Estimating skills level in recreational skiing: development and validation of a practical multidimensional instrument

Luppino, F.S.; Hollander-Gijsman, M.E. den; Dekker, F.W.; Bartlema, K.A.; Diepen, M. van

Received: 1 February 2022 | Revised: 13 September 2022 | Accepted: 9 October 2022

DOI: 10.1111/sms.14245

ORIGINAL ARTICLE

WILEY

Estimating skills level in recreational skiing: Development and validation of a practical multidimensional instrument

Floriana Samantha Luppino[1] | **Margien Elisabeth den Hollander-Gijsman**[2] | **Friedo Wilhelm Dekker**[3] | **Kornelis Alexander Bartlema**[4] | **Merel van Diepen**[3]

[1]Eurocross Assistance, Leiden, the Netherlands

[2]Department of Information Technology and Digital Innovation, Leiden University Medical Centre, Leiden, the Netherlands

[3]Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, the Netherlands

[4]Department of Traumatology, Leiden University Medical Centre, Leiden, the Netherlands

Correspondence
Floriana Samantha Luppino, Eurocross Assistance, Dellaertweg 1, 2316 WZ Leiden, The Netherlands.
Email: floriana.luppino@eurocross.nl

Abstract

Skiing and snowboarding are both popular recreational alpine sports, with substantial injury risk of variable severity. Although skills level has repeatedly been associated with injury risk, a validated measure to accurately estimate the actual skills level without objective assessment is missing. This study aimed to develop a practical validated instrument, to better estimate the actual skills level of recreational skiers, based on the criteria of the Dutch Skiing Federation (DSF), and covering five different skill domains. A sample of Dutch recreational skiers ($n = 84$) was asked to fill in a questionnaire reflecting seven, a priori chosen predictors by expert opinion, to ski downhill and to be objectively evaluated by expert assessors. The instrument was developed to have a multidimensional character and was validated according to the TRIPOD guideline (Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis). The sample reported an overall incorrect self-reported estimation of their skills, compared with the observed skill score. The instrument showed good calibration and underwent multiple validation methods. The estimated skills score showed to be closer to the observed scores, than self-reportage. Our study provides a practical, multidimensional, and validated instrument to estimate the actual skills level. It proved to better reflect the actual skills levels compared with self-reportage among recreational skiers.

KEYWORDS

prediction model, recreational skiers, skiing, skills level assessment, validation

1 | INTRODUCTION

Alpine skiing is a popular recreational sport, with a significant injury risk, but with considerably varying injury rates, in both mild and severe injuries.[1–4] Based on the concomitant impact on the individual burden of disease and healthcare systems, many epidemiological studies aimed to identify which risk factors contribute to the risk of (severe) injury.

An important risk factor that has repeatedly been associated with injury risk is skills level. The results however remain inconclusive. Among the less experienced skiers, increased injury risk has been described for overall injury risk,[5–8] and for more severe injuries.[9] Among more skilled skiers, higher injury levels have been reported,[10] also for severe injuries.[11–13]

The inconsistent results may be explained by the heterogenous skills level assessment. First, the subdivision of categorical instruments differ, varying from three[8] to five categories[10,12] not clearly defining the criteria for the categorical assignment.

Second, most skills level assessments are based on self-reported scores, which differ from observed scores, both among skiers and in other populations, and with respect to different performance outcomes: general fitness level among a general population,[14] physical activity among skiers[15] and skiing ability among recreational skiers.[16,17] Sulheim et al. validated five self-reported questions intended to measure actual skills level, reporting a poor to moderate correlation. Within the context of self-reportage, the perception of one's abilities is generally incorrect. The risk of bias between the perceived and actual performance is captured by the Dunning–Kruger effect (DKE),[18] suggesting that low ability leads to a significant overestimation, while a good ability leads to an accurate estimation (or marginal underestimation). Existing scales do not correct for the DKE-effect.

Third, to our knowledge, only one study reported assessing skiing skills level through a validated scale,[12] suggesting that its assessment in other studies was not validated. Instrument validation is important in itself to establish its reliable use,[19] but even more so in self-reported scales, in light of the DKE-effect and the inconsistent and poor correlations with observed values.

An accurate risk factor assessment is essential when aiming at risk prevention. Without an accurate skills level assessment, its association with injury risk lacks accuracy, potential specific preventive strategies may be overlooked, and therefore, its effect on the reported association, is disputable. To investigate its actual influence on injury risk, skills level should ideally be assessed objectively. As this is in practice not achievable and self-reportage inaccurate, a practical validated instrument is needed to better estimate actual skills level. Additionally, as the technical skills embrace and reflect abilities of other domains, its character should be multidimensional.

In this study, we aimed to develop a validated and multidimensional instrument for a more accurate estimation of the actual skills level, compared with self-reportage. This may serve as a first step in establishing a new standard for measuring skiing ability in future epidemiological studies.

2 | MATERIAL AND METHODS

2.1 | Study design and participants

This cross-sectional observational study was based on a Dutch sample of recreational skiers. The study was approved by the board of Eurocross Assistance, and all participants provided written informed consent.

The aim was to examine at least 6 predictors. Further, we estimated that an inclusion of 50 subjects a day would be achievable. We therefore aimed to recruit at least 100 Dutch participants. As injuries occur across all ages and levels, these were not considered exclusion criteria. We only considered important that participants were able to understand the questions. This was done upon reasonable estimation by the recruiters.

Participants were recruited at two different time points. The first sample was obtained during a pilot study, including 18 participants skiing in a Dutch indoor ski hall with artificial snow, October 2019. The second sample ($n = 66$) was recruited in Flachau, Austria, February 2020.

Both study design and data assessment were developed in consultation with an expert team: co-author and trauma surgeon K.B., head instructor of the Dutch Ski Federation (DSF) R.M. and personal trainer with special winter sport accreditation J.M.

2.2 | Data assessment

The pilot study, needed to identify and overcome practical flaws or unforeseen issues, was performed in a Dutch artificial snow indoor ski hall. Despite the static slope conditions, the needed technique is most similar to the one used on real slopes. Indoor treadmill ski halls present equal static conditions, but require a longer adaptation period, even among experienced skiers, thus resulting less fit for this study. The main study was performed during two days in the mountains (Flachau, Austria), under natural snow and slope conditions.

After giving informed consent, participants were asked to rate themselves (self-reported skill score, SRSS) and to fill in a questionnaire, intended to assess the predictors. Afterwards, participants were instructed to descend the appointed slope, with low to intermediate difficulty, in order not to exclude more unexperienced subjects. Their observed skill score (OSS) was independently assessed by two registered ski-instructors and was calculated by taking the mean value of both instructors.

Apart from the different slope conditions, assessments were identical in both samples. No a priori exclusion criteria were applied.

2.3 | Outcome and predictors

The OSS was considered to reflect the actual skill score, and to be the golden standard.

The OSS was based on the technical skills. For their systematical assessment, we developed an evaluation form (Supplementary material), based on the DSF evaluation methodology.[20] Several technical aspects were evaluated (central balance, inclination of the ski's, shape of the turn, level of parallel skiing, symmetry of movement, level of knee joint flexion, core stability, choice, and control of the speed), resulting in an overall score (i.e., the OSS). After evaluation of the pilot study, the items "level of knee joint flexion" and "technical preparation" were added to the Flachau sample.

The concept of "a good skier" is not yet clearly described. Our expert team defined it as someone able to handle him/herself under any circumstance. In line with this definition, the predictors were a priori chosen to cover five skill domains (i.e., self-reported skill level score, experience, avoidant behavior, equipment, and preparation), based on experience of R.M. and J.M. and relevant literature. First, the self-reported skill score was included as predictor. Second, since men seem to be more prone to injury, gender was included.[21] Third, we chose the number of skiing days per year, and not the number of skiing years, as predictor for experience, as some people do not ski every day of their holiday. Fourth, more unexperienced skiers tend to avoid skiing on cloudy or snowy days; therefore "avoidance of certain weather conditions" was included. Fifth, among all weather conditions, foggy circumstances demand more of one's skills, as fog impairs one's vision. Sixth, we reasoned that less experienced skiers would not be inclined to buy their own gear, and underestimate the importance of proper boot fitting, also in terms of risk reduction.[8] Seventh, experienced skiers often take more steep or less well prepared slopes. This requires a certain strength of the upper legs. A good physical preparation, ensures that the technical performance is maintained for a longer period of time, thus reducing the risk of exhaustion and the corresponding risk of falling.[22] We expected experienced skiers to be more aware of this, and therefore more physically prepared. The questionnaire, developed to assess the predictors (Supplementary material), comprised more questions than used in this study, to be able to use them in future investigations.

Although age has been suggested to play a role in both overestimation[23] and injury risk, being higher among youth and elderly,[21] the expert team decided not to immediately add "age" as a predictor. In their experience, the chosen predictors were thought to be more relevant.

Reporting was done according to the TRIPOD guidelines.[24]

2.4 | Statistical analyses

All statistical analyses were performed using IBM SPSS Statistics version 26 (IBM, New York, NY, USA). Analyses were performed on the total sample (i.e., the pilot and Flachau sample combined). General sample characteristics were analyzed using means and standard deviations (SD) for quantitative variables and percentages for categorical variables. Missing data were excluded from calculations. To investigate potential differences between the pilot and the Flachau sample, independent-samples t-test for continuous variables or the χ^2-test for categorical variables were performed. To estimate to which extent the self-reported scale showed a good fit with the observed scores, this association was plotted in a scatter plot and in a Bland–Altmann plot to investigate the closeness of agreement. To investigate agreements between instructors, the Cohen's Kappa coefficients were calculated.

To develop the skills level estimation instrument, a multivariable linear regression analysis was performed using the seven a priori chosen predictors and the OSS as outcome, referred to as the "expert opinion model" (EOM).

The performance of the model was assessed through measures of goodness of fit and calibration, in line with the TRIPOD-guidelines.[19] To assess the model fit we calculated the R^2, the adjusted R^2, and the mean squared prediction error (MSPE). To assess the calibration of the model, we constructed the calibration plot, and computed the calibration-in-the-large and the calibration slope. In addition, we calculated the percentage of people for whom our predicted score was within 1 point and within 0.5 point of the observed score.

To assess the validity of our model, we performed several validation analyses. First, we internally validated the model through a random split procedure (i.e., random spit model, RSM). We randomly divided our sample in two halves, estimated our model in one half and assessed all performance measures in the other half. In addition, we assessed the performance of our original "expert opinion model" stratified by sample (Flachau and Pilot). Next, as external validation, we estimated our model in the larger Flachau sample, and assessed its performance in the pilot sample (i.e., external validation model, EVM).

Next, to further assess robustness of our results, we performed a sensitivity analysis where we estimated a model using a backward selection procedure of all potential predictors, as opposed to our a priori selection of predictors (backward selection model, BSM). Again, performance of the model was assessed.

Finally, we assessed whether our newly developed and validated instrument outperformed the self-reported skills score in terms of agreement with the actual OSS, by calculating for what percentage of cases our score was closer to the OSS than the self-reported score.

3 | RESULTS

3.1 | General descriptives

Sample characteristics are presented in Table 1. In the end, 84 participants were included. Approximately 57% of the sample was male, 79.8% was adult. In one case, information about age was missing. The mean number of lessons was 69.1 hours, the mean experience 15 years, with a mean number of skiing days 8.5 a year, and 5.7 hours a day. 42.9% stated to know the rules on the slopes. 9.5% prepared themselves physically. The mean self-reported score was 7.1 (SD 1.0) versus a 6.0 (SD 1.4) observed skill score (OSS). 78.6% of all cases overestimated their skill level, with a mean level of 1.1 points (SD 1.0). In 78,6% of the cases, no disagreement between the instructors was present. Cohen's Kappa coefficient for the instructors' OSS was 0.76.

When comparing the Flachau to the pilot group, both groups significantly differed with respect to gender, mean self-reported and OSS, number of hours a day spent skiing, avoidance of "certain slopes" and black slopes, ski boot rental, knowledge about the slope rules, physical preparation, and all technical aspects.

Figure 1 illustrates the association between the self-reported score and the OSS. The Bland–Altmann plot (Figure 2) additionally showed that the limits of agreement for the self-reported skill score and the OSS were not met. Additionally, the mean difference between self-reported and observed scores was significantly different from zero ($p = <0.001$), indicating that the self-reported score was not representative for the actual objectified level, and decreased towards a higher mean OSS.

3.2 | Skills level estimation instrument: development and validation of the "Expert Opinion Model"

The multivariable regression analysis with the a priori chosen predictors as independent variables resulted in the following "Expert Opinion Model."

$$-0.850 + (\text{self} - \text{reported skill score}^{*}0.969) + (\text{gender}^{*} - 0.510) + (\text{skiing daysperyear}^{*}0.037) + (\text{avoidance offog}^{*} - 0.121) + (\text{avoidance of certain weather conditions}^{*} - 0.611) + (\text{skiboot rental}^{*}0.491) + (\text{physical preparation}^{*}0.779).$$

Explanation of the model: Female = 0, Male = 1.
For other dichotomous variables: No = 0, Yes = 1.

All performance measures are shown in Table 2. The estimated coefficients and standard errors of all models are shown in supplementary Table S1. All calibration plots in supplementary Figure S2.

The goodness of fit of the EOM is moderate and decreases in the external validation. However, calibration of the model is good with calibration slopes consistently close to 1 and calibration in the large indicating close agreement of our instrument with OSS on average. This is reiterated by the calibration plots. In addition, in around 80% of cases, our instrument was within 1 point of the actual OSS, and in almost half of the cases, the difference was smaller than 0.5 point.

Finally, our instrument was closer to the OSS than the subjects' self-reported scores in 73.8% of the cases.

4 | DISCUSSION

In this study, we developed and validated an instrument to estimate a recreational skier's skills level.

Our instrument included seven predictors, chosen by expert opinion, and was developed based on a sample of 84 Dutch recreational skiers. The sample was mainly male, showed poor physical and technical preparation, reported low levels of ski boot rental and an incorrect estimation of its skills. Comparison of the Flachau and pilot sample showed the pilot group to be overall better prepared, to have more experience and better technical skills, and to be less avoidant.

To our knowledge, this study is first of its kind. Despite previous attempts to estimate the actual skiing ability based on self-reported items,[16,17] transparent validation has rarely been done. The importance of this study lies in the scope to accurately assess the association between skills level and injury risk, which, based on the suggested inaccurate self-reported skills level assessment, is still lacking.

4.1 | Robustness of the model

Several findings suggest that our instrument can be considered valuable. First, the R^2 of our model was 0.640. Although not perfect, it can be appreciated as "reasonable." Second, calibration of our model was good, as indicated by the calibration in the large, slope and plot, and our instrument was within 1 point of the actual scores in almost 80% of cases. Third, the calibration plot (Supplementary Figure S2) showed a (slightly) better agreement of our instrument with the OSS than of the self-reported score (Figure 1), indicating that our model better estimates the actual skill score than self-reportage. This was reiterated by the fact that our instrument was closer to the OSS than the self-reported score in 73.8% of the cases.

TABLE 1 General characteristics of the sample ($n = 84$), and comparison between Pilot and Flachau sample

	Total sample $n = 84$ (% or mean, SD)	Flachau $n = 66$	Pilot $n = 18$	p-value
Gender[‡]				
Male	48 (57.1%)	34 (51.5%)	14 (77.8%)	0.046*
Female	36 (42.9%)	32 (48.5%)	4 (22.2%)	
Age				
Mean age	34.9 (15.2)	35.2 (15.9)	33.8 (12.9)	0.69
Missing	1 (1.2%)	1 (1.5%)	0 (0.0%)	
0–18	16 (19.0%)	14 (21.2%)	2 (11.1%)	0.32
18+	67 (79.8%)	51 (77.3%)	16 (88.9%)	
Mean self-reported skill score (SRSS) (from 1 to 10)[‡]	7.1 (1.0)	6.9 (1.0)	7.9 (0.6)	<0.001*
Mean observed skill score (OSS) (from 1 to 10)	6.0 (1.4)	5.7 (1.4)	7.2 (0.80)	<0.001*
Overestimation				
Yes	66 (78.6%)	52 (78.8%)	14 (77.8%)	0.93
No	18 (21.4%)	14 (21.2%)	4 (22.2%)	
Experience				
Lesson, number of hours	69.1 (69.8)	70.3 (70.3)	65.2 (69.8)	0.79
Years of skiing	15.3 (11.3)	14.3 (10.8)	18.9 (12.7)	0.18
Skiing days per year[‡]	8.5 (6.8)	7.8 (4.0)	10.8 (12.7)	0.33
Hours per day	5.7 (1.1)	5.5 (1.0)	6.4 (1.2)	0.011*
Avoidance, (yes)				
Any avoidance	71 (84.5%)	57 (86.4%)	14 (67%)	0.37
Certain slopes	32 (38.1%)	31 (47.0%)	1 (5.6%)	0.001*
Black slopes	24 (28.6%)	24 (36.4%)	0 (0.0%)	0.002*
Certain snow conditions	36 (42.9%)	31 (47.0%)	5 (27.8%)	0.15
Icy slopes	20 (23.8%)	24 (36.4%)	4 (22.2%)	0.26
Certain weather conditions[‡]	19 (22.6%)	17 (25.8%)	2 (11.1%)	0.12
Fog	3 (3.6%)	3 (4.5%)	1 (5.6%)	0.86
Crowded spots	46 (54.8%)	35 (53.0%)	11 (61.1%)	0.54
Ski boot rental[‡]				
No, has own ski boots	79 (94.0%)	66 (100.0%)	13 (72.2%)	<0.001*
Yes, rents ski boots	5 (6.0%)	0 (0.0%)	5 (27.8%)	
Knows the rules				
Yes	36 (42.9%)	19 (28.8%)	17 (94.4%)	<0.001*
No	48 (57.1%)	47 (71.2%)	1 (5.6%)	
Physical preparation before skiing[‡]				
Yes	8 (9.5%)	2 (3.0%)	6 (33.3%)	<0.001*
No	76 (90.5%)	64 (97.0%)	12 (66.7%)	
Technical preparation before skiing[†]				
Yes	2 (3.0%)	2 (3.0%)	NA	–
No	64 (97.0%)	64 (97.0%)	NA	
Technical aspects				
Central balance	5.8 (1.6)	5.4 (1.5)	7.1 (0.8)	<0.001*

(Continues)

TABLE 1 (Continued)

	Total sample $n = 84$ (% or mean, SD)	Flachau $n = 66$	Pilot $n = 18$	p-value
Inclination of the ski's	5.8 (1.6)	5.5 (1.6)	6.8 (1.1)	<0.001*
Form of the curve	6.1 (1.7)	5.8 (1.7)	6.9 (1.0)	0.002*
Level of parallel skiing	6.5 (1.7)	6.3 (1.8)	7.4 (0.7)	<0.001*
Symmetry of movements	5.4 (1.7)	5.1 (1.6)	6.9 (1.0)	<0.001*
Level of knee joint flexion[†]	4.6 (1.7)	4.6 (1.7)	NA	–
Core stability	5.7 (1.8)	5.3 (1.7)	7.1 (0.9)	<0.001*
Speed choice and control	6.7 (1.7)	6.6 (1.8)	7.3 (0.9)	0.04*

Abbreviation: NA, not applicable.

[†]Flachau sample only.

[‡]A priori chosen predictors.

*$p < 0.05$.

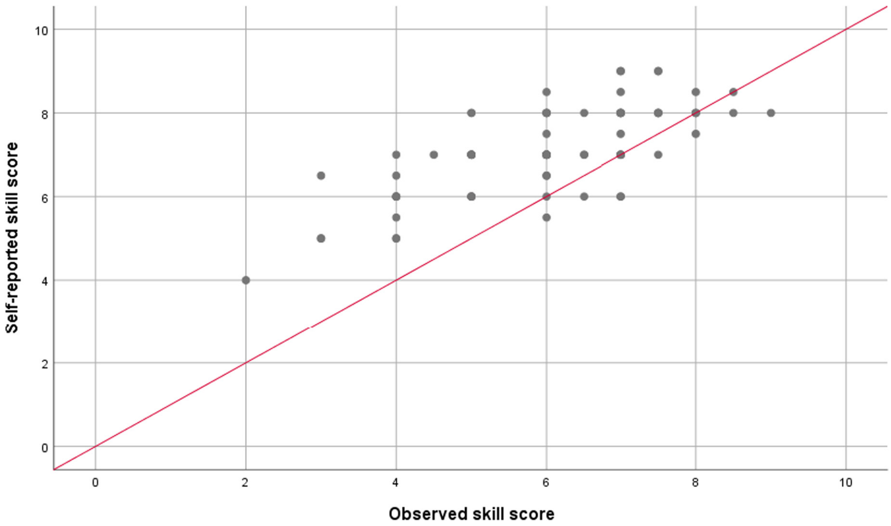

FIGURE 1 Scatter plot of self-reported and the observed skill score

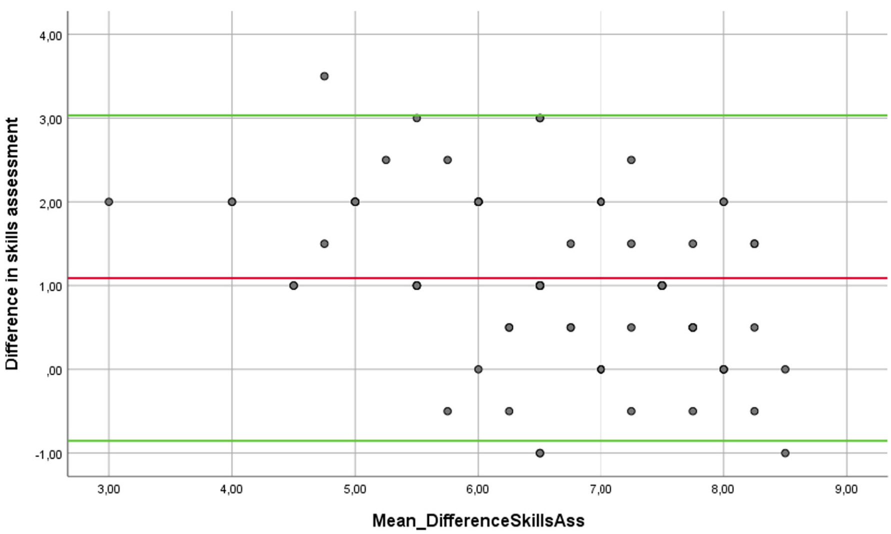

FIGURE 2 Bland–Altmann plot for self-reported and observed skill score. The red reference line represents the mean difference between SRSS and OSS. Both green lines represent the limits of agreement (+ 1.96 SD; − 1.96 SD)

4.2 | Additional testing

To further test our instrument, we performed several validation and sensitivity analyses, adding to face validity. Internal validation showed similar performance, reducing potential concerns of overfitting. Notably, stratification on the pilot sample showed the best performance measures of all. As shown in Table 1, the pilot group scored significantly higher on the OSS. This might explain the better performance of the model among the pilot group, most likely due to a better self-estimation in this group, in line with the DKE.

TABLE 2 Performance measures of all models

	Expert Opinion model (EOM)	Random Split (Internal validation) (IVM)	Stratified model on Flachau group	Stratified model on Pilot group	External validation model (EVM)	Backward selection model (sensitivity analysis) (BSM)
R^2	0.640	0.702	0.565	0.777	0.420	0.655
Adjusted R^2	0.635	0.695	0.520	0.621	0.384	0.651
Mean squared prediction error (MSPE)	0.723	0.594	0.846	0.271	0.455	0.661
Calibration slope	1.00	0.917	1.01	0.819	0.659	1.00
Calibration-in-the-large (Pred vs. Obs)	6.01 vs. 6.01	6.17 vs. 6.07	5.71 vs. 5.70	7.12 vs. 7.17	6.99 vs. 7.17	6.05 vs. 6.01
% with less than 1 point difference	78.6%	78.6%	75.8%	94.4%	83.3%	79.7%
% with less than 0,5 point difference	42.9%	46.4%	46.4%	46.4%	51.2%	51.2%

Abbreviations: Pred, predicted; Obs, observed.

In external validation, performance decreased. The explanation for the lower performance may lie in two factors. First, the model was developed based on the Flachau sample. As mentioned above, the Flachau and pilot sample differed significantly on several variables. That might explain why developing a model on the Flachau sample alone shows a lesser performance when testing it in the pilot group. Nonetheless, the purpose of the study was to develop a model applicable to a Flachau-like population, and not only a more prepared group.

Second, the fact that the predictor "ski boot rental" was negatively answered by all subjects in the Flachau sample, weakened the model validation. Additionally, the pilot sample size was very small, causing one value to have great impact.

Finally, as a check for our chosen predictors, we performed a statistical selection (BSM) as well, which largely resulted in the same, though longer, list of predictors. Although this model showed a slightly better performance, it is important to note that overfitting is a serious concern when statistically selecting predictors in small samples and a priori selection is therefore preferred.[25] Additionally, 4 out of the 9 predictors were about avoidance, and 2 about experience. Correlations between self-reportage and observed singular abilities show limited value,[17] suggesting an additional value of a multidimensional approach.

4.3 | Strengths and limitations

Our study presents some limitations. The major limitation is the sample size. First, this might have led to a lower accuracy, based on a limited power and goodness of fit. Second, it restricted the number of predictors we could include, thus possibly omitting other relevant predictors. Third, unlike others[8] "ski boot rental" was not found to be a proxy for lower skill scores. This might be attributable to the small sample, among whom 5 subjects reported to rent ski boots, while having a relative high OSS. Finally, we found significant differences between the pilot and Flachau sample. This seems to reflect actual differences but could also be the result of lack of power. The small sample size might be attributable to a number of reasons. First, the weather conditions: the second day, it was cold and it was snowing. Compared with the first recruitment day, the number of skiers who could have been asked was scarce. Second, the COVID pandemic just made its entrance in Europe, and the news started spreading that especially some winter sport resorts were subject to infection outbreaks. And third, our second and last recruitment day was also the second-last day of the Dutch holiday. Possibly, people were physically tired and had to prepare for the trip back home.

A second limitation was the chosen slope. Although a low to medium difficulty slope is more accessible for participation, it might not appeal strongly enough to one's skills, therefore having limited discriminative value. Third, among top and race skiers, the technical skill level is objectively high, and injury prevention focusses on further increasing one's skills, physical strength, and further equipment improvement.[26] Therefore, this model is not applicable to assess or compare a professional skier's ability.

Third, the linear assumption might not be appropriate. The DKE, trades of which are seen in our sample, suggests non-linearity. Therefore, this model might not be fully adequate for the more skilled recreational skiers and could be improved upon.

Despite the limitations, our study has several strengths, and sets an important first step towards a new skills level approach, providing a new path for future studies to improve and adjust this model. First, this model is first of its kind, proposing a better alternative for skills level assessment than self-reportage, without the need to objectively assess skiers. Second, the model's calibration performance was good, and was, extensively tested by various degrees of validation thus strongly adding to face validity. Furthermore, it was developed in line with the TRIPOD-guideline. Third, the model is multidimensional, covering multiple skills needed to fulfill the definition of "a good skier." Fourth, the OSS-assessment was reliable, likely due to the systematic evaluation form. Finally, misinterpretation, and more specifically, overestimation of circumstantial factors or perceived skills, have been reported in non-Dutch recreational downhill skiers before.[27–29] Although perceived ability could theoretically vary across different nationalities and cultures, our conclusions can be considered generalizable to the larger sample of recreational skiers. Combined with its practical nature, our instrument is directly and broadly applicable in studies that aim skills level examination and its role in relation to injury risk.

5 | CONCLUSION

Our study provides a practical, multidimensional, and validated prediction instrument of the actual skills level among recreational skiers. Its outcomes are more accurate than self-estimation. This model is a first step towards a new skills level assessment approach, and directly applicable in studies investigating the association between the skills level and ski injury.

6 | PERSPECTIVE

A replication of our study with a larger sample is advisable. As the number of included predictors reflects the sample size to prevent overfitting, the advisable sample size is dependent on the aimed analyses. Predictors as age or sport (snowboarding vs. skiing) would be relevant predictors, resulting in a sample size of at least 100–110 participants. Time and expenses permitting, a minimum of 200 participants would be ideal, in order to broader examine the role of other potential predictors.

First, a larger sample may give the opportunity to discriminate between self-reported level categories, where higher levels may better estimate the skiing ability. Second, more power could identify other important predictors, such as age, improving the model's accuracy, or allow a more in-depth examination on the discriminative value of different slopes.

Despite the smaller proportion and different injury patterns,[13] snowboarders show significantly higher incidence levels of injuries requiring hospital treatment.[30] Our focus lied on skiers, but in term of injury risk prevention, future research should also examine how (perceived) skills levels relate to injury risk among snowboarders.

To better evaluate the role of the skills level in injury risk, the current model could be implemented in existing studies. By replacing the self-reported skills level with our instrument, the outcomes can be compared allowing a better insight in the role of skills level in skiing risk injury. Applying our instrument in future studies may add to establish new epidemiological standards in the future.

ACKNOWLEDGEMENTS

We would like to thank the members of the expert team: co-author and trauma surgeon Kees Bartlema, for his clinical input and revision of the paper; Jasper Mulder, personal trainer with special winter sport accreditation, for his valuable input in the development of the questionnaire choice of predictors, and his active participation in participants recruitment. We also gratefully thank Richard Molenaar, head instructor of the Dutch Ski Federation, for his important role in the questionnaire development, choice of predictors and explanation on the importance of performing the pilot study in an indoor hall with artificial snow.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Floriana Samantha Luppino https://orcid.org/0000-0001-5061-0737

REFERENCES

1. McBeth PB, Ball CG, Mulloy RH, Kirkpatrick AW. Alpine ski and snowboarding traumatic injuries: incidence, injury

patterns, and risk factors for 10 years. *Am J Surg*. 2009;197:560-563. discussion 563–564.
2. Statistics Netherlands (Centraal Bureau voor de Statistiek, CBS). *Trend Report Tourism, Recreation and Leisure: Less Skiing and More Hiking during the Winter Holiday (Minder skien en meer wandelen tijdens wintersport)*. Centraal Bureau voor de Statistiek, CBS. 2018. Accessed September 13, 2022. https://www.cbs.nl/nl-nl/nieuws/2018/08/minder-skien-en-meer-wandelen-tijdens-wintersport
3. Stenroos A, Handolin L. Incidence of recreational alpine skiing and snowboarding injuries: six years experience in the largest ski resort in Finland. *Scand J Surg*. 2015;104:127-131.
4. Weinstein SKM, van Baak K. Common skiing and snowboarding injuries. *Curr Sports Med Rep*. 2019;18:394-400.
5. Bissell BT, Johnson RJ, Shafritz AB, Chase DC, Ettlinger CF. Epidemiology and risk factors of humerus fractures among skiers and snowboarders. *Am J Sports Med*. 2008;36:1880-1888.
6. Johnson RJ, Ettlinger CF, Shealy JE. Myths concerning alpine skiing injuries. *Sports Health*. 2009;1:486-492.
7. Sran R, Djerboua M, Romanow N, et al. Ski and snowboard school programs: injury surveillance and risk factors for grade-specific injury. *Scand J Med Sci Sports*. 2018;28:1569-1577.
8. Sulheim S, Holme I, Rødven A, Ekeland A, Bahr R. Risk factors for injuries in alpine skiing, telemark skiing and snowboarding--case-control study. *Br J Sports Med*. 2011;45:1303-1309.
9. Goulet C, Régnier G, Grimard G, Valois P, Villeneuve P. Risk factors associated with alpine skiing injuries in children. A case-control study. *Am J Sports Med*. 1999;27:644-650.
10. Coury T, Napoli AM, Wilson M, Daniels J, Murray R, Milzman D. Injury patterns in recreational alpine skiing and snowboarding at a mountainside clinic. *Wilderness Environ Med*. 2013;24:417-421.
11. Goulet C, Hagel B, Hamel D, Légaré G. Risk factors associated with serious ski patrol-reported injuries sustained by skiers and snowboarders in snow-parks and on other slopes. *Can J Public Health*. 2007;98:402-406.
12. Goulet C, Hagel BE, Hamel D, Légaré G. Self-reported skill level and injury severity in skiers and snowboarders. *J Sci Med Sport*. 2010;13:39-41.
13. Maat SC, Luppino FS, Schipper IB, Krijnen P, Bartlema KA. Injury patterns after skiing and snowboarding sports accidents. *J Sports Med Phys Fitness*. 2020;60:119-124.
14. Mikkelsson L, Kaprio J, Kautiainen H, Kujala UM, Nupponen H. Associations between self-estimated and measured physical fitness among 40-year-old men and women. *Scand J Med Sci Sports*. 2005;15:329-335.
15. Tzetzis G, Avgerinos A, Vernadakis N, Kioumourtzoglou E. Differences in self-reported perceived and objective measures of duration and intensity of physical activity for adults in skiing. *Eur J Epidemiol*. 2001;17:217-222.
16. Ruedl G, Brunner F, Woldrich T, et al. Factors associated with the ability to estimate actual speeds in recreational alpine skiers. *Wilderness Environ Med*. 2013;24:118-123.
17. Sulheim S, Ekeland A, Bahr R. Self-estimation of ability among skiers and snowboarders in alpine skiing resorts. *Knee Surg Sports Traumatol Arthrosc*. 2007;15:665-670.
18. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol*. 1999;77:1121-1134.
19. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162:W1-W73.
20. Dutch Ski Federation (Nederlandse Ski Vereniging). *What is Your Level? (Wat is jouw niveau?)*. Nederlandse Ski Vereniging. 2013. Accessed September 13, 2022. https://www.wintersport.nl/forum/topic/118761
21. Davey A, Endres NK, Johnson RJ, Shealy JE. Alpine skiing injuries. *Sports Health*. 2019;11:18-26.
22. Suchomel TJ, Nimphius S, Stone MH. The importance of muscular strength in athletic performance. *Sports Med*. 2016;46:1419-1449.
23. van der Ham IJ, van der Kuil MN, Claessen MH. Quality of self-reported cognition: effects of age and gender on spatial navigation self-reports. *Aging Ment Health*. 2021;25:873-878.
24. Collins GSRJ, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD). *The TRIPOD Statement*. 2015;350:g7594. Accessed September 13, 2022. https://www.tripod-statement.org/adherence/
25. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Springer Science and Business Media, LLC; 2009.
26. Spörri J, Kröll J, Gilgien M, Müller E. How to prevent injuries in alpine ski racing: what do we know and where do we go from here? *Sports Med*. 2017;47:599-614.
27. Bailly N, Abouchiche S, Masson C, Donnadieu T, Arnoux PJ. *Recorded Speed on Alpine Slopes: How to Interpret Skier's Perception of their Speed?*. Springer International Publishing; 2017:163-174.
28. Ruedl G, Posch M, Niedermeier M, et al. Are risk-taking and ski helmet use associated with an ACL injury in recreational alpine skiing? *Int J Environ Res Public Health*. 2019;16:3107.
29. Willick SE, Wagner G, Ericson D, Josten G, Teramoto M, Davis J. Helmet use and risk-taking behavior among skiers and snowboarders. *Clin J Sport Med*. 2019;29:329-335.
30. Rønning R, Gerner T, Engebretsen L. Risk of injury during alpine and Telemark skiing and snowboarding. The equipment-specific distance-correlated injury index. *Am J Sports Med*. 2000;28:506-508.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Luppino FS, den Hollander-Gijsman ME, Dekker FW, Bartlema KA, van Diepen M. Estimating skills level in recreational skiing: Development and validation of a practical multidimensional instrument. *Scand J Med Sci Sports*. 2023;33:55-63. doi: 10.1111/sms.14245