POLITECNICO DI TORINO
Repository ISTITUZIONALE

Evaluating objective measures of impairment to trunk strength and control for cross-country sit skiing

Original
Evaluating objective measures of impairment to trunk strength and control for cross-country sit skiing / Rosso, V.; Linnamo, V.; Vanlandewijck, Y.; Rapp, W.; Fasel, B.; Karczewska-Lindinger, M.; Lindinger, S.; Gastaldi, L.. - In: SPORTS ENGINEERING. - ISSN 1369-7072. - ELETTRONICO. - 24:1(2021). [10.1007/s12283-021-00342-5]

Availability:
This version is available at: 11583/2888738 since: 2021-04-12T09:39:42Z

Publisher:
Springer Science and Business Media Deutschland GmbH

Published
DOI:10.1007/s12283-021-00342-5

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Springer postprint/Author's Accepted Manuscript

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s12283-021-00342-5

(Article begins on next page)

12 October 2024

Sports Engineering (2021) 24:6
https://doi.org/10.1007/s12283-021-00342-5

ORIGINAL ARTICLE

Evaluating objective measures of impairment to trunk strength and control for cross-country sit skiing

Valeria Rosso[1,2] · Vesa Linnamo[2] · Yves Vanlandewijck[3] · Walter Rapp[4] · Benedikt Fasel[5] · Magdalena Karczewska-Lindinger[6] · Stefan Lindinger[7] · Laura Gastaldi[8]

Accepted: 4 March 2021
© The Author(s) 2021

Abstract
In Paralympic cross-country sit skiing, athlete classification is performed by an expert panel, so it may be affected by subjectivity. An evidence-based classification is required, in which objective measures of impairment must be identified. The purposes of this study were: (i) to evaluate the reliability of 5 trunk strength measures and 18 trunk control measures developed for the purposes of classification; (ii) to rank the objective measures, according to the largest effects on performance. Using a new testing device, 14 elite sit-skiers performed two upright seated press tests and one simulated poling test to evaluate trunk strength. They were also subjected to unpredictable balance perturbations to measure trunk control. Tests were repeated on two separate days and test–retest reliability of trunk strength and trunk control measures was evaluated. A cluster analysis was run and correlation was evaluated, including all strength and control measures, to identify the measures that contributed most to clustering participants. Intraclass correlations coefficients (ICC) were $0.71 < \text{ICC} < 0.98$ and $0.83 < \text{ICC} < 0.99$ for upright seated press and perturbations, respectively. Cluster analysis identified three clusters with relevance for strength and balance control measures. For strength, in upright seated press peak anterior pushing force without backrest (effect size = 0.77) and ratio of peak anterior pushing force without and with backrest (effect size = 0.72) were significant. For balance control measures, trunk range of motion in forward (effect size = 0.81) and backward (effect size = 0.75) perturbations also contributed. High correlations ($-0.76 < r < -0.53$) were found between strength and control measures. The new testing device, protocol, and the cluster analysis show promising results in assessing impairment of trunk strength and control to empower an evidence-based classification.

Keywords Evidence-based classification · Paralympics · Sitting sport · Core stability · Nordic skiing

1 Introduction

In cross-country skiing, athletes with structural or functional impairment at the lower limbs, pelvis or trunk compete on a sit-ski. As these sit-skiers have full upper limb function, they generate propulsion with two poles using the double poling technique. To compete in cross-country sit skiing, athletes have to be affected by at least one of the seven eligible physical impairments. Eligible physical impairments for cross-country sit skiing are impaired muscle power, impaired passive range of movement, limb deficiency, leg length difference, hypertonia, ataxia, and athetosis.

✉ Valeria Rosso
valeria_rosso@polito.it

1 Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin, Italy

2 Biology of Physical Activity, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

3 Department of Rehabilitation Sciences, KU Leuven, Leuven, Belgium

4 Olympic Training Centre Freiburg, Freiburg, Germany

5 Laboratory of Movement Analysis and Measurement, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

6 Department of Physiology, Center of Health and Performance at the Department of Food and Nutrition and Sport Science, University of Gothenburg, Göteborg, Sweden

7 Center of Health and Performance, Department of Food and Nutrition and Sport Science, University of Gothenburg, Göteborg, Sweden

8 Department of Mathematical Sciences, Politecnico di Torino, Turin, Italy

Published online: 25 March 2021

With a view to assuring equal and fair competition, sit-skiers with different eligible physical impairments go through a functional classification process. As a results, skiers are appointed to one of the five classes LW10, LW10.5, LW11, LW11.5, and LW12 [1] according to the effects the impairment has on cross-country sit skiing performance [2]. Athletes in class LW10 are unable to maintain an unsupported sitting position. Athletes in classes LW10.5, LW11, and LW11.5 have some ability to control the trunk and to keep balance while sitting that increases with class level. Athletes in class LW12 have complete trunk function [3].

Functional classification has the advantage to be sport-specific, although it is subjective. Athletes are physically and technically judged by at least two experienced classifiers, as described in the World Para Nordic Skiing Classification Rules and Regulations [3]. In addition, athletes may be visually observed during a competition by the classification panel [3]. High precision and reliability may be challenging to achieve with the subjective current classification system for cross-country sit-skiers, making the validity of the method questionable [2]. Indeed, in 2007, the International Paralympic Committee recognized the need for an objective and evidence-based classification for all sports, by stating it in the International Paralympic Committee Classification Code [4].

For each Paralympic sport, evidence-based classification aims to define eligible types of impairment, to set minimum impairment criteria, and to identify and profile classes [5]. To achieve these goals, it is mandatory to develop (i) reliable measures of impairment for classification purposes and, (ii) sport-specific measures of performance determinants [5]. Dealing with measures of impairment, core stability is a key issue for many seated sports. Methodology to evaluate trunk impairment can be applied to different sports (e.g., wheelchair racing, wheelchair rugby, and para-canoeing) [6, 7]. Trunk strength impairments affect force generation during maximal voluntary contractions [8]. Trunk control impairments influence trunk balance and range of motion during unpredictable balance perturbations [9].

Several Paralympic sports have initiated evidence-based classification research taking into consideration measures of both physical impairment and performance. Among them there are wheelchair racing [10, 11], wheelchair rugby [12–15], athletics [16, 17], throwing [18], and swimming [19]. In cross-country sit skiing few studies aim to propose scientific evidence to support a classification system. Regarding measures of impairment, unpredictable balance perturbations in the sagittal plane have been used to assess trunk control ability [20]. In [21–23], an adapted and sensorized ergometer was used for measuring performance determinants. To the best of our knowledge, only one study has been conducted on the relationship between impairment and performance, and it involved athletes with intellectual impairment [24].

Measures of impairment reported in [20] represent an advancement towards an evidence-based classification process for cross-country sit skiing. However, this study has some limitations related to standardization. Participants were tested using their personal sit-ski, with a customized seat and personalized constraints, such as straps and backrest. Based on these considerations, in this work, a new device to measure trunk strength and trunk control impairments in standardized conditions [25] is presented.

The present study aimed first to assess if the proposed measurements of trunk strength and control could be used for classification purposes, in compliance with the Position Stand [2] and following documents [5, 26, 27]. The second purpose was to identify the most influential measures to objectively assess trunk strength and control. A cluster analysis [20, 23] was used to group athletes based on impairment and performance. Then, a minimum set of measures was defined to assess trunk strength and trunk control. The authors hypothesized that the use of a cluster analysis would group sit-skiers according to trunk impairments. Indeed, trunk strength and trunk control affect the ability of a sit-skier to generate propulsive force and balance on the sit-ski, respectively.

2 Methods

2.1 Participants

Fourteen elite cross-country sit-skiers (nine males and five females, mean ± standard deviation: 32 ± 6 years, 160 ± 18 cm, 55 ± 13 kg) volunteered as participants. Participants had different impairment conditions (spinal cord injury = five, spina bifida = two, lower limbs amputation = seven). According to the functional classification process [1], they were assigned to the classes: LW10.5 = one, LW11 = two, LW11.5 = three, and LW12 = eight. Research methods and protocols were approved by the University of Jyväskylä Ethics Committee. The procedures were performed in accordance with the Declaration of Helsinki. Before the tests, participants were informed about the test purpose and procedures and they signed a consent form.

2.2 Materials

The tests took place in the Vuokatti Sports Technology Unit laboratory (University of Jyväskylä, Finland). The testing device was composed of a seat mounted on a motorized sledge and surrounded by an aluminum frame (Fig. 1). The device was integrated with a rubber cover for the seat to prevent participant from sliding or falling, two belts to fix their

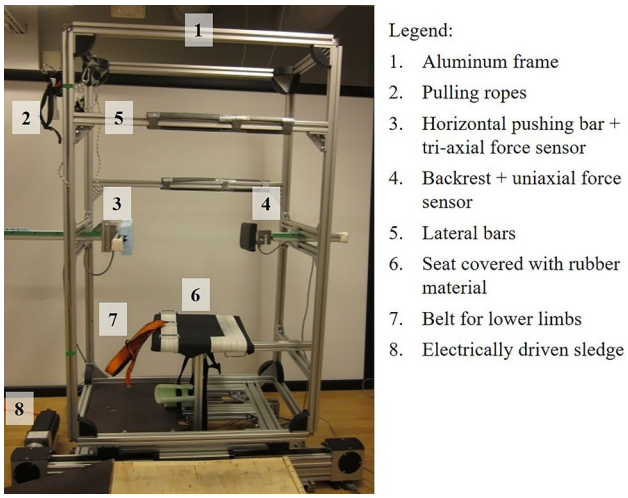

Fig. 1 A new device for testing trunk strength and control impairment in standardized conditions

pelvis and thighs to the seat, two padded belts to hold their shanks, and two lateral bars to prevent lateral falls (Fig. 1).

A horizontal pushing bar and a backrest were fixed to the frame, anteriorly and posteriorly, respectively. Both the bar and the backrest were adjustable in height and depth to fit with the participants' anthropometry. A tri-axial force sensor (K3D120; Me-Meßsysteme GmbH, Germany, full range 2 kN, resolution 0.4 N) was mounted between the horizontal pushing bar and the frame. A uniaxial force sensor (TB5; Lahti Precision, Finland, full range 5 kN, resolution 0.5 N) was mounted between the backrest and the frame. Two ropes with instrumented handles (uniaxial force sensors, University of Jyväskylä, Finland, full range 1 kN, resolution 0.3 N) [28] were hung from the anterior top of the frame. All force signals were synchronized (CED Power1401-3; CED Ltd., England) and collected with a sample frequency of 100 Hz.

The sledge was electrically driven on a linear rail (LF 12S; Bosch Rexroth, Germany). It was actuated by a three-phase motor (MSK060C; Bosch Rexroth, Germany) and a servo drive (HCS01.1E; Bosch Rexroth, Germany). Sledge movement parameters (acceleration, velocity, and amplitude) were set by an operator using LabVIEW (National Instruments, USA). Maximum sledge acceleration, velocity, and displacement were set to ± 3 m/s^2, ± 1 m/s, and ± 0.4 m, respectively.

An inertial sensor (Physilog® 4; GaitUp, Switzerland) was fixed on the posterior right corner of the sledge, with another on the participants' seventh cervical vertebra. The inertial sensors sampled 3D acceleration (full range ± 16 g, sensitivity 2.048 LSB/g) and 3D angular velocity [full range ± 2000°/s, sensitivity 16.4 LSB/(°/s)] at 500 Hz. The sensors were synchronized by an electronic trigger [29].

2.3 Protocol

All the tests were conducted during the International Paralympic Committee World Cup in March 2016 in Vuokatti (Finland). The protocol consisted of isometric maximum voluntary contraction tests to evaluate strength impairment, and balance control tests to evaluate trunk control impairment [25]. To evaluate test–retest reliability, maximal voluntary contraction and balance control tests were repeated twice (day 1 and day 2) with 24–36 h in between. To assure standard and comparable measures between days, no training or competitions were sustained by the participants in the 4 h before the tests. Tests conditions and the participants' posture were standardized to increase methodological consistency, reducing any effect of external factors.

The maximal voluntary contraction tests (Fig. 2) were: (a) upright seated press without using the backrest, (b) upright seated press using the backrest, and (c) simulated poling in which the participant pulled the two ropes isometrically.

The pelvis was fixed to the seat by the two straps to prevent the participants from sliding on the seat. The lower limbs were suspended under the seat using two padded belts. Belts avoided voluntary or involuntary legs

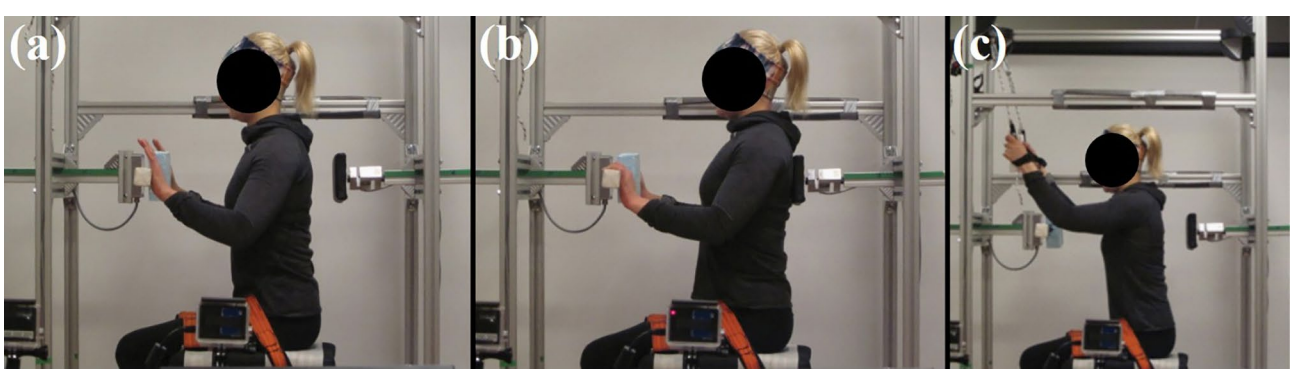

Fig. 2 Isometric maximal voluntary contraction tests conditions: **a** upright seated press without using the backrest; **b** upright seated press using the backrest; **c** simulated poling

movements and prevented the participants from pushing their feet against the ground to generate more force by the upper body.

In the upright seated press tests (a) and (b), participants were requested to keep their hands on the anterior pushing bar. Intra hands distance was defined to have shoulder abduction close to zero in the frontal plane. The anterior pushing bar was then regulated to fit participants' anthropometry: (i) height corresponding to the participants' mesosternum, (ii) antero-posterior distance set so that in the sagittal plane flexion angles were 90° for the elbow, and 30° with respect to the trunk for the shoulder.

The participant's back was in contact with the backrest in test (a) and 75 mm away from the backrest in test (b). The distance was set using a calibrated block interposed between the participant's back (straight trunk) and the backrest. The block was removed during data collection. Therefore, in upright seated press with backrest, participants could use the backrest to generate more force. In upright seated press without backrest, with the trunk straight and no backrest, participants generated force mostly relying on their core muscles.

The poling test (c) was designed to simulate the beginning of the pushing phase during double poling on flat terrain [30]. The arms were posed in the sagittal plane with flexion angles of 120° for the elbow, and 90° with respect to the trunk for the shoulder. For all the maximal voluntary contraction tests, participants were instructed to exert the maximum force and to hold it for 3 s. Each participant performed three repetitions for each test, with an interval of 30 s between repetitions. A resting interval of 5–10 min separated maximal voluntary contraction tests.

The balance control test consisted of forward and backward unpredictable perturbations, obtained using the electrically driven sledge. Participants were seated on the testing device, with their pelvis and lower limbs strapped, as for the maximal voluntary contraction tests. Participants were instructed to maintain their trunk in a straight position and their arms in a neutral position, relaxed along the body. Participants were forbidden to use their arms and hands to keep or recover the initial position.

The balance control test was performed imposing to the sledge three accelerations: ± 0.5, ± 1, and ± 2.5 m/s^2 and a stroke of ± 0.3 m. For each acceleration, five forward and five backward stimuli were given. Stimuli were administered in a random order and inter-stimuli time. In general, forward (backward) stimuli produce a backward (forward) trunk movement before physiological control mechanisms start influencing the movement. After each stimulus, the sledge returned to the initial position, in the middle of the rail.

2.4 Data analysis

For each maximal voluntary contraction test, the best attempt of the three repetitions (highest force peak) using unfiltered data was considered for analysis. The maximal voluntary contraction tests gave five measures of strength. The upright seated press without backrest yielded one measure of strength: the peak of anterior pushing force (PFa_{wo}). The upright seated press with backrest, yielded two measures of strength: (i) the peak of anterior pushing force (PFa_w) and (ii) the peak of posterior pushing force (PFp_w). For PFa_{wo} and PFa_w, the magnitude of the generated force measured along the three axes was calculated, while the PFp_w was measured along the longitudinal axis. The simulated poling yielded one measure of strength that was calculated as the mean value between the left and right peaks of the pulling force (PF_p). The fifth measure of strength was calculated as the ratio between PFa_{wo} and PFa_w (PFa_{wo}/PFa_w).

In the balance control test, the inertial sensor data were used to compute measures of trunk control. The inertial sensors were functionally calibrated based on an assisted upright posture (to define the trunk's longitudinal axis) and on controlled forward–backward sledge movements (to define the trunk's anterior–posterior axis) [29, 31]. Segment orientation was obtained with the strapdown integration with static drift correction and the azimuth orientation constrained to zero [32]. Trunk angle was computed as the dot product between trunk's longitudinal axis and vertical axis in the fixed global frame (e.g., gravity). Trunk forward movements were defined as positive, whereas backward movements were defined as negative.

The balance control tests gave 18 measures of trunk control. For each of the three acceleration values (0.5, 1.0, 2.5 m/s^2) and two perturbation directions (backward and forward), the following measures were evaluated: (i) trunk angle before perturbations ($Trunk_{rest}$), (ii) trunk reaction time (TRT), and (iii) trunk range of motion in the sagittal plane ($Trunk_{rom}$). $Trunk_{rest}$ was calculated as the mean of trunk angle in the 500 ms before each backward ($Trunk_{rest\text{-}b}$) and each forward ($Trunk_{rest\text{-}f}$) perturbation for the three accelerations. TRT was calculated as the time between when the trunk started moving and when it inverted the motion. These two instants were automatically detected on the trunk angular velocity signal, respectively, as (i) when the signal started increasing sharply from a baseline value and (ii) when the signal crossed the zero. $Trunk_{rom}$ was calculated as the difference between (i) the trunk angle when the trunk inverted the motion for the first time and (ii) the $Trunk_{rest}$.

For each participant and acceleration value, the mean of $Trunk_{rest}$, TRT, and $Trunk_{rom}$ was calculated for backward ($Trunk_{rest\text{-}b}$, TRT_b, and $Trunk_{rom\text{-}b}$) and forward ($Trunk_{rest\text{-}f}$, TRT_f, and $Trunk_{rom\text{-}f}$) perturbations. All analyses were performed using custom script in MatLab Software (MatLab

and Release 2015, The MathWorks, Inc., Natick, Massachusetts, USA).

2.5 Statistical analysis

Tests repeatability between days 1 and 2 was calculated to assess whether the proposed measures of impairment could be used for classification purposes. To assess test–retest reliability, statistical difference between the 2 days (Wilcoxon test), intraclass correlation coefficient (ICC, 2-way mixed, 95% confidence interval), and standard error of measurement were computed for all the analyzed measures [33, 34]. ICC was defined as poor ($ICC < 0.50$), moderate ($0.50 \leq ICC < 0.75$), good ($0.75 \leq ICC < 0.90$), or excellent ($ICC \geq 0.90$) [34].

To identify the most important measures that naturally group participants according to their trunk strength and control, a cluster analysis was performed on measures of strength (maximal voluntary contraction tests) and control (balance control tests). A similar cluster analysis approach was previously adopted to group cross-country sit-skiers. Trunk impairment in balance perturbation test [20] and performance during simulated action of poling on the ergometer test [23] were considered. In addition, a similar cluster analysis approach was used to evaluate strength impairment in wheelchair racers [6].

Data of the 14 participants in day 1 were considered. Before conducting the cluster analysis, data were prepared as follows. Data were checked for outliers using the method of ± 3 standard deviations. Only measures with a coefficient of variability over 5% were included in the cluster analysis. Because the measures had different scales, before running the cluster analysis they were normalized using the z score.

A k-means cluster analysis [35] was run to empirically categorize participants on the basis of their trunk strength and control. To run k-means cluster analysis is required to predefine the number of clusters. Studies on classification in cross-country sit skiing [20, 23] have worked best with two or three clusters. Therefore, in the present study, it was hypothesized a priori to divide participants into three clusters (cluster_1, cluster_2, and cluster_3). With respect to considering only two clusters, this solution is expected to represent a greater variance of impairment (i.e., absent, partial, or full trunk control).

The cluster analysis was validated using internal and external criteria. For the internal validation, principal component analysis [35] was used to represent data in the space of the first two principal components to visualize the formation of clusters. For external validation, k-means output (participants gathered in clusters) was compared to a priori information [36] (participants gathered in groups). So, three groups of participants were a priori defined according to their trunk control ability as described in [1]. Group_1: participants with absent or limited trunk control and ability to keep balance (LW10.5). Group_2: participants with fair trunk control and ability to keep the balance (LW11). Group_3: participants with complete or near to complete trunk control and ability to keep balance (LW11.5–LW12) [1]. The external validation was assessed to calculate accuracy, precision, and sensitivity [37] for the used k-means classifier. In addition, the macro-averaging F score was calculated to evaluate the harmonic mean between precision and sensitivity for the classifier. The score ranges between 0 and 1 [38].

The contribution of each measure in clusters formation was quantified using Kruskal–Wallis test (Fisher's least significant difference post hoc) and the effect size was assessed. Statistical significance for the analysis was set at an alpha of 0.05 (MatLab and Release 2015, The MathWorks, Inc., Natick, Massachusetts, USA). To determine statistical power, a post hoc power analysis (software = GPower; test = ANOVA: fixed effects, omnibus, one-way test; $\alpha = 0.05$) was conducted on the measures that contributed most to clustering athletes. The Spearman correlation was calculated between maximal voluntary contraction and balance control test measures.

3 Results

Concerning the assessment of measurements repeatability, very good results were found for test–retest reliability (Table 1). Test–retest reliability was evaluated on 12 out of 14 participants, because two of them performed the test protocol only on day 1. For all maximal voluntary contraction and balance control measures no significant differences were found between days 1 and 2.

ICC was good to excellent for all maximal voluntary contraction measures ($0.71 < ICC < 0.98$), with the lowest value for PFa_{wo}/PFa_{w} ($ICC = 0.71$). In the balance control test, for both stimuli directions good to excellent values of ICC were found for $Trunk_{rom}$ at the three accelerations ($0.94 < ICC < 0.96$) and for TRT for ± 1 and ± 2.5 m/s^2 ($0.83 < ICC < 0.99$) (Table 1) [34]. In contrast, poor ICC was found for $Trunk_{rest}$ in backward and forward perturbations at the three accelerations, and for TRT in both perturbation directions at the lowest acceleration (0.5 m/s^2). Low standard error of measurement was found for all maximal voluntary control and balance control measures (Table 1).

Cluster analysis allocated participants in the three clusters as follow: one in cluster_1 (high impact of impairment), four in cluster_2 (middle impact of impairment), and nine in cluster_3 (low impact of impairment). No outliers were found in the data with all the measures showing a coefficient of variability higher than the threshold.

Table 1 Results of test–retest reliability between day 1 and day 2

Tests	Measures	ICC	Standard error of measurement
Maximal voluntary contraction	PFa_w	**0.98*****	44.9 N
	PFp_w	**0.95*****	80.7 N
	PFa_{wo}	**0.93*****	49.8 N
	PF_p	**0.94*****	33.0 N
	PFa_{wo}/PFa_w	**0.71***	0.1
Balance control at 0.5 m/s^2	$Trunk_{rest\text{-}b}$	− 0.17	2.4°
	$Trunk_{rom\text{-}b}$	**0.96*****	0.9°
	TRT_b	− 0.22	158.2 ms
	$Trunk_{rest\text{-}f}$	− 0.74	3.0°
	$Trunk_{rom\text{-}f}$	**0.94*****	0.6°
	TRT_f	0.30	192.8 ms
Balance control at 1 m/s^2	$Trunk_{rest\text{-}b}$	− 0.45	3.3°
	$Trunk_{rom\text{-}b}$	**0.95*****	1.2°
	TRT_b	**0.83****	27.8 ms
	$Trunk_{rest\text{-}f}$	0.10	2.6°
	$Trunk_{rom\text{-}f}$	**0.95*****	1.5°
	TRT_f	**0.90****	40.8 ms
Balance control at 2.5 m/s^2	$Trunk_{rest\text{-}b}$	0.03	4.2°
	$Trunk_{rom\text{-}b}$	**0.96*****	1.1°
	TRT_b	**0.99*****	12.6 ms
	$Trunk_{rest\text{-}f}$	0.03	4.2°
	$Trunk_{rom\text{-}f}$	**0.94*****	1.5°
	TRT_f	**0.97*****	15.4 ms

Intra-class correlation coefficient (ICC) and standard error of measurement are reported for maximal voluntary contraction tests and balance control tests

PFa_w anterior pushing force in upright seated press with backrest, *PFp_w* posterior pushing force in upright seated press with backrest, *PFa_{wo}* anterior pushing force in upright seated press without backrest, *PF_p* mean pulling force between the left and right peaks in simulated poling, *$Trunk_{rest}$* trunk at rest, *$Trunk_{rom}$* trunk range of motion, *TRT* trunk reaction time, *b* backward perturbation direction, *f* forward perturbation direction

Statistical significance: ***$p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Therefore, all five measures of strength and the 18 measures of trunk control were considered in the k-means.

Internal validation results were represented for visualization purposes in the space of the first two principal components (Fig. 3) and three clusters were clearly visible, supporting the hypothesis. External validation results showed an accuracy of 86%, and results for precision and sensitivity are reported in Table 2. The lowest precision was found for cluster_2 (50%) and the lowest sensitivity for group_3 (81.8%). Macro-averaging *F* score was equal to 0.89, showing an overall balance between precision and sensitivity for the classification model used.

For each identified cluster, means ± standard deviations for all measures were calculated (Table 3). Kruskal–Wallis test and effect size allowed identification of the most relevant measures in clusters aggregations (last two columns of Table 3).

Concerning the maximal voluntary contraction test, PFa_{wo} and PFa_{wo}/PFa_w were the two measures that contributed

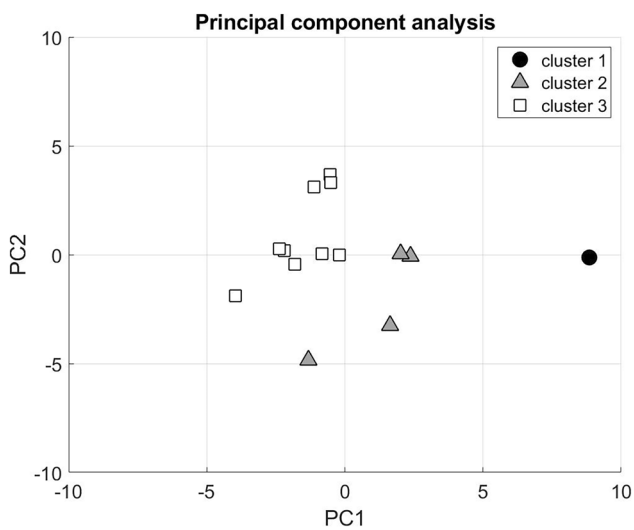

Fig. 3 Internal validation. Data representation in the space of the first two components of principal component analysis (PC1 and PC2) clearly showed three clusters

Table 2 Results for external validation

	Group_1 (LW10-LW10.5)	Group_2 (LW11)	Group_3 (LW11.5-LW12)	Precision (%)
Cluster_1 (high impact of impairment)	1	0	0	100
Cluster_2 (middle impact of impairment)	0	2	2	50
Cluster_3 (low impact of impairment)	0	0	9	100
Sensitivity (%)	100	100	81.8	

Table 3 Results for measures relevance

Tests	Measures	Cluster_1 ($n=1$)	Cluster_2 ($n=4$)	Cluster_3 ($n=9$)	p value	Effect size
Maximal voluntary contraction	PFa_w (N)	554.0 ± 0	889.1 ± 337.7	1097.4 ± 287.6	0.14	0.53
	PFp_w (N)	606.7 ± 0	974.7 ± 399.0	1217.5 ± 311.8	0.14	0.53
	PFa_{wo} (N)	291.3 ± 0	325.7 ± 152.5	623.4 ± 140.3	**0.02**[b]	**0.77**
	PF_p (N)	105.7 ± 0	236.9 ± 181.7	355.9 ± 105.8	0.15	0.52
	PFa_{wo}/PFa_w	0.52 ± 0	0.37 ± 0.09	0.58 ± 0.07	**0.03**[b]	**0.72**
Balance control at 0.5 m/s^2	$Trunk_{rest\text{-}b}$ (°)	3.3 ± 0	1.1 ± 2.7	1.5 ± 2.3	0.69	0.23
	$Trunk_{rom\text{-}b}$ (°)	20.6 ± 0	10.4 ± 3.8	4.9 ± 1.6	**0.04**[b]	**0.69**
	TRT_b (ms)	902.0 ± 0	833.9 ± 31.7	835.6 ± 122.6	0.25	0.45
	$Trunk_{rest\text{-}f}$ (°)	1.5 ± 0	0.7 ± 3.7	1.8 ± 2.3	0.82	0.17
	$Trunk_{rom\text{-}f}$ (°)	10.4 ± 0	6.7 ± 1.1	4.0 ± 1.2	**0.02**[a,b]	**0.75**
	TRT_f (ms)	854.0 ± 0	785.2 ± 167.7	614.9 ± 238.3	0.31	0.41
Balance control at 1 m/s^2	$Trunk_{rest\text{-}b}$ (°)	4.4 ± 0	1.4 ± 1.5	2.5 ± 1.5	0.34	0.39
	$Trunk_{rom\text{-}b}$ (°)	27.4 ± 0	10.0 ± 2.8	5.9 ± 1.3	**0.02**[a,b]	**0.75**
	TRT_b (ms)	717.6 ± 0	562.4 ± 15.4	526.1 ± 48.7	0.08	0.60
	$Trunk_{rest\text{-}f}$ (°)	3.9 ± 0	0.2 ± 1.5	2.1 ± 1.8	0.16	0.50
	$Trunk_{rom\text{-}f}$ (°)	32.1 ± 0	8.3 ± 2.5	4.7 ± 1.2	**0.01**[a,b]	**0.81**
	TRT_f (ms)	627.5 ± 0	489.0 ± 73.9	414.6 ± 127.0	0.19	0.49
Balance control at 2.5 m/s^2	$Trunk_{rest\text{-}b}$ (°)	2.8 ± 0	− 1.5 ± 4.5	3.5 ± 2.8	0.14	0.53
	$Trunk_{rom\text{-}b}$ (°)	28.6 ± 0	12.0 ± 2.0	7.7 ± 1.2	**0.01**[a,b]	**0.81**
	TRT_b (ms)	475.6 ± 0	475.9 ± 177.4	284.2 ± 26.3	**0.02**[b]	**0.77**
	$Trunk_{rest\text{-}f}$ (°)	4.0 ± 0	− 1.8 ± 4.2	3.1 ± 2.7	0.07	0.62
	$Trunk_{rom\text{-}f}$ (°)	27.7 ± 0	10.8 ± 1.0	8.7 ± 2.8	0.16	0.51
	TRT_f (ms)	559.2 ± 0	286.7 ± 24.3	262.4 ± 28.6	0.12	0.55

For all the measures, results for the three clusters are reported as mean ± standard deviation

Kruskal–Wallis test with post hoc results and effect size are reported in the last two columns, respectively

PFa_w anterior pushing force in upright seated press with backrest, PFp_w posterior pushing force in upright seated press with backrest, PFa_{wo} anterior pushing force in upright seated press without backrest, PF_p mean pulling force between the left and right peaks in simulated poling, $Trunk_{rest}$ trunk at rest, $Trunk_{rom}$ trunk range of motion, *TRT* trunk reaction time, *b* backward perturbation direction, *f* forward perturbation direction

For Kruskal–Wallis post hoc: [a] significant difference between cluster_1 and cluster_3, [b] significant difference between cluster_2 and cluster_3

most to distinguish between cluster_2 and cluster_3 (effect size > 0.72). Concerning the balance control test, $Trunk_{rom}$ allowed discriminating cluster_3 from cluster_1 and from cluster_2 in both forward and backward stimuli given at 1 m/s^2 (effect size > 0.75). No trunk control measures discriminated between cluster_1 and cluster_2. For the four measures that contributed most to clustering participants, post hoc power analysis exceeded 80%.

Spearman correlation results showed that PFa_{wo} and $Trunk_{rom}$ had an inverse correlation for all balance control accelerations and for forward and backward stimuli directions (Table 4). $Trunk_{rom}$ in both stimuli directions also correlated with PFa_w and with PFa_{wo}/PFa_w in balance control at 1 m/s^2.

Table 4 Results for Spearman correlation between maximal voluntary contraction and balance control measures

Balance control measures	Maximal voluntary contraction measures				
	PFa_w	PFp_w	PFa_{wo}	PF_p	PFa_{wo}/PFa_w
0.5 m/s^2					
$Trunk_{rest\text{-}b}$	– 0.45	– 0.37	– 0.13	– 0.28	0.29
$Trunk_{rom\text{-}b}$	**– 0.66***	– 0.49	**– 0.70****	– 0.51	– 0.28
TRT_b	– 0.02	– 0.04	0.26	– 0.11	0.52
$Trunk_{rest\text{-}f}$	– 0.05	– 0.39	– 0.12	– 0.28	0.40
$Trunk_{rom\text{-}f}$	– 0.42	– 0.42	**– 0.54***	– 0.35	– 0.40
TRT_f	0.05	– 0.19	– 0.08	– 0.13	– 0.13
1 m/s^2					
$Trunk_{rest\text{-}b}$	– 0.43	– 0.40	0.02	– 0.33	0.49
$Trunk_{rom\text{-}b}$	**– 0.58***	– 0.49	**– 0.71****	– 0.42	**– 0.56***
TRT_b	– 0.32	– 0.24	– 0.48	– 0.37	– 0.31
$Trunk_{rest\text{-}f}$	– 0.35	– 0.34	0.11	– 0.34	0.48
$Trunk_{rom\text{-}f}$	**– 0.55***	**– 0.57***	**– 0.76****	**– 0.57***	**– 0.57***
TRT_f	0.05	– 0.06	– 0.02	0.03	– 0.06
2.5 m/s^2					
$Trunk_{rest\text{-}b}$	– 0.17	– 0.13	0.24	– 0.16	0.52
$Trunk_{rom\text{-}b}$	– 0.42	– 0.32	**– 0.74****	– 0.35	**– 0.64***
TRT_b	– 0.47	– 0.41	**– 0.68****	– 0.47	**– 0.63***
$Trunk_{rest\text{-}f}$	– 0.22	– 0.15	0.27	– 0.12	**0.70****
$Trunk_{rom\text{-}f}$	**– 0.57***	– 0.44	**– 0.53***	– 0.31	– 0.16
TRT_f	– 0.44	– 0.38	– 0.52	– 0.22	– 0.18

PFa_w anterior pushing force in upright seated press with backrest, PFp_w posterior pushing force in upright seated press with backrest, PFa_{wo} anterior pushing force in upright seated press without backrest, PF_p mean pulling force between the left peak and right peaks in simulated poling. $Trunk_{rest}$ trunk at rest, $Trunk_{rom}$ trunk range of motion, *TRT* trunk reaction time, *b* backward perturbation direction, *f* forward perturbation direction

Correlation significance: $^{**}p < 0.01$; $^{*}p < 0.05$

4 Discussion

Overall, ICC, standard error of measurement, and Wilcoxon test showed very good results for the repeatability of impairment measures. Accuracy, precision, sensitivity, and macro-averaging *F* score showed very good results for the cluster analysis validation. Results of Kruskal–Wallis test, statistical power, and Spearman correlation showed that two measures of strength and two measures of trunk control were relevant in grouping participants in three clusters according to their impairment.

4.1 Measures for the classification purposes

The Position Stand and following documents for the development of an evidence-based classification systems [2, 5, 27] state the importance of identifying measures of impairment that can be used for classification purposes. Measures of impairment should be specific for only one of the eligible impairments, reliable, representative for the maximum variance possible in sport performance, quantitative, and ratio scaled [5]. These requirements are discussed singularly here.

The present study evaluated two different measurements, each of which is specific for an eligible impairment (trunk strength and trunk control). The tests to assess trunk strength (upright seated press and simulated poling) are selective and allow the participants to exert maximum force production [13, 39]. In addition, isometric maximal voluntary contraction tests were selected to evaluate impairment of strength, because isometric tests have been successfully adopted for this purpose in wheelchair racing [6, 8]. Finally, maximal voluntary contraction tests have been recommended to assess muscles strength for the purpose of classification in Paralympic sports, as reported in [39].

Core stability has been defined as the ability to maintain or recovery the trunk equilibrium after a perturbation [40]. Therefore, unpredictable perturbations have been used in the literature to assess trunk stability and trunk control ability [9, 41]. Trunk stability can be maintained by a combination of passive stiffness (ligaments and bones) and active stiffness (muscles), with the latter playing a greater role [42]. Stability can be increased by co-contraction of trunk flexor and extensor muscles and by increasing intra-abdominal pressure [43]. Athletes with physical impairment, such as spinal cord injury, have reduced or absent ability to voluntarily activate core muscles depending on the level of physical impairment.

Previously, balance control tests have been used to evaluate trunk control in cross-country sit-skiers, while seated on their personal sledge [20]. In the current study, balance control tests were performed at three different accelerations to find the one that was comfortable for the participants, but high enough to be influenced by the impairment. During skiing, when sudden balance loss occurs, trunk accelerations are similar to the one used in this study. However, the trunk acceleration at the beginning of the poling phase is expected to be much lower. This condition was not taken into account in the balance control test.

The proposed measurements of impairment of trunk strength and trunk control seem also to be reliable due to the test–retest results between days 1 and 2 (absence of statistical differences, very high intraclass correlation coefficient, and low standard error of measurement). There were some intraclass correlation coefficient results that deserve further discussion. First, the maximal voluntary contraction tests. The ratio between the anterior pushing force without and with a backrest reached a lower intraclass correlation coefficient (0.71) compared to the other four measures of strength (> 0.93). This might be explained because the ratio is a calculated measure, whereas the other four measures were directly acquired. The variability of the two measures

used to calculate the ratio may reduce the intraclass correlation coefficient, which nevertheless is still a good value [34].

The other two results of intraclass correlation coefficient that require further discussion concern the balance test. The low intraclass correlation coefficient of trunk angle at rest in both perturbation directions and three accelerations might suggest that the variability within a day and between 2 days was similar. This finding could occur because evaluating trunk angle at rest before the stimulus onset did not highlight participants' behavior in response to the perturbations given during the tests. Indeed, it might be possible that within participant variability had greater contribution. Therefore, trunk angle at rest does not seem to be a suitable measure from a reliability perspective.

Moreover, it is interesting to notice that at the lowest acceleration (0.5 m/s^2), trunk reaction time had low intraclass correlation coefficient in both perturbation directions. This finding may suggest that the lowest acceleration could be a test condition easily managed and controlled by participants. Therefore, 0.5 m/s^2 seems to be the least suitable of the three candidate accelerations to measure trunk control.

The proposed measurements of impairment are representative of the greatest variance possible in sport performance [5]. Indeed, trunk strength and trunk control can be assessed with maximal voluntary contraction and balance control tests using one device. Finally, measures of impairment should be quantitative and ratio scaled. These requirements are also satisfied, since numeric (instead of qualitative or grade) scale measures were used to assess impairment level.

4.2 Measures to assess trunk impairment

The use of a cluster analysis is a transparent scientific process and if good results were found during its validation it could contribute in defining a class structure for classification in Paralympic sports [6]. This was also demonstrated in studies focused on cross-country sit skiing [20, 23]. In the present study, cluster analysis validation showed promising results for the proposed classification model. Three clusters were disclosed by the internal validation (Fig. 3). External validation results (accuracy, precision, sensitivity, and F score) showed an overall agreement between the three clusters (cluster analysis output) and the three groups (a priori information defined on participants' trunk control and balance ability [1]).

Although in cluster_1 there was only one participant (classified as LW10.5), cluster analysis results seem to be in line with the current classification system. Only 2 participants out of 14 were allocated in a cluster different than expected. They were expected to belong to cluster_3, since they were LW11.5 or LW12. However, one of these two has partial trunk control ability due to a spinal cord injury, which seems to be more coherent with the description of partial trunk control ability of cluster_2. Whether cluster analysis offers a similar number of classes as the current system of classification remains to be studied with more participants. Indeed, such a study would benefit from including more sit-skiers from lower classes (LW10 and LW10.5) and able-bodied skiers as upper benchmark.

A second important outcome of the cluster analysis approach was the identification of the measures that mostly discriminated participants' impairment (Table 3). Among the measures of trunk strength tests, neither anterior nor posterior force in the upright seated press with the backrest significantly contributed to discriminate between clusters. The posterior force was collected as a control that all forces were consistent. However, since posterior force was close to anterior force, it is advisable not to consider it further.

The anterior force in the upright seated press without backrest and the ratio between the two anterior forces (without/with backrest) had the greatest effect in clusters differentiation. Therefore, they are sufficient for classification purposes, so simulated poling test may be unnecessary. However, the ratio of the two anterior forces was the measure that showed the lowest reliability. In further studies, to reduce the variability and increase measure reliability, a possible solution could be to assess the mean value over a time interval instead of considering only the peaks. A suitable time interval would need to be determined.

Concerning the balance control test, trunk angle at rest could not be considered as a measure of trunk control because of the low reliability and the non-significant contribution in discriminating among clusters. Trunk reaction time, being significant only for backward perturbations at 2.5 m/s^2, does not seem to be a suitable measure of trunk control either. Trunk range of motion showed the greatest effect size for the three accelerations and for both forward and backward stimuli, with the only exception in forward stimuli at 2.5 m/s^2. Comparing the three accelerations, $\pm$1 m/s^2 was the one with the greatest discrimination among the three clusters, whereas the other two ($\pm$0.5 and $\pm$2.5 m/s^2) distinguish only between cluster_2 and cluster_3.

These findings are also in line with correlation results (Table 4) that highlight a very high correlations between the two trunk control measures and the two trunk strength measures. The trunk range of motion in forward and backward perturbations at 1 m/s^2 correlated with the anterior force without backrest and with the ratio of the two anterior forces without/with back support, collected in strength tests. Therefore, these four measures can be identified as the most descriptive of the participants' characteristics. However, it is hard to compare this result with previous work that used unpredictable perturbations for classification [20]. In that

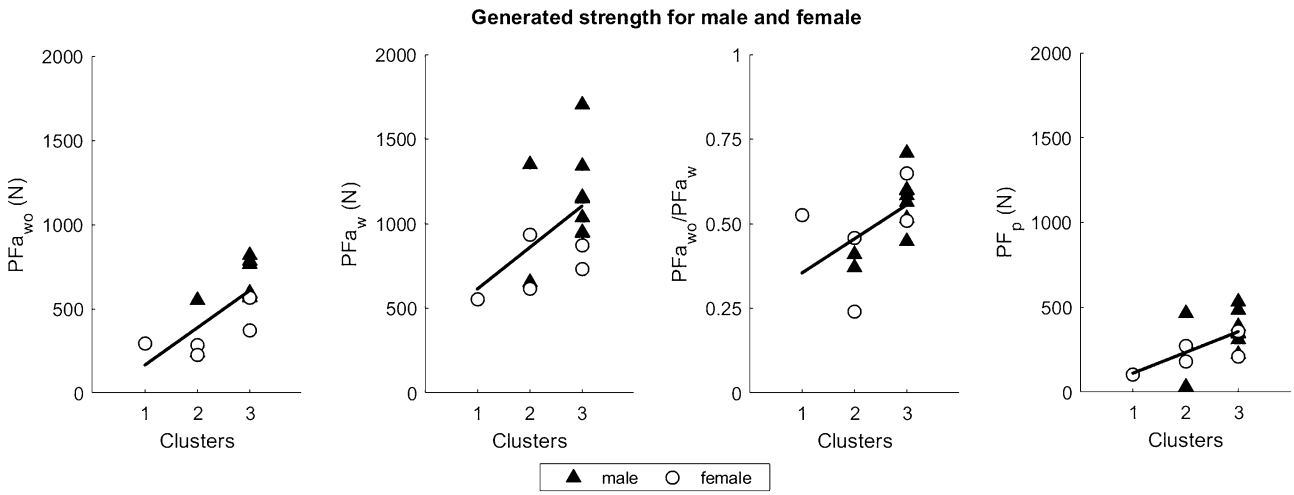

Fig. 4 Generated strength for male (black triangle) and female (white circle) participants in the three clusters for the maximal voluntary contraction tests

study, only ± 2.5 m/s^2 acceleration was used, two rather than three clusters were found, and participants were tested on their own sit-ski instead of a standard seat.

Overall, the results indicate that a shorter protocol and fewer measures than used here might be proposed as a measure of trunk strength and control impairment. However, to validate this finding, further analysis with more participants and considering possible gender effects is required.

4.3 Limitations

In this study, there are two main limitations that should be discussed. The first limitation is the small and skewed sample size, especially towards participants with low impairment. This kind of population may lead to difficulties in cluster analysis. However, the four measures that contributed most to cluster participants had power over 80%. Testing more sit-skiers would corroborate actual results about cluster_2 and cluster_3 and would allow generalizing results of cluster_1, considering that only one sit-skier was included in this cluster. Since worldwide the number of elite cross-country sit-skiers is low, athletes from other sitting sports, especially those with high trunk impairment, could be included.

The second main limitation of this study is related to the inclusion of both male and female participants, especially in maximal voluntary contraction tests. However, to limit gender effect on force, the ratio between anterior forces was calculated. Both genders were included in the study because strength generated in all maximal voluntary contraction tests have an overall increasing trend between clusters versus impairment reduction (from cluster_1 to cluster_3). This effect seems to be more evident with respect to gender effect (Fig. 4).

5 Conclusion

The development of an evidence-based classification in Paralympic sports has become mandatory. An evidence-based classification system is based on measures of impairment that can be used for classification purposes and sport-specific measures of performance. This study aimed to develop and evaluate measures of impairment of trunk strength and trunk control that can be used for classification purposes. Considering the current findings, it seems that performing trunk strength and balance control tests, and adopting a clustering approach, sit-skiers could be gathered in three clusters coherently with their trunk control and balance ability. However, a validation of the current findings including a larger sample size and considering effects of gender should be conducted. Since impairment of trunk is crucial in all sports, the importance of this study is that results found for cross-country sit skiing may also be extended to other sitting sports.

Acknowledgements The authors would thank the staff of the University of Jyväskylä for the technical support and all the athletes who volunteered in this study for participating.

Funding Open access funding provided by Politecnico di Torino within the CRUI-CARE Agreement. The authors would thank the Finnish Ministry of Education and Culture (OKM/51/626/2015) and IPC for approving this research and for financial support.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. International Paralympic Committee (2018) World Para Nordic Skiing—Classification. http://www.paralympic.org/nordic-skiing/rules-and-regulations/classification. Accessed 26 Nov 2018
2. Tweedy SM, Vanlandewijck YC (2011) International Paralympic Committee position stand-background and scientific principles of classification in Paralympic sport. Br J Sports Med 45:259–269
3. International Paralympic Committee (2017) World Para Nordic Skiing - Classification Rules and Regulations. https://www.paralympic.org/sites/default/files/document/170803114654801_World%252BPara%252BNordic%252BSkiing%252BClassification%252BRules%252Band%252BRegulations_0.pdf. Accessed 26 Nov 2018
4. International Paralympic Committee (2007) IPC classification code and international standards. http://www.paralympic.org/sites/default/files/document/120201084329386_2008_2_Classification_Code6.pdf. Accessed 26 Nov 2018
5. Tweedy SM, Mann D, Vanlandewijck YC (2016) Research needs for the development of evidence-based system of classification for physical, vision and intellectual impairments. In: Vanlandewijck Y, Thompson W (eds) Training and coaching the Paralympic athlete, 1st edn. Wiley, pp 122–149
6. Connick MJ, Beckman EM, Vanlandewijck YC et al (2017) Cluster analysis of novel isometric strength measures produces a valid and evidence-based classification structure for wheelchair track racing. Br J Sport Med 52:1123–1129
7. Rosén JS, Arndt A, Goosey-Tolfrey VL et al (2019) The impact of impairment on kinematic and kinetic variables in Va'a paddling: towards a sport-specific evidence-based classification system for Para Va'a. J Sports Sci 37:1942–1950
8. Vanlandewijck YC, Verellen J, Beckman EM et al (2011) Trunk strength effect on track wheelchair start: implications for classification. Med Sci Sports Exerc 43:2344–2351
9. Borghuis J, Hof AL, Lemmink KAPM (2008) The importance of sensory-motor control in providing core stability: implications for measurement and training. Sport Med 38:893–916
10. Vanlandewijck YC, Verellen J, Tweedy SM (2010) Towards evidence-based classification—the impact of impaired trunk strength on wheelchair propulsion. Postep Rehabil 3:1–5
11. Vanlandewijck YC, Verellen J, Tweedy SM (2011) Towards evidence-based classification in wheelchair sports: impact of seating position on wheelchair acceleration. J Sports Sci 29:1089–1096
12. van der Slikke RMA, Bregman DJJ, Berger MAM et al (2017) The future of classification in wheelchair sports: can data science and technological advancement offer an alternative point of view? Int J Sports Physiol Perform 13:742–749
13. Beckman EM, Newcombe P, Vanlandewijck YC et al (2014) Novel strength test battery to permit evidence-based Paralympic classification. Medicine (Baltimore) 93:e31
14. Altmann VC, van Limbeek J, Hart AL, Vanlandewijck YC (2014) Improvement of the classification system for wheelchair rugby: Athlete priorities. Adapt Phys Act Q 4:377–389
15. Chua JJC, Fuss FK, Subic A (2017) Activity identification and classification in wheelchair rugby using fractal dimensions. Sport Eng 20:1–15
16. Beckman EM, Tweedy SM (2009) Towards evidence-based classification in Paralympic athletics: evaluating the validity of activity limitation tests for use in classification of Paralympic running events. Br J Sports Med 43:1067–1072
17. Connick MJ, Beckman E, Deuble R, Tweedy SM (2016) Developing tests of impaired coordination for Paralympic classification: normative values and test–retest reliability. Sport Eng 19:147–154
18. Frossard L (2012) Performance dispersion for evidence-based classification of stationary throwers. Prosthet Orthot Int 12:348–355
19. Nicholson VP, Spathis JG, Hogarth LW et al (2018) Establishing the reliability of a novel battery of range of motion tests to enable evidence-based classification in Para Swimming. Phys Ther Sport 32:34–41
20. Rosso V, Gastaldi L, Rapp W et al (2019) Balance perturbations as a measurement tool for trunk impairment in cross-country sit skiing. Adapt Phys Act Q 36:61–76
21. Rosso V, Linnamo V, Rapp W, et al (2016) Trunk kinematics during cross country sit-skiing ergometry: skiing strategies associated to neuromusculoskeletal impairment. In: 2016 IEEE International Symposium on medical measurements and applications. IEEE, Benevento, pp 149–154
22. Rosso V, Linnamo V, Rapp W, et al (2018) Different sitting positions influence cross country sit skiers performance: sitting position influence on force generation and cycle characteristics. In: 2018 IEEE International Symposium on Medical Measurements & Applications. IEEE, Rome, pp 146–151
23. Rosso V, Linnamo V, Rapp W et al (2019) Simulated skiing as a measurement tool for performance in cross-country sit skiing. Proc Inst Mech Eng Part P J Sport Eng Technol 233:455–466
24. Blomqvist S, Van Biesen D, Vanlandewijck YC (2018) Gear selection between techniques in freestyle cross-country skiing in athletes with intellectual impairment: a pilot study. Int J Sports Sci Coach 13:1150–1155
25. Rosso V, Gastaldi L, Rapp W et al (2017) A new testing device for the role of the trunk in force production and in balance control in disabled sitting athletes. In: Ferraresi C, Quaglia G (eds) Proceedings of the 26th International Conference on Robotics in Alpe-Adria-Danube Region. Springer, Turin, pp 925–932
26. Tweedy SM, Beckman EM, Connick MJ (2014) Paralympic classification: conceptual basis, current methods, and research update. Pm R 6:S11–S17
27. Tweedy SM, Connick MJ, Beckman EM (2018) Applying scientific principles to enhance Paralympic classification now and in the future: a research primer for rehabilitation specialists. Phys Med Rehabil Clin N Am 29:313–332
28. Rosso V, Gastaldi L, Rapp W et al (2017) Biomechanics of simulated versus natural cross-country sit skiing. J Electromyogr Kinesiol 32:15–21
29. Fasel B, Spörri J, Schütz P et al (2017) Validation of functional calibration and strap-down joint drift correction for computing 3D joint angles of knee, hip, and trunk in alpine skiing. PLoS ONE. https://doi.org/10.1371/journal.pone.0181446

30. Bernardi M, Janssen T, Bortolan L et al (2013) Kinematics of cross-country sit skiing during a Paralympic race. J Electromyogr Kinesiol 23:94–101
31. Favre J, Aissaoui R, Jolles BM et al (2009) Functional calibration procedure for 3D knee joint angle description using inertial sensors. J Biomech 42:2330–2335
32. Favre J, Jolles BM, Siegrist O (2006) Animian K (2006) Quaternion-based fusion of gyroscopes and accelerometers to improve 3D angle measurement. Inst Eng Technol 42:612
33. Weir JP (2005) Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 19:231–240
34. Koo TK, Li MY (2016) A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15:155–163
35. Everitt BS, Landau S, Leese M, Stahl D (2011) Optimization clustering techniques. In: Shewhart WA, Wilks SS (eds) Cluster analysis, 5th ed. Wiley, Hoboken
36. Xu R, Wunsch DC (2008) Cluster validity. Clustering. Wiley-IEEE Press, pp 263–278
37. Beleites C, Salzer R, Sergo V (2013) Validation of soft classification models using partial class memberships: an extended concept of sensitivity & Co. applied to grading of astrocytoma tissues. Chemom Intell Lab Syst 122:12–22
38. Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Inf Process Manag 45:427–437
39. Beckman EM, Connick MJ, Tweedy SM (2017) Assessing muscle strength for the purpose of classification in Paralympic sport: a review and recommendations. J Sci Med Sport 20:391–396
40. Zazulak BT, Hewett TE, Reeves NP et al (2007) The effects of core proprioception on knee injury: a prospective biomechanical-epidemiological study. Am J Sports Med 35:368–373
41. Thigpen MT, Cauraugh J, Creel G et al (2009) Adaptation of postural responses during different standing perturbation conditions in individuals with incomplete spinal cord injury. Gait Posture 292:113–118
42. Akuthota V, Nadler SF (2004) Core strengthening. Arch Phys Med Rehabil 85:86–92
43. Arokoski JP, Valta T, Airaksinen O, Kankaanpaa M (2001) Back and abdominal muscle function during stabilization exercises. Arch Phys Med Rehabil 82:1089–1098

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.