. 1a**) indicates separated clustering of FIS points not related to any of the analyzed variables (not located in the same area of the plot). The  $R^2/Q^2$  plot indicates a low predictive power ( $R^2 = 0.27$  to  $0.43$ ,  $Q^2 = -0.8$  to  $-0.17$ ) for both slalom and giant slalom (► **Fig. 1b**). Cross-validation by permutation (► **Fig. 2a–b**) confirms the low predictive power demonstrated in ► **Fig. 1a**. Consensus: None of the included physiological tests are important for FIS points in any of the disciplines at any time.

A PCA loading plot (► **Fig. 3**) indicates that physiological test results and corresponding scoring points from Fysprofilen are interchangeable (specific physiological test results and their corresponding scoring points are located in the same area of the plot). Consensus: Specific physiological test results and their corresponding score points from Fysprofilen are highly correlated and can, therefore, be used interchangeably.

The Mann-Whitney  $U$  test was not significant for any of the analyzed disciplines (SL;  $U = 99.5$ ,  $z = -0.07$ ,  $p = 0.945$ ; GS;  $U = 80$ ,  $z = -0.83$ ,  $p = 0.408$ ). This suggests that the distribution of the Fysprofilen Index for the Lowest group was not significantly different from the distribution of the Fysprofilen Index for the Highest category in either SL or GS. These results are illustrated by the box



► **Fig. 1 a-b.** Correlations between physiological test results and SL and GS ranking. Data from fourteen senior elite female alpine skiers between the years 2012 and 2014. **a**) Orthogonal partial least squares (OPLS) loading scatter plot ( $X = 15$ ,  $Y = 2$ ;  $n = 14$ ) visualizes correlations between variables. Physiological test results and FIS-rankings located in the same or opposite part of the loading plot are correlated. The horizontal axis displays the X- and Y-loadings of the predictive component, and the vertical axis the X- and Y-loadings for the first Y-orthogonal component. A high value ( $\max = 1$ ) means that the component is aligned with the original variable, a value close to zero shows that it has no influence. A low value ( $\min = -1$ ) indicates an opposite influence. **b**) The X/Y overview plot shows the models' cumulated  $R^2$  (variance explained) and  $Q^2$  (variance predicted) values. The plot indicates a low predictive power of the models ( $Q^2 < 0.1$ ).

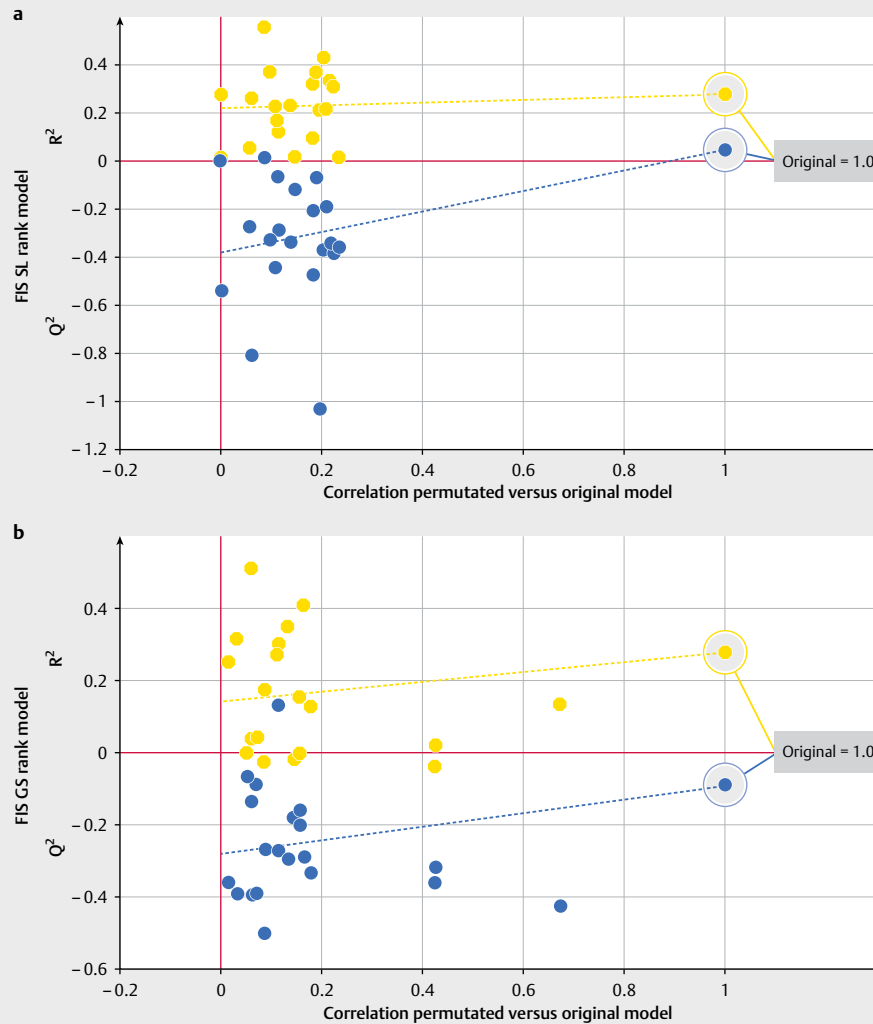
plot and paired dot-line plot (► **Fig. 4**), which demonstrates similar median values between groups and a substantial individual variation in the Fysprofilen Index for senior elite female alpine skiers. Consensus: The Fysprofilen Index and competitive performance over time are independent due to a substantial individual variation in the Fysprofilen Index, regardless of skiing performance.

## Discussion

The main finding of this study indicates that the physiological variables determined as part of the Fysprofilen test battery cannot predict future performance among senior elite female alpine skiers on a group level. These results are consistent with findings in a similar study by Impellizzeri, Rampinini, Freschi, et al. [47], where no correlation between physiological tests and competitive performance could be identified, but in contrast to others where correlations between skiing performance (time-trial and ranking) and results of physical test were demonstrated [10, 15, 21, 33, 34, 48]. However,

the results from these studies are not homogeneous, presumably due to differences in study designs and different testing protocols, which impede the overall interpretation of their findings.

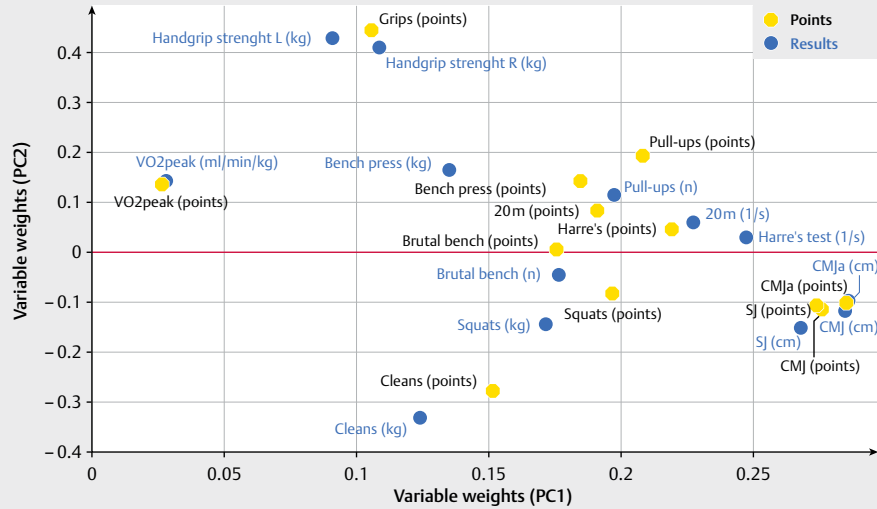
Several studies investigating the usefulness of one or more of the physiological tests included in this study also reported inconsistent results. For example, White and Johnson [3], Emeterio and Gonzalez-Badillo [28] and von Duvillard and Knowles [48] found a low to moderate correlation between performance in vertical jump tests and skiing performance, whereas Impellizzeri, Rampinini, Freschi, et al. [47] found no such relationship. In addition, some studies [14–16] have reported significant correlations between skiing performance and VO<sub>2</sub>max, whereas others [43, 49, 50] have not found similar results. The contradictory results could be attributed to the use of different statistical approaches, selection of test variables, heterogeneity of tested subjects, and the absence of statistical validation. One issue with the different statistical approaches utilized may lie in the use of parametric versus non-parametric data where, for example, FIS points demonstrate a non-linear (sigmoid-



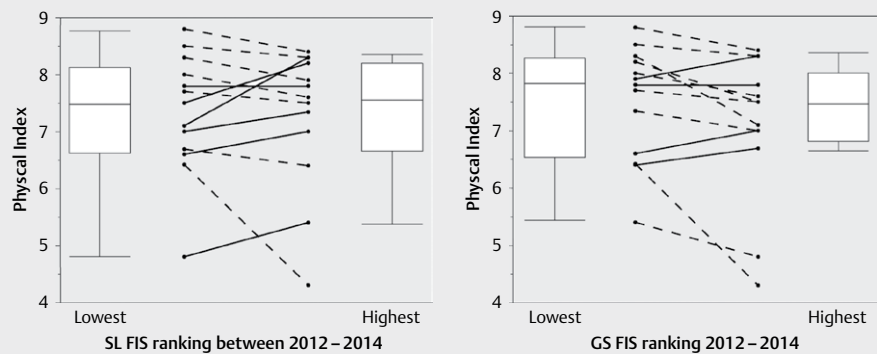
► **Fig. 2 a-b.** Validation of predicted FIS SL **a** and GS **b** ranking. Both models show cross-validation by repeated permutations. The plots indicate the risk that the OPLS model is spurious, i. e., the model just fit the training set well but does not predict Y well for new observations. Goodness-of-fit ( $R^2$  and  $Q^2$ ) of the original model is compared with the goodness-of-fit of models based on data where the order of the Y observations has been randomly permuted, whereas the X matrix has been kept intact. For the selected Yvariable (SL and GS), on the vertical axis for both models, the values of  $R^2$  and  $Q^2$  for the original model (far to the right) and of the Y-permuted models further to the left. The horizontal axis shows the correlation between the permuted Y-vectors and the original Y-vector for the selected Y. The original Y correlates 1.0 with itself, defining the high point on the horizontal axis. The plots above strongly indicate that the original models are NOT valid. The criteria for validity are that all blue  $Q^2$  values to the left are lower than the original points to the right, or the blue regression line of the  $Q^2$  points intersects the vertical axis (on the left) at or below zero.

dal) distribution [43], and as such linear statistical methods do not apply. Another concern is the lack of cross-validation whereby correlations may be due to chance [51], and when applying models to novel datasets correlations cannot be repeated. Contradicting results can also be found despite apparently standardized protocols, exemplified by findings from two different studies in which results from a 30-s maximal sprint Wingate cycling test were correlated [3] and not correlated [10] to alpine skiing performance. Lack of correlation on a group level, despite standardized testing protocols such as the Wingate test, indicate that the importance of such tests have large individual variations, suggesting that individual profiling is needed.

The lack of performance-predictive power of commonly used tests is clearly a consequence of the fact that alpine skiing is a complex sport [1, 15], in which performance is the result of several different interacting qualities [10, 15, 18, 19, 52], many of them not tested in this study. For example, previous research suggests that variables such as balance [53, 54], coordination [18], flexibility [18], and eccentric and isometric muscle actions [6, 13, 22, 55, 56] are essential for the sport-specific performance in alpine skiing. The performance variables examined in the present study were conducted according to Fysprofilen's instructions and were therefore not initially intended to test the unique requirements of alpine skiing; rather it was the athletes' physical capacity in a variety of variables that were tested. Some previous studies (one of which is a



► **Fig. 3** Principal component analysis of physiological test results and corresponding score points. Loading scatter plot visualizes the correlation between variables. Physiological test results and score points located in the same area of the loading plot are correlated. Physiological test results, green squares; Score points; blue squares. Included variables (X) = 13.



► **Fig. 4** Boxplot and paired dot-line plot for Fysprofilen Index, grouped by Lowest and Highest SL and GS FIS ranking between the years 2012 and 2014. Solid lines indicate a positive correlation between Fysprofilen Index and SL/GS ranking, while dashed lines indicate a negative correlation. Boxplots show no differences in Fysprofilen Index between Lowest and Highest FIS ranking in either discipline. The ends of the box represent the 25<sup>th</sup> and 75<sup>th</sup> quantiles (interquartile range). The whiskers extend from the ends of the box to the outermost data point that falls within the distances computed as follows: 1st quartile - 1.5 \* (interquartile range), 3rd quartile + 1.5 \* (interquartile range).

published master’s thesis) have used a similar approach, where the correlation of commonly used training exercises (such as squats) to competitive performance has been investigated [28, 33]. However, the use of parametric statistics or the lack of validation limit the usefulness and applicability of these results.

Consequently, the lack of predictive power (i. e., cross-validated Q<sup>2</sup>) of the tests included in this study is perhaps not surprising, despite selecting tests well described for prediction of physical performance in general. In alpine skiing, the use of nonvalidated test batteries seems common, and the scientific rationale behind this study was specifically to investigate one of these test batteries and thereby be able to illustrate this potential problem further. Indeed, attempts to enhance the correlation and predictive power of conducted physical tests by introducing a scoring or ranking system,

such as the investigated Fysprofilen Index, will ultimately not change the inherent lack of correlation between test results and athletic performance. Instead, our results indicate that the sport-specific performance and Fysprofilen Index for alpine skiers are independent (► **Fig. 4**) and that physiological test results, and the given point by each specific test, can be used interchangeably (► **Fig. 3**).

One limitation of the present as well as previous studies is the use of FIS points as the outcome of race performance. Because the FIS points formula is a mathematical function that displays a non-linear distribution [43], parametric statistical methods cannot be applied. The system has also been criticized for not being fair [57], primarily due to how the points are calculated and that it encourages opportunistic behavior in less skilled skiers in that they con-

sciously attend competitions where only a few elite athletes participate. An alternative outcome could be time trial performance on a standardized intervention course [21], with a limitation in sample size that is not representative of a genuinely competitive situation. Individual competition results can be another outcome variable, but due to technical limitations and strategical reasons [10, 57], this will not represent performance over time. Furthermore, the use of two different  $\dot{V}O_{2\text{peak}}$  protocols can be viewed as a limitation. However, the potential variation in results caused by the use of different protocols (ergometer cycling vs. treadmill running) is alleviated by the low statistical impact of  $\dot{V}O_{2\text{peak}}$  observed in the models.

This study has demonstrated that physiological test variables cannot predict competitive performance of senior elite female alpine skiers on a group level, neither alone nor in combination. Even when the test scores are converted into a summarizing Fysprofilen Index, they fail to discriminate athletes on an individual level. Future studies should include validation of selected test batteries before implementation, and careful attention must be paid when predicting future competitive results. As this study has also suggested, correlations in one dataset do not imply a prediction in another. Individual profiling is needed and more specific testing procedures need to be developed.

## Acknowledgements

The authors wish to thank Roger Andersson, Lennart Burlin, and Mikael Therell at the Sports Medicine Laboratory at Umeå University for invaluable help during the data collection. We are also grateful to SKI TEAM SWEDEN ALPINE for making data available for analysis.

## Conflict of Interest

The authors declare that they have no conflict of interest.

## References

- [1] Turnbull JR, Kilding AE, Keogh JW. Physiology of alpine skiing. *Scand J Med Sci Sports* 2009; 19: 146–155. doi:10.1111/j.1600-0838.2009.00901.x
- [2] Supej M, Hebert-Losier K, Holmberg HC. Impact of the steepness of the slope on the biomechanics of World Cup slalom skiers. *Int J Sports Physiol Perform* 2015; 10: 361–368. doi:10.1123/ijsp.2014-0200
- [3] White AT, Johnson SC. Physiological aspects and injury in elite Alpine skiers. *Sports Med* 1993; 15: 170–178
- [4] Szmedra L, Im J, Nioka S et al. Hemoglobin/myoglobin oxygen desaturation during Alpine skiing. *Med Sci Sports Exerc* 2001; 33: 232–236
- [5] Karlsson J, Eriksson A, Forsberg A et al. *The Physiology of Alpine Skiing*. 1st edition. Park City, UT: The United States Ski Coaches Association; 1978
- [6] Tesch PA. Aspects on muscle properties and use in competitive Alpine skiing. *Med Sci Sports Exerc* 1995; 27: 310–314
- [7] Hintermeister RA, O'Connor DD, Lange GW et al. Muscle activity in wedge, parallel, and giant slalom skiing. *Med Sci Sports Exerc* 1997; 29: 548–553
- [8] Davidson TM, Laliotis AT. Alpine skiing injuries. A nine-year study. *West J Med* 1996; 164: 310–314
- [9] Ferguson RA. Limitations to performance during alpine skiing. *Exp Physiol* 2010; 95: 404–410. doi:10.1113/expphysiol.2009.047563
- [10] Bacharach DW, von Duvillard SP. Intermediate and long-term anaerobic performance of elite Alpine skiers. *Med Sci Sports Exerc* 1995; 27: 305–309. doi:10.1249/00005768-199503000-00003
- [11] Gilgien M, Sporri J, Chardonens J et al. Determination of external forces in alpine skiing using a differential global navigation satellite system. *Sensors* 2013; 13: 9821–9835. doi:10.3390/s130809821
- [12] Reid RC. *A kinematic and kinetic study of alpine skiing technique in slalom* [Doctoral Dissertation]. Oslo, Norway: Norwegian School of Sport Sciences; 2010
- [13] Hintermeister RA, O'Connor DD, Dillman CJ et al. Muscle activity in slalom and giant slalom skiing. *Med Sci Sports Exerc* 1995; 27: 315–322
- [14] Song TMK. Relationship of physiological characteristics to skiing performance. *Phys Sportsmed* 1982; 10: 96–102. doi:10.1080/00913847.1982.11947396
- [15] Neumayr G, Hoertnagl H, Pfister R et al. Physical and physiological factors associated with success in professional alpine skiing. *Int J Sports Med* 2003; 24: 571–575. doi:10.1055/s-2003-43270
- [16] Veicsteinas A, Ferretti G, Margonato V et al. Energy cost of an energy sources for alpine skiing in top athletes. *J Appl Physiol Respir Environ Exerc Physiol* 1984; 56: 1187–1190. doi:10.1152/jap-1984.56.5.1187
- [17] Spring H, Jordan K. [Maximal and high-velocity power. A study in Swiss male and female national ski athletes]. *Schweiz Z Med Traumatol* 1994; 27–29
- [18] Andersen RE, Montgomery DL. Physiology of Alpine skiing. *Sports Med* 1988; 6: 210–221
- [19] Hebert-Losier K, Supej M, Holmberg HC. Biomechanical factors influencing the performance of elite Alpine ski racers. *Sports Med* 2014; 44: 519–533. doi:10.1007/s40279-013-0132-z
- [20] Kröll J, Birklbauer J, Stricker G et al. Technique training in alpine ski racing: forced movement changes by a specific device. In: 24<sup>th</sup> International Symposium on Biomechanics in Sports; Salzburg, Austria. 2006; 559–563
- [21] Andersen RE, Montgomery DL, Turcotte RA. An on-site test battery to evaluate giant slalom skiing performance. *J Sports Med Phys Fitness* 1990; 30: 276–282
- [22] Haymes EM, Dickinson AL. Characteristics of elite male and female ski racers. *Med Sci Sports Exerc* 1980; 12: 153–158
- [23] Jasmin BJ, Montgomery DL, Hoshizaki TB. Applicability of the hexagonal obstacle test as a measure of anaerobic endurance for alpine skiers. *Sports Med Train Rehabil* 1989; 1: 155–163. doi:10.1080/15438628909511870
- [24] Stark RM, Reed AT, Wenger HJ. Power curve characteristics of elite slalom and downhill skiers performing a modified 90 s Wingate test. *Can J Sport Sci* 1987; 12: 24–25
- [25] Piper F, CHT Ward, McGinnis P M et al. Prediction of alpine ski performance based upon selected anthropometrical and motor dexterity parameters. *J Sports Med Phys Fitness* 1988; 27: 478–482
- [26] Patterson C, Raschner C, Platzer HP. The 2.5-minute loaded repeated jump test: evaluating anaerobic capacity in alpine ski racers with loaded countermovement jumps. *J Strength Cond Res* 2014; 28: 2611–2620. doi:10.1519/JSC.0000000000000436
- [27] Bosco C, Cotelli F, Bonomi R et al. Seasonal fluctuations of selected physiological characteristics of elite alpine skiers. *Eur J Appl Physiol Occup Physiol* 1994; 69: 71–74



- [28] Emeterio CA, Gonzalez-Badillo JJ. The physical and anthropometric profiles of adolescent alpine skiers and their relationship with sporting rank. *J Strength Cond Res* 2010; 24: 1007–1012. doi:10.1519/JSC.0b013e3181cbabb5
- [29] Patterson C, Raschner C, Platzer HP. Power variables and bilateral force differences during unloaded and loaded squat jumps in high performance alpine ski racers. *J Strength Cond Res* 2009; 23: 779–787. doi:10.1519/JSC.0b013e3181a2d7b3
- [30] Raschner C, Platzer HP, Patterson C et al. The relationship between ACL injuries and physical fitness in young competitive ski racers: a 10-year longitudinal study. *Br J Sports Med* 2012; 46: 1065–1071. doi:10.1136/bjsports-2012-091050
- [31] Sveriges Olympiska Kommitté. Frågor & Svar om Fysprofilen. In: 2016
- [32] Muller E, Benko U, Raschner C et al. Specific fitness training and testing in competitive sports. *Med Sci Sports Exerc* 2000; 32: 216–220
- [33] Heikkinen D. Physical testing characteristics and technical event performance of junior alpine ski racers. *Electronic Theses and Dissertations*. 2003
- [34] Vermeulen B, Clijsen R, Fässler R et al. Event-specific body characteristics of elite alpine skiers in relation to international rankings. *Advances in Anthropology* 2017; 07: 94–106. doi:10.4236/aa.2017.72007
- [35] Harriss DJ, Macsween A, Atkinson G. Standards for ethics in sport and exercise science research: 2018 update. *Int J Sports Med* 2017; 38: 1126–1131. doi:10.1055/s-0043-124001
- [36] Sveriges Olympiska Kommitté. Testbeskrivningar Fysprofilen Utmanare. In: 2016
- [37] Hoyek N, Champely S, Collet C et al. Age and gender-related differences in the temporal congruence development between motor imagery and motor performance. *Learn Individ Diff* 2009; 19: 555–560. doi:10.1016/j.lindif.2009.07.003
- [38] Svensson K, Alricsson M, Olausson M et al. Physical performance tests – a relationship of risk factors for muscle injuries in elite level male football players. *J Exerc Rehabil* 2018; 14: 282–288. doi:10.12965/jer.1836028.014
- [39] Carlsson M, Carlsson T, Hammarstrom D et al. Validation of physiological tests in relation to competitive performances in elite male distance cross-country skiing. *J Strength Cond Res* 2012; 26: 1496–1504. doi:10.1519/JSC.0b013e318231a799
- [40] Tilley NR, Macfarlane A. Effects of different warm-up programs on golf performance in elite male golfers. *Int J Sports Phys Ther* 2012; 7: 388–395
- [41] Stoggl T, Muller E, Ainegren M et al. General strengt h and kinetics: fundamental to sprinting faster in cross country skiing? *Scand J Med Sci Sports* 2011; 21: 791–803. doi:10.1111/j.1600-0838.2009.01078.x
- [42] Ski Fld Inside FIS - Document Library. In: 2017
- [43] Nilsson R, Lindberg A-S, Theos A et al. Aerobic variables for prediction of alpine skiing performance – a novel approach. *Sports Med Int Open* 2018; 02: E105–E112. doi:10.1055/a-0655-7249
- [44] Severini TA. *Analytic Methods In Sports: Using Mathematics and Statistics to Understand Data From Baseball, Football, Basketball, and Other Sports*. Boca Raton: CRC Press; 2015
- [45] Lindberg AS, Oksa J, Antti H et al. Multivariate statistical assessment of predictors of firefighters' muscular and aerobic work capacity. *PloS one* 2015; 10: e0118945. doi:10.1371/journal.pone.0118945
- [46] Eriksson L, Byrne T, Johansson E et al. *Multi- and Megavariate Data Analysis: Basic Principles and Applications*. Malmö: MKS Umetrics. 2013;
- [47] Impellizzeri FM, Rampinini E, Freschi M et al. Identification of the physical characteristics that discriminate between competitive levels and specialities of alpine skiing. In: Müller E, Lindinger S, Stöggl T Eds. *Science and Skiing IV*. Meyer & Meyer Sport (UK) Ltd; 2009: 272–280
- [48] von Duvillard SP, Knowles WJ. Relationship of anaerobic performance tests to competitive alpine skiing events. In Müller E, Schwameder H, Kornelx E, Eds. *Science and Skiing*. Meyer & Meyer Sport (UK) Ltd.; 1997: 299–310
- [49] Brown SL, Wilkinson JG. Characteristics of national, divisional, and club male alpine ski racers. *Med Sci Sports Exerc* 1983; 15: 491–495
- [50] White AT, Johnson SC. Physiological comparison of international, national and regional alpine skiers. *Int J Sports Med* 1991; 12: 374–378. doi:10.1055/s-2007-1024697
- [51] Loken E, Gelman A. Measurement error and the replication crisis. *Science* 2017; 355: 584–585. doi:10.1126/science.aal3618
- [52] Saibene F, Cortili G, Gavazzi P et al. Energy sources in alpine skiing (giant slalom). *Eur J Appl Physiol Occup Physiol* 1985; 53: 312–316
- [53] Wojtyczek B, Paslawska M, Raschner C. Changes in the balance performance of Polish recreational skiers after seven days of alpine skiing. *J Hum Kinet* 2014; 44: 29–40. doi:10.2478/hukin-2014-0108
- [54] Cigrovski V, Franjko I, Rupčić T et al. Comparison of standard and newer balance tests in recreational alpine skiers and ski novices. *Monten J Sports Sci Med* 2017; 6: 49–55
- [55] Ropret R. The role of eccentric regime of leg muscle work in alpine skiing. *Fizicka kultura* 2017; 71: 5–11. doi:10.5937/fizkul1701005R
- [56] Berg HE, Eiken O. Muscle control in elite alpine skiing. *Med Sci Sports Exerc* 1999; 31: 1065–1067
- [57] Maisano D, Botta A, Franceschini F. On the rating system in alpine skiing racing: Criticism and new proposals. *Proceedings of the Institution of Mechanical Engineers, Part P. J Sport Eng Tech* 2016; 230: 253–263. doi:10.1177/1754337115621818