This is the published version of a paper published in .

Citation for the original published paper (version of record):

Aerobic Variables for Prediction of Alpine Skiing Performance – A Novel Approach
Sports Medicine International Open, 2(4): E105-E112
https://doi.org/10.1055/a-0655-7249

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-154284
Aerobic Variables for Prediction of Alpine Skiing Performance – A Novel Approach

Introduction

Competitive alpine skiing is regarded as the world’s fastest non-motorized sport [5] and consists primarily of four different disciplines: slalom (SL), giant slalom (GS), super giant slalom (SG) and downhill (DH), the latter two being referred to as speed events. Each of these disciplines differs in turning radius, gate distance, speed, and length of the course [41]. The two more technical disciplines (SL and GS) run on relatively steep terrain at speeds between 20–60 km·h⁻¹ for 45–90 s, whereas the speed disciplines are carried out on long, steep slopes at speeds reaching 130–160 km·h⁻¹ and lasting 120–180 s [10, 32]. Due to the intermittent nature of the sport with rapid shifts in directions at high-speed [27], alpine ski racing is a demanding and multi-faceted sport requiring high levels of physical and technical competence [37].
Elite performance in alpine skiing requires a wide range of physical qualities, including high muscular strength and endurance, as well as a broad range of neuromuscular skills such as balance, speed, and agility [2, 20, 41]. Anthropometric characteristics may also affect performance [1, 2, 37], especially in young athletes [8, 14]. High aerobic capacity has been described as an important physical quality of an alpine skier [23, 31] and today it is considered, at least by coaches, athletes and federations, an important physiological determinant of competitive success. However, few findings support this assertion. Inconsistent findings regarding required metabolic capacities for optimal competitive performance are reported; some researchers find a significant correlation between maximum absolute and relative aerobic capacity (VO₂max) and competitive performance in elite alpine skiers [13, 21], whereas others do not [41]. Furthermore, VO₂max could not discriminate between skiers at different performance levels [6, 40].

Instead, several contemporary studies show a correlation between anaerobic power and ski performance [3, 6, 13, 39, 41]. With race times of 45–180 s, both aerobic and anaerobic energy systems are utilized [29], which possibly contributes to the disagreements between studies [24]. Hence, the importance of the respective metabolic system is still a matter of debate [19, 22, 24, 30]. Technical events (SL, GS) have a larger relative anaerobic component compared to the faster and longer-duration disciplines (SG, DH) [38, 39]. Differences in skiing technique, mechanical work, and overall skills will also result in differences in the relative utilization of energy systems [27, 38].

In alpine skiing, long-term, or summarized, competitive performance is quantified by the Fédération Internationale de Ski (FIS) point scoring system, based on practitioners’ results in each discipline during the racing season. In brief, FIS points are matched so that the best skier in the world in every discipline has 0 points and the 30th has 6 points, per season. The point system is therefore a measure of where each practitioner stands against other practitioners. The points system is adjusted several times annually and can be found, along with detailed information about the ranking system, at www.fis-ski.com.

Because of the complexity of the sport, with a mixed utilization of energy systems and lack of consistent research finding, the selection of useful sport-specific tests of physiological capacities is challenging. Consequently, coaches and athletes are dependent on non-reliable tests for planning and evaluating training. This study will use advanced multivariate statistics, the validity of which has been demonstrated in other similar contexts [17], to investigate relationships between results from commonly used aerobic tests and anthropometric variables and their value to predictive alpine skiing performance. By focusing on aerobic tests, results from this study will guide future research towards more specific testing procedures. Omitting all or any irrelevant tests will save time and money for sports federations and sports clubs alike, as well as help coaches and athletes to optimize training, evaluation and racing performance.

Materials & Methods

Subjects

Anthropometric data and aerobic test results from a total of twenty-three elite junior alpine skiers, aged 16–17 years, were included in the study (▶Table 1). All participants provided their written, informed consent for participation and parental/guardian consent was obtained for participants under 18 years of age.

Table 1: Anthropometric data and aerobic test results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Male</th>
<th></th>
<th>Female</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 17 yr (n = 10)</td>
<td>Age 16 yr (n = 13)</td>
<td>Age 17 yr (n = 6)</td>
<td>Age 16 yr (n = 10)</td>
<td></td>
</tr>
<tr>
<td><strong>Body weight (kg)</strong></td>
<td>75.4 ± 5.3</td>
<td>69.2 ± 5.5</td>
<td>68.1 ± 3.7</td>
<td>69.7 ± 3.4</td>
</tr>
<tr>
<td><strong>Body height (m)</strong></td>
<td>1.78 ± 0.5</td>
<td>1.78 ± 0.5</td>
<td>1.70 ± 0.6</td>
<td>1.72 ± 0.4</td>
</tr>
<tr>
<td><strong>BMI (kg·m⁻²)</strong></td>
<td>24 ± 1</td>
<td>22 ± 1</td>
<td>24 ± 1</td>
<td>24 ± 2</td>
</tr>
<tr>
<td><strong>HR at aerobic ventilatory threshold (bmp)</strong></td>
<td>128.8 ± 18.0</td>
<td>126.5 ± 21.4</td>
<td>129.6 ± 17.9</td>
<td>125.7 ± 12.9</td>
</tr>
<tr>
<td><strong>VO₂ at aerobic ventilatory threshold (L·min⁻¹)</strong></td>
<td>1.6 ± 0.20</td>
<td>1.5 ± 0.5</td>
<td>1.3 ± 0.4</td>
<td>1.1 ± 0.2</td>
</tr>
<tr>
<td><strong>HR at aerobic ventilatory threshold by Wasserman (bmp)</strong></td>
<td>163.2 ± 10.7</td>
<td>157.3 ± 19.5</td>
<td>162.3 ± 12.7</td>
<td>155 ± 14</td>
</tr>
<tr>
<td><strong>VO₂ at anaerobic threshold by Wasserman (L·min⁻¹)</strong></td>
<td>2.6 ± 0.5</td>
<td>2.4 ± 0.4</td>
<td>2 ± 0.3</td>
<td>1.8 ± 0.3</td>
</tr>
<tr>
<td><strong>HR at 2.00 mmol·L⁻¹ [HLa]ₘₜ</strong></td>
<td>159.9 ± 12.4</td>
<td>160.1 ± 11.3</td>
<td>160.9 ± 15.3</td>
<td>169.2 ± 8.3</td>
</tr>
<tr>
<td><strong>VO₂ at 2.00 mmol·L⁻¹ [HLa]ₘₜ</strong></td>
<td>2.4 ± 0.5</td>
<td>2.5 ± 0.3</td>
<td>2.0 ± 0.3</td>
<td>2.2 ± 0.3</td>
</tr>
<tr>
<td><strong>HR at RER 1.00</strong></td>
<td>181.6 ± 9.9</td>
<td>175.2 ± 14.8</td>
<td>175.8 ± 13.5</td>
<td>177 ± 10.4</td>
</tr>
<tr>
<td><strong>VO₂ at RER 1.00</strong></td>
<td>3.3 ± 0.4</td>
<td>3.0 ± 0.6</td>
<td>2.5 ± 0.3</td>
<td>2.5 ± 0.3</td>
</tr>
<tr>
<td><strong>HR at OBLA</strong></td>
<td>182.2 ± 7.4</td>
<td>179.9 ± 11.7</td>
<td>180.3 ± 12.3</td>
<td>185.5 ± 8.6</td>
</tr>
<tr>
<td><strong>VO₂ at OBLA</strong></td>
<td>3.3 ± 0.3</td>
<td>3.2 ± 0.4</td>
<td>2.6 ± 0.2</td>
<td>2.9 ± 0.2</td>
</tr>
<tr>
<td><strong>Maximal HR</strong></td>
<td>199.8 ± 5.6</td>
<td>199.7 ± 8.5</td>
<td>195.8 ± 8.2</td>
<td>193.3 ± 8.1</td>
</tr>
<tr>
<td><strong>Maximal absolute VO₂ (L·min⁻¹)</strong></td>
<td>4.4 ± 0.4</td>
<td>4.1 ± 0.4</td>
<td>3.3 ± 0.3</td>
<td>3.2 ± 0.2</td>
</tr>
<tr>
<td><strong>Maximal relative VO₂ (ml·min·kg⁻¹)</strong></td>
<td>58.2 ± 4.4</td>
<td>58.6 ± 2.6</td>
<td>48.4 ± 2.3</td>
<td>46.5 ± 2.4</td>
</tr>
</tbody>
</table>

HR = heart rate; bpm = beats per minute; VO₂ = volume of oxygen uptake; [HLa]ₘₜ = concentration of blood lactate; RER = respiratory exchange ratio; OBLA = onset of blood lactate accumulation 4.00 mmol·L⁻¹. All respiratory variables and HR are the average during 30 s. Mean ± SD.
obtained for minors. Some participants were tested more than once, as indicated in the data. Ethical permission Dnr 2011-236-31 M was granted by the ethical committee for northern Sweden at Umeå University, 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 [12].

Methodology

Analyzed data were compiled from both aerobic and anthropometric tests conducted at the Department of Sports Medicine, Umeå University, Sweden. All participants included in the study were asked to refrain from any strenuous physical activity the day before the tests and to follow the same routine with respect to e.g., sleep and nutritional intake, before all test occasions. Before each test occasion, the participants were asked to complete a health questionnaire regarding previous exercise training, nutritional intake and disease history. Exclusion criteria included any injury and/or history of hormonal, metabolic or cardiovascular diseases.

Slalom and GS FIS ranking points were collected twice, in December (6th list) and April (11th list) and correlated to the pre-season testing in June-July and October-December. Data were sorted into categories by SL, GS, sex, year of birth and national ranking (Sweden). In brief, FIS ranking was used as a measurement of competitive performance (the Y-variable in all statistical analyses), and maximal physiological capacity including VO2peak, heart rate, and blood lactate were measured during a maximal cycling ergometer test (X-variables). Anthropometric variables recorded were body mass and stature. BMI calculation was based on the following formula:

\[
\text{BMI} = \frac{\text{body mass (kg)}}{\text{stature (m)}}^2
\]

Before each test session, a short medical exam was performed, including resting blood pressure. Body mass was measured using a standard weight scale (Soehnle weighing scale, Leifheit AG, Nasaus, Germany), and body stature was measured using a wall-mounted scale (Fosamax stadiometer, Merck & Co. Inc., Kenilworth, NJ, USA). A peripheral venous catheter (Optiva 2 radiopaque I.V. catheter, L = 32 mm, ø = 1.10 mm, Smiths Medical, London, England) was placed in the antecubital vein, and an airtight facial mask (Hans Rudolph Inc., Shawnee, KS, USA) was fitted to cover the subject’s airways. At rest, blood samples were collected and subsequently analyzed for hemoglobin concentration (HB) with a HemoCue Hb 201+ (HemoCue AB, Angelholm, Sweden) and blood lactate concentration [HLa] with a YSI 1500 sport analyzer (YSI Life Sciences, Yellow Springs, OH, USA). Maximum performance tests were carried out on a Monark 839E cycle ergometer (Monark, Vargberg, Sweden) with an increasing load (+40 and 30 W every 3 min, with a starting load of 40 W for males and 30 W for females) until exhaustion. Oxygen consumption (VO2) was tested using a calibrated Oxymon Pro Jaeger System (VIASYS Healthcare, Würzburg, Germany) set on mixing chamber mode, and the mean value for the final 30 s before exhaustion registered as VO2peak. Heart rate was monitored telemetrically using a Polar Electro s610i (Polar Electro Oy, Kempele, Finland) pulse watch. Blood lactate samples were collected every 3 min (equivalent to the 3rd minute of each exercise load). Test results are X-variables in the statistical analysis.

Multivariate statistics – justification and simplified description of PCA and OPLS

In sports research, the use of bivariate and multivariate linear regression is common [28]. However, using traditional sports research methods may in some cases be an inefficient approach because multivariate data is often hidden and some variables can interact and elicit a specific response [9]. A complex interaction between different qualities determines the performance of an alpine skier [16, 21, 37], making a complete identification and interpretation of valid performance data by means of conventional statistics impossible. Thus, the data in this study have been analyzed using multivariate data analysis (MVDA). Multivariate statistical methods use two or more variables collectively to investigate interactive outcomes. To examine the relationship between anthropometric and physiological variables and the sport-specific performance of alpine skiing, principal component analysis (PCA) and orthogonal projections to latent structures (OPLS) statistical methods were applied. Principal component analysis is a relatively simple method that can be used to investigate how observations and variables relate to each other and to find hidden structures or patterns in data, all by reducing data dimensions [9, 15].

Orthogonal projections to latent structures is a statistical method that is considered an extended version of PCA and is often used to find the linear relationship between two groups of variables [9]. Similar to PCA, observations in OPLS are assumed to be affected only by a few indivisible underlying variables. To facilitate the identification of these variables, estimations of the observations are therefore calculated. The difference from PCA, however, is that each observation in the data matrices is now expressed as two separate point coordinates, with one projected into the X space and the other into the Y space [45]. Orthogonal projections to latent structures is then used to examine whether there is a correlation between these point coordinates, one in each multivariate space, to predict Y based on X [35]. Thus, the regression in OPLS is calculated by the covariance between Y and X instead of the variance within X, iteratively by testing the predictive ability of previous components until the new model is not considered significant when the procedure stops. Cross-validation by permutation is used to determine the number of regression components that should be included in the model [35]. For more in-depth reading on the statistical methods, we refer to the published literature [7, 9, 15, 36, 44].

Statistical analysis

Prediction of FIS ranking (Y-variables) was achieved from anthropometric and physiological test results (X-variables). R2VY is the cumulative percent of the variation of the response explained by the model after the last component. R2 is a measure of fit, i.e., how well the model fits the data. R2VYAdj is the cumulative percent of the variation of the response, adjusted for degrees of freedom, explained by the model after the last component. Q2VY is the cumulative percent of the variation of the response predicted by the model, after the last component, according to cross-validation. Q2 indicates how well the model predicts new data. A useful model should have a large R2 and Q2. To evaluate the importance of variables for FIS ranking, an analysis of variable influence on projection (VIP) was executed. In an OPLS model, VIP summarizes the impor-
tance of the X-variables, both for the X- and Y-models. VIP is
normalized, and the average squared VIP value is 1; thus a VIP > 1 in-
dicates that the variable is important for the projection, and values
lower than 0.5 indicate that the variable is unimportant for the pro-
jection. An R^2 and a Q^2 > 0.60 were deemed valid. The significance
is set by rules 1, 2 and 3, where Q^2 > Limit (indicated as R1, R2 and
R3 in Results, where the Limit depends on a number of components
for PCA and Y-variables for OPLS (extract from the SIMCA-P + Hand-
book)).

Physiological and anthropometric variables were distributed
normally according to Shapiro-Wilk goodness-of-fit test (p < 0.05).
Because FIS points and rankings are ordinal, parametric statistics
cannot be applied. If treated as continuous data, as in other publi-
cations, FIS points do not have a normal distribution (Fig. 1: Sha-
piro-Wilk goodness-of-fit test p > 0.01), which also suggests the
use of non-parametric methods. Data were analyzed using SIMCA
14.0 (MKS AB, Umeå, Sweden) and JMP 13.1 (SAS Institute Inc.,
Cary, NC, USA).

Results

Anthropometric and physiological test results are presented in
Table 1. There was no significant correlation between VO\textsubscript{2peak} and
FIS ranking (Fig. 2). Table 2 shows OPLS models with cross-valida-
tion for each year by sex and discipline. Fig. 3 is a PCA scatter plot
(A) and loading plot (B) showing a clustering of sex based on physical
performance (Fig. 3a), where the loading plot indicates higher
VO\textsubscript{2peak} and lower heart rate in males compared to females (Fig. 3b).
The correlation of FIS ranking with physical tests and anthropometric
data by OPLS in Fig. 4 indicates a clustering of FIS rank not related to
any analyzed variables (not located in the same area of the plot). Of all
variables analyzed, BMI is located closest to the FIS rank cluster, but
with a weight of 0.2 it is not of significant importance in the models.
BMI scores a 2.0 but with large 95% jackknife uncertainty bars (Fig.
5b). Separation into FIS rank of young elite alpine skiers could not be
reliably predicted using results from physiological and anthropomet-
ric measurements (Table 2). The explanation of variation in the re-
gression models yielded R^2 > 0.5 (more than 50% of the error explained
by the components). However, prediction of rank (by cross-validation)
could not be made with high confidence when Q^2 < 0.3 in all OPLS anal-
yses (Table 2). As an example, the model for slalom rank in females
aged 16 is shown in Fig. 5, where a significant regression model was
observed (Fig. 5a), with BMI as the most important variable (Fig.
5b). However, the R^2/Q^2 plot (Fig. 5c) demonstrates a low predic-
tive power (Q^2 = −0.29), and cross-validation by permutation (Fig.
6) confirms the low predictive power and rejects the regression model
in Fig. 5a.
The main finding of this study is that aerobic and anthropometric variables cannot predict alpine skiing performance, even when multivariate statistics are applied. In conformity with some previous studies [6, 40] but in contrast to others [11, 13, 21], we could not demonstrate a strong correlation between aerobic work capacity (V\text{̇}O_2\text{peak}) and competitive performance, as indicated by FIS ranking (▶ Fig. 2). Because none of the aerobic or anthropometric variables cluster with any of the FIS rankings (▶ Fig. 4), the overall interpretation of the findings in this study must be that factors other than those investigated predict long-term performance in alpine skiing among adolescents. This outcome is not surprising, given the large variation in reported aerobic demands during high-intensity alpine skiing, ranging from around 80 % up to 200 % of V\text{̇}O_2\text{max} [26, 33, 38]. Furthermore, the lack of correlation in our calculations can to some extent be a result of the overall skill set of the individual athlete, because differences in skiing economy between practitioners most likely result in a significant variation in the importance of various physiological skills on an individual level. This claim is supported by the fact that the practitioners in the present study, who compete at the national junior level, have similar V\text{̇}O_2 values (▶ Table 1), with a mean of 58 ml · min · kg\textsuperscript{−1} for men and 49 ml · min · kg\textsuperscript{−1} for women. These numbers are similar to world-class male [21] and female (Swedish national women’s team, unpublished data) skiers. Thus, in agreement with previous studies [6, 40], it seems that V\text{̇}O_2\text{max} is not a discriminating factor between practitioners at different levels and that a relative V\text{̇}O_2\text{max} of ≥ 50 ml · kg\textsuperscript{−1} · min\textsuperscript{−1} appears sufficient to be competitive at the highest international level.

As a measurement of anaerobic workload both Saibene et al. [26] and Tesch et al. [33] conclude that high [HLa]b is reached during both SL and GS, without affecting competitive performance ei-

**Table 2** Multivariate models.

<table>
<thead>
<tr>
<th>Age</th>
<th>FIS rank</th>
<th>Sex</th>
<th>R\textsuperscript{2}/Q\textsuperscript{2} Adj</th>
<th>R\textsupersignificance</th>
<th>Q\textsuperscript{2} cross-validated</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 (n = 23)</td>
<td>Slalom</td>
<td>Female (n = 10)</td>
<td>0.73/−0.29</td>
<td>NS</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male (n = 13)</td>
<td>0.61/−0.58</td>
<td>NS</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Giant Slalom</td>
<td>Female (n = 10)</td>
<td>0.65/−0.73</td>
<td>R1</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male (n = 13)</td>
<td>0.51/−0.51</td>
<td>NS</td>
<td>Yes</td>
</tr>
<tr>
<td>17 (n = 16)</td>
<td>Slalom</td>
<td>Female (n = 6)</td>
<td>0.86/0.18</td>
<td>R1</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male (n = 10)</td>
<td>0.84/−0.08</td>
<td>R1</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Giant Slalom</td>
<td>Female (n = 6)</td>
<td>0.80/−0.35</td>
<td>NS</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male (n = 10)</td>
<td>0.82/−0.19</td>
<td>NS</td>
<td>No</td>
</tr>
</tbody>
</table>

Prediction of final (11\textsuperscript{th}) FIS ranking in April each year. Q\textsuperscript{2} indicates the overall fit and the predictive power of the model. R\textsuperscript{2} and Q\textsuperscript{2} should be >0.5 for well-modeled data. Q\textsuperscript{2} cross-validated; assessment of the risk that the model is spurious, i.e., the model fits only the training set well but does not predict Y well for new observations. Permutation (n = 15) validated models if intercept <0 or all permuted Q\textsuperscript{2} values are below original model value. Thus, a "yes" indicates that the predictive model is valid, but not that it is significant. NS; Not Significant. R\textsupersignificance; significant model when Q\textsuperscript{2}> limit, with the limit increased with subsequent components to account for the loss in degrees of freedom in each OPLS model. A valid model should have R\textsupersquare >0.8, Q\textsupersquare >0.5, significance by R1 and "yes" for Q\textsupersquare cross-validation. No model fulfills these criteria. For further details, we refer to SIMCA Support at https://umetrics.com/sites/default/files/kb/multivariate_faq.pdf.

**Fig. 3** Principal component analysis of performance tests. a) Score scatter plot visualizes variation in the performance testing data, here seen as between-group differences and within-group differences. Age 16 and 17 data for male and female. Included variables (X) = 16. Each data point is the total score of one subject. b) Loading scatter plot visualizes correlations between variables: Physical tests located in the same part of the loading plot are correlated. The score plot and the loading plot communicate: subjects located in the same area in the score plot (A), as variables in the loading plot (B), have a high performance within these tests. n = 39. Consensus: Male and female can be separated based on tests results, but no test result is of significant importance.
ther in elite or sub-elite athletes. Our findings support this conclusion, because no variable measured, including \([HLa]_b\), is of significant importance in any model (▶ Table 2 and ▶ Fig. 3–6). However, during extensive preparatory periods, \([HLa]_b\) should be considered as an indicator of anaerobic muscle workload, because high \([HLa]_b\) levels have been associated with decreased skiing performance during repeated training runs [42].

As in the present study, Neumayr et al. [21] did not find a correlation between anthropometric variables and performance in elite adult male and female skiers. In contrast, Emeterio et al. [8] observed significant correlations between anthropometric variables and national rank (Spain) in male adolescent skiers. In both studies, conclusions were made that female and male skiers are limited by different performance factors (without any discriminant analysis shown), and that very few quantifiable variables predict future performance among female alpine skiers. This finding is supported by the separation of sexes in ▶ Fig. 3a.

Suggestions have been made that body control and body composition are important for performance in alpine skiing [2, 37]. Thus, it is not surprising that percent body fat is correlated to performance variables in alpine skiing, because smaller and leaner athletes perform better in SL [13], and skiers with a greater fat mass have an advantage in DH [13, 21]. Calculated BMI was not a significant factor in our models (▶ Fig. 3–5), yet among the ranked variables (▶ Fig. 5b) it ranks as number one but with high variability. Our results, therefore, suggest that body composition may be of importance, but that BMI is an uncertain tool for evaluation.

▶ Fig. 5 Predicted FIS slalom ranking for females age 16. a) Correlation between actual and predicted FIS ranking in April (11th list) based on pre-season physical testing (model from ▶ Table 2). b) VIP summarizes the importance of the X-variables (physical tests), both for the X- and Y-models (X = 16, Y = 4; n = 39). Tests with VIP > 1 are the most relevant for explaining Y. The VIP values reflect the importance of terms in the model both with respect to Y, i.e., its correlation to all the responses and with respect to X (the projection). The plot is displayed with 95% jackknife uncertainty bars. c) Overview plot shows the cumulated R² and Q² values for the model. Consensus: The observed correlation between actual and predicted FIS ranking (A) occurred by chance due to a large variation in the data (B) and low predictive power (C).
the statistical analysis, because somatotypes [8] and various anthropometric measurements, including body mass and stature were recorded, we chose to include calculated BMI in the model. However, because BMI is a composite measure, it does not accurately predict body fat percentage (such as athletes) [25, 43]. However, BMI is a valid and reliable predictor of performance, especially because BMI tends to overestimate adipose tissue in individuals with large muscle mass and is lacking, both based on present and previously published data. Thus, future research directed towards screening for valid and reliable physiological tests for evaluation of current, and prediction of future, athletic performance.

In summary, none of the included variables predicted competitive alpine skiing performance. Cross-validation by permutation confirmed the lack of validity in observed multivariate statistical models. We suggest that the relevance of current modes of aerobic and anaerobic testing be considered. A valid and reliable test battery that can predict performance in alpine skiing seems to be lacking, both based on present and previously published data. Thus, future research directed towards screening for valid components of athletic performance is required. The results of this study should encourage future investigations to consider the predictive power of included test variables for the long-term, sport-specific performance in alpine skiing.

Acknowledgements

The authors wish to thank Daniel Lindberg, Sofie Henriksson, Malin Andersson, Sofia Pettersson and Catalina Fahlman for invaluable advice. The authors wish to thank Daniel Lindberg, Sofie Henriksson, Malin Andersson, Sofia Pettersson and Catalina Fahlman for invaluable advice. The authors wish to thank Daniel Lindberg, Sofie Henriksson, Malin Andersson, Sofia Pettersson and Catalina Fahlman for invaluable advice. The authors wish to thank Daniel Lindberg, Sofie Henriksson, Malin Andersson, Sofia Pettersson and Catalina Fahlman for invaluable advice. The authors wish to thank Daniel Lindberg, Sofie Henriksson, Malin Andersson, Sofia Pettersson and Catalina Fahlman for invaluable advice. The authors wish to thank Daniel Lindberg, Sofie Henriksson, Malin Andersson, Sofia Pettersson and Catalina Fahlman for invaluable advice. The authors wish to thank Daniel Lindberg, Sofie Henriksson, Malin Andersson, Sofia Pettersson and Catalina Fahlman for invaluable advice.

Conflict of Interest:

The authors declare that they have no conflict of interest.

References


Even though both the aerobic and anaerobic energy systems help to maintain energy levels during alpine skiing [3, 6, 21, 26, 33, 38], it has repeatedly been shown that maximal aerobic tests cannot adequately predict future competitive performance. A possible explanation is that improvements in VO2max, lactate threshold, and maximal anaerobic power, beyond a certain point, ultimately do not lead to a significant increase in race performance. Furthermore, in agreement with several previous findings [2, 21, 24, 30, 34], this study shows that the average alpine skier does not exhibit exceptionally high VO2peak values. An explanation for the inconsistencies between studies may stem from small sample sizes in noisy global data, resulting in a type I error [18]. Because there are fundamental differences in study designs and participants’ characteristics, including VO2max, 21, 24, 26, 30, 34], standardized, common testing criteria are also difficult to propose. Cross-validation by permutation supports these suggestions (Fig. 6), where the observed regression model is not confirmed, and the weak prediction power validated.

One limitation of this study is the use of BMI as an anthropometric predictor of performance, especially because BMI tends to overestimate adipose tissue in individuals with large muscle mass and a low body fat percentage (such as athletes) [25, 43]. However, because no measurement of body composition other than body mass and stature were recorded, we chose to include calculated BMI in the statistical analysis, because somatotypes [8] and various anthropometric variables (including BMI) has previously been shown to affect the sport-specific performance in alpine skiing [4, 40]. The focus of this study was aerobic variables as these are commonly used predictors of future athletic performance by alpine coaches and federations. Still, the lack of comprehensive anaerobic test results can be viewed as a limitation.

To be of real practical use, exercise testing must also be valid and reliable. Therefore, the presented workflow and analytical procedures, including multivariate statistical methods, can be used as a starting point for a global, more holistic view on performance evaluation [17]. Considering the time, effort and resources allocated to exercise testing of athletes around the world, validated procedures should be the minimum requirement of federations, coaches, and athletes. Well-executed meta-analyses for the selection of candidate tests, followed by larger-scale interventions, can find valid and reliable physiological tests for evaluation of current, and prediction of future, athletic performance.

Fig. 6 Validation of predicted FIS slalom ranking for females age 16. Female’s slalom model shown as an example of cross-validation by 15 permutations (one less cycle than X, the number of physical tests). The plot indicates the risk that the current OPLS model is spurious, i.e., the model just fits the training set well but does not predict Y well for new observations. Goodness-of-fit (R2 and Q2) of the original model is compared with the goodness-of-fit of 15 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 Y-variable (female slalom), on the vertical axis, the values of R2 and Q2 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 plot above strongly indicates that the original model is NOT valid. The criteria for validity are: All blue Q2 values to the left are lower than the original points to the right, or the blue regression line of the Q2 points intersects the vertical axis (on the left) at or below zero.

FIS slalom rank model

Q2 R2

0.2 0.4 0.6 0.8 1

Correlation permuted versus original model


