The future of classification in wheelchair sports
Can data science and technological advancement offer an alternative point of view?

van der Slikke, Rienk; Bregman, Daan; Berger, Monique A.M.; De Witte, Annemarie M.H.; Veeger, DirkJan

DOI
10.1123/ijspp.2017-0326

Publication date
2017

Document Version
Final published version

Published in
International Journal of Sports Physiology and Performance

International Journal of Sports Physiology and Performance, 2018, 13, 742-749
https://doi.org/10.1123/ijspp.2017-0326

Human Kinetics
ORIGINAL INVESTIGATION

The Future of Classification in Wheelchair Sports: Can Data Science and Technological Advancement Offer an Alternative Point of View?

Rienk M.A. van der Slikke, Daan J.J. Bregman, Monique A.M. Berger, Annemarie M.H. de Witte, and Dirk-Jan (H.) E.J. Veeger

Purpose: Classification is a defining factor for competition in wheelchair sports, but it is a delicate and time-consuming process with often questionable validity. New inertial sensor-based measurement methods applied in match play and field tests allow for more precise and objective estimates of the impairment effect on wheelchair-mobility performance. The aim of the present research was to evaluate whether these measures could offer an alternative point of view for classification. ***Methods:*** Six standard wheelchair-mobility performance outcomes of different classification groups were measured in match play (n = 29), as well as best possible performance in a field test (n = 47). ***Results:*** In match results, a clear relationship between classification and performance level is shown, with increased performance outcomes in each adjacent higher-classification group. Three outcomes differed significantly between the low- and mid-classified groups, and 1, between the mid- and high-classified groups. In best performance (field test), there was a split between the low- and mid-classified groups (5 out of 6 outcomes differed significantly) but hardly any difference between the mid- and high-classified groups. This observed split was confirmed by cluster analysis, revealing the existence of only 2 performance-based clusters. ***Conclusions:*** The use of inertial sensor technology to obtain objective measures of wheelchair-mobility performance, combined with a standardized field test, produced alternative views for evidence-based classification. The results of this approach provide arguments for a reduced number of classes in wheelchair basketball. Future use of inertial sensors in match play and field testing could enhance evaluation of classification guidelines, as well as individual athlete performance.

Keywords: Paralympic sports, wheelchair basketball, inertial sensors, big data

In most Paralympic sports, a classification system is used to attain fair competition among athletes with various levels of impairment. The Paralympic classification systems aim to promote sports participation of people with disabilities by minimizing the impact of eligible types of impairment on competition outcome.[1] Ideally, the classification should only cover the effect of impairment on game performance. However, the magnitude of that effect is hard to estimate accurately given the number of confounding factors.[2] To determine the level of impairment itself, most classification systems categorize based on function levels rather than pathology.[3] Functional assessment is either based on isolated function tests, with assumptions about their effect on game performance, or the classification system is based on match observation. Given the diversity of functions, it is nearly impossible to determine the effect of each impairment level on game performance. The latter argument pledges to use match observation-based classification, but match-related confounders (field position, opponent, tactics) affect the functional assessment of those systems.

Wheelchair basketball was the first disability sport to use a functional classification system. Although functional classification is now a common practice, the wheelchair basketball system still stands out because the function-level assessment is based on match observation of "volume of action," instead of isolated function tests. The wheelchair basketball classification system (International Wheelchair Basketball Federation; www.iwbf.org) started out as a medical-based system (3 classes), but with the conversion to a function-based system, the number of classes was extended to 8 to take the increasing heterogeneity of participants into account. Classifications range from 1 point (most impaired) to 4.5 points (no functional limitation), with a team of 5 athletes composed of a maximum of 14 points. Although used since 1982,[4] there is an ongoing quest to provide scientific knowledge for more evidence-based classification guidelines.[2,5,6] The advantage of a match observation-based classification is that the assessments are made in an ecologically valid way, but observation methods also have their flaws and limitations. Actions like ball handling are well observed, but estimations of speed, acceleration, and force cannot be assessed accurately on observation alone. Another contaminating factor in the current observations is that match-specific factors like field position (guard, forward, center), opponent, and coach instructions are known to impact performance.[2] Indeed, more impaired players (low classification) are often positioned in physically less demanding field positions, possibly masking their potential best performance levels. Therefore, assessment of performance in a match alone provides a narrowed image, possibly disregarding best possible performance levels. By contrast, testing best performance in an isolated field test or lab setting alone does not provide information on how well an athlete is able to make use of his performance capacities during the course of a match. Therefore,

van der Slikke, Berger, and de Witte are with Human Kinetic Technology, Faculty of Health, Nutrition and Sports, The Hague University of Applied Sciences, The Hague, the Netherlands. van der Slikke, Bregman, and Veeger are with the Dept of Biomechanical Engineering, Delft University of Technology, Delft, the Netherlands. de Witte and Veeger are also with the Dept of Human Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. van der Slikke (r.m.a.vanderslikke@hhs.nl.) is corresponding author.

research on the relationship between match and best conditions is needed to determine if measurements in only 1 condition are sufficient for well-founded classification.

Several researchers investigated the effect of impairment on performance as expressed in the current classification, both in match conditions as well as in a field test to measure best possible performance. Vanlandewijck et al[5] assessed the wheelchair basketball performance of differently classified players during a match based on the Comprehensive Basketball Grading System next to the physical fitness in a laboratory test. Based on their results, they considered a reduced number of classes viable. In a similar study by Vanlandewijck et al[2] based on the Comprehensive Basketball Grading System scores of match performance, the relationship between class and position in the field was determined to be one of the factors for the absence of significant performance differences between 2 adjacent classes. In a study by Molik et al,[7] a Wingate Anaerobic Test was used to assess indexes of upper-extremity anaerobic performance, which also led to the conclusion that a reduced number of classes was recommendable. So, in research a relationship between classification and different performance measures is acknowledged in various conditions. However, to identify the true effect of impairment on performance and to explore the relationship between match and best performance, a single outcome measure should be used in both conditions.

A recently introduced method based on inertial sensors allows for objective performance estimations in both match and best conditions in a reliable and unobstructive way.[8] This method quantifies the wheelchair mobility performance, that is, the ability to maneuver the wheelchair. This measure of the wheelchair–athlete combination is one of the most important performance aspects[9] contributing to overall game performance as described by Byrnes et al.[10] In elite wheelchair basketball, van der Slikke et al[11] confirmed the clear relationship between classification and wheelchair mobility performance, but so far only in match conditions, not yet in best conditions (field test). In this study, wheelchair basketball athletes were measured in a sport-specific wheelchair mobility performance field test[12] that was first tested for reliability. Once the reliability had been ascertained, 47 elite athletes of all classifications were tested for best wheelchair mobility performance in this field test to rule out possible match-related confounding factors on wheelchair mobility performance.

The present study explores the relationship between wheelchair mobility performance in both match and best conditions, and its interaction with classification. The current classification is then compared with clusters derived from wheelchair mobility performance analysis in best conditions to outline a suitable number of performance-based classes. Finally, we will evaluate whether such clustering may provide an alternative point of view to classification systems.

Methods

Subjects

Wheelchair mobility performance was measured in a match[11] for the first group of elite wheelchair basketball athletes (n = 29) and in a standardized field test for a second group of athletes (n = 47, Table 1). Some of the athletes (n = 12) were measured in both conditions, forming a third data set for analysis of the relationship between match and field test performance. For the purpose of reliability testing, 23 of the athletes performed the field test twice. Results of this test–retest analysis are described in Appendix II. This study was approved by the ethical committee of the Department of Human Movement Sciences: ECB-2014-2. All participants signed an informed consent after being informed of the aims and procedures of the experiment.

Methodology

Each athlete's own sports wheelchair was equipped with 3 inertial sensors (xIMU for match; X-IO technologies and Shimmer3 for field test, Shimmer Sensing; Figure 1), 1 on each rear wheel axis and 1 on the rear frame bar. The frame sensor was used for measuring forward acceleration as well as rotation of the frame in the horizontal plane (heading direction). The combined signals of wheel sensor acceleration and gyroscope were used to estimate wheel rotation, which in turn provided frame displacement given the wheel circumference.

Estimates of frame rotations in the horizontal plane were used to correct the wheel gyroscope signal for wheel camber angle, as described by Pansiot et al,[13] Fuss,[14] and van der Slikke et al.[8] Furthermore, a skid correction algorithm was applied to reduce the effect of single or concurrent wheel skidding.[15]

Table 1 The Distribution of Classification and Age (Years) per Competition-Level Group of Athletes Measured in the Field Test

			Classification						
Level	**Mean**	**SD**	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**4.0**	**4.5**
National male									
Class	3.3	1.2	2	1	1	1	2	7	4
Age	23.7	10.1							
International male									
Class	3.0	1.2	2	1	1	4	3	2	4
Age	26.4	7.8							
International female									
Class	2.8	1.2	1	2	1	2	3	1	2
Age	32.9	8.0							
Total			5	4	3	7	8	10	10
Group total			Low = 9			Mid = 18		High = 20	

Figure 1 — Measurement setup, with inertial sensors on wheels and frame, and measurements during a match (photograph by www.frankvanhollebeke.be).

Based on inertial sensor outcomes for each measurement, a wheelchair mobility performance plot was generated, showing the 6 key outcomes of wheelchair performance.[11] The outcomes are as follows: average speed; average best speed (of best 5 in a match and of best 2 in the field test); average acceleration in the first 2 m from standstill; average rotational speed during forward movement; average best rotational speed during a turn on the spot (of best 5 in a match and of best 2 in the field test); and average rotational acceleration.

Statistical Analysis

To test for classification effects on wheelchair mobility performance, athletes were split into 3 classification groups: low (1–1.5), mid (2–3), and high (4–4.5). These classification group boundaries were chosen in line with earlier research regarding wheelchair mobility performance. In the paper by van der Slikke et al,[11] they chose to separate class I (1–1.5) as a single group, given their distinct performance levels,[2,5] and to separate class IV (4–4.5) from the class II and III athletes, since they also show (to a lesser extent) distinct performance levels.[2,5] Visual inspection of the distribution and a subsequent Kolmogorov–Smirnov test were applied to test for normal distribution[16] of all 6 wheelchair mobility performance outcomes to verify for the use of parametric statistics. A 1-way analysis of variance was used to test for group differences in the 6 standard mobility performance outcomes. For both field test (n = 47) and match data (n = 29), post hoc Bonferroni tests were applied to identify between which groups significant differences occurred.[17] The magnitudes of the classification group differences in the field test were also expressed in the smallest detectable difference (SDD 95%) as determined by the test–retest reliability (Appendix II). For the 12 athletes measured in both field test and match, a Pearson correlation was calculated for all 6 outcomes of the wheelchair mobility performance, combined with a paired-sample *t* test to verify if there were structural differences.

TwoStep clustering analysis was applied[18–20] to the complete field test performance data set without the split in classification groups (Appendix III). The TwoStep method is an exploratory tool designed to reveal natural groupings within a data set that would otherwise not be apparent.[21] Given the small sample size, a log-likelihood distance measure was combined with the Schwartz's Bayesian criterion.[22] As the maximal number of clusters is arbitrary, it was set in alignment to the current classification system (n = 8).

Results

For the 29 athletes measured in match play, classification group averages are displayed in the standardized wheelchair mobility performance plot (Figure 2).[11] The plot range was slightly enlarged to allow display of the best wheelchair mobility performance outcomes per classification group of the 47 athletes measured in the field test (Figure 3).

The differences among wheelchair mobility performance outcomes in the field test are also expressed in a factor of the SDD 95% (Table 2). The lowest factors of SDD 95% appear between the mid and high classification groups (0–1.0), and the highest factors show between the low and high classification groups (1.3–6.5).

Classification groups showed significant ($P < .05$) differences in all 6 wheelchair mobility performance outcomes in the match and in 5 outcomes in the field test measurements (Table 3). Post hoc Bonferroni tests revealed that, in the match, 3 out of 6 outcomes differed significantly ($P < .05$) between the low- and mid-classified athletes, and only best forward speed differed between the mid- and high-classified groups (Table 3). For best performance as measured in the field test, 5 wheelchair mobility performance outcomes differed significantly between low- and mid-classified athletes, and no outcomes differed between mid- and high-classified athletes.

For the 12 athletes measured in both match and field test conditions, the Pearson correlations for all 6 wheelchair mobility performance outcomes are displayed in Table 4. Three outcomes were significantly ($P < .05$) higher in the field test compared with the match performance, and 2 outcomes were higher on average, but not significant. The average best speed was significantly lower in the test compared with the match performance.

The TwoStep analysis revealed 2 clusters from a model that was considered "good" based on the cluster quality (silhouette of cohesion and separation ≥0.5). Most important model predictors were all forward movement–based outcomes (factor 0.93–1), whereas the importance of rotational outcomes ranged from a factor of 0.35 to 0.51. If analyzed for class allocation (Table 5), the first cluster (A) shows clear agreement with the low-classified group, although

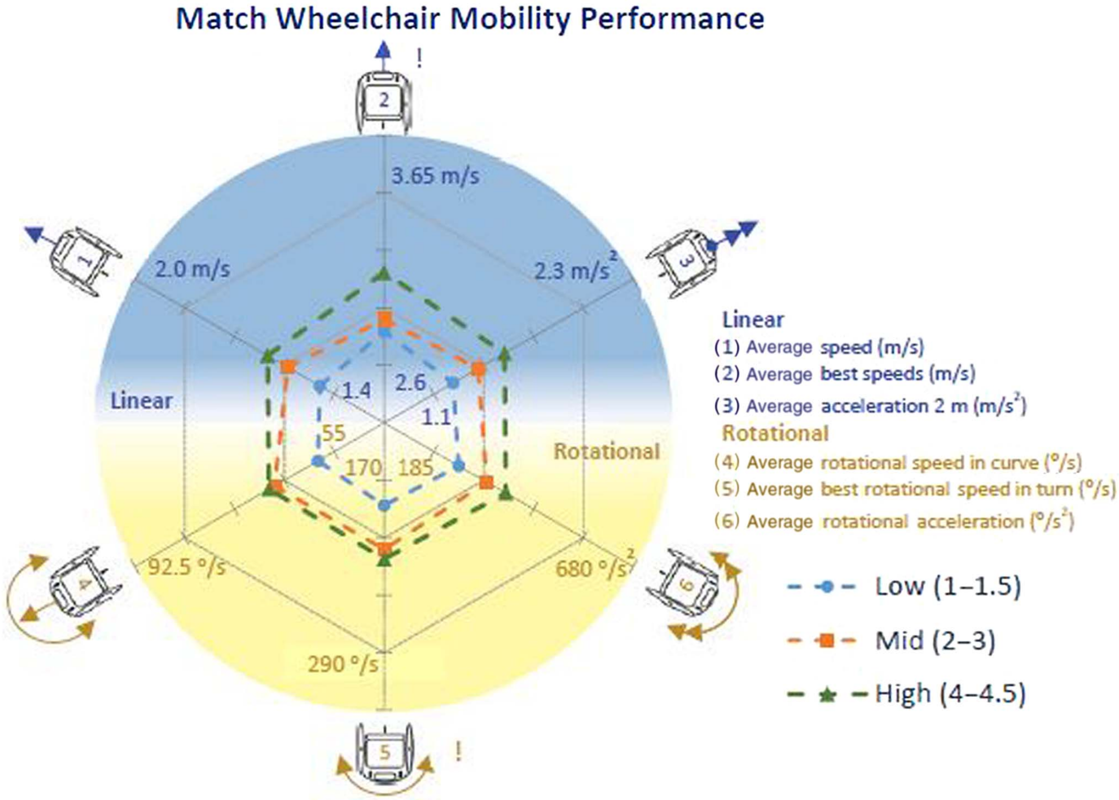

Figure 2 — Wheelchair-mobility performance plot for match performance of the 3 classification groups, adapted from van der Slikke et al.[11] The upper half of the plot shows the 3 kinematics regarding forward motion, and the lower half shows the rotational aspects. The low-classified (1–1.5) athletes performed below average on all 6 kinematic outcomes. The high-classified (4–4.5) athletes performed best on all outcomes.

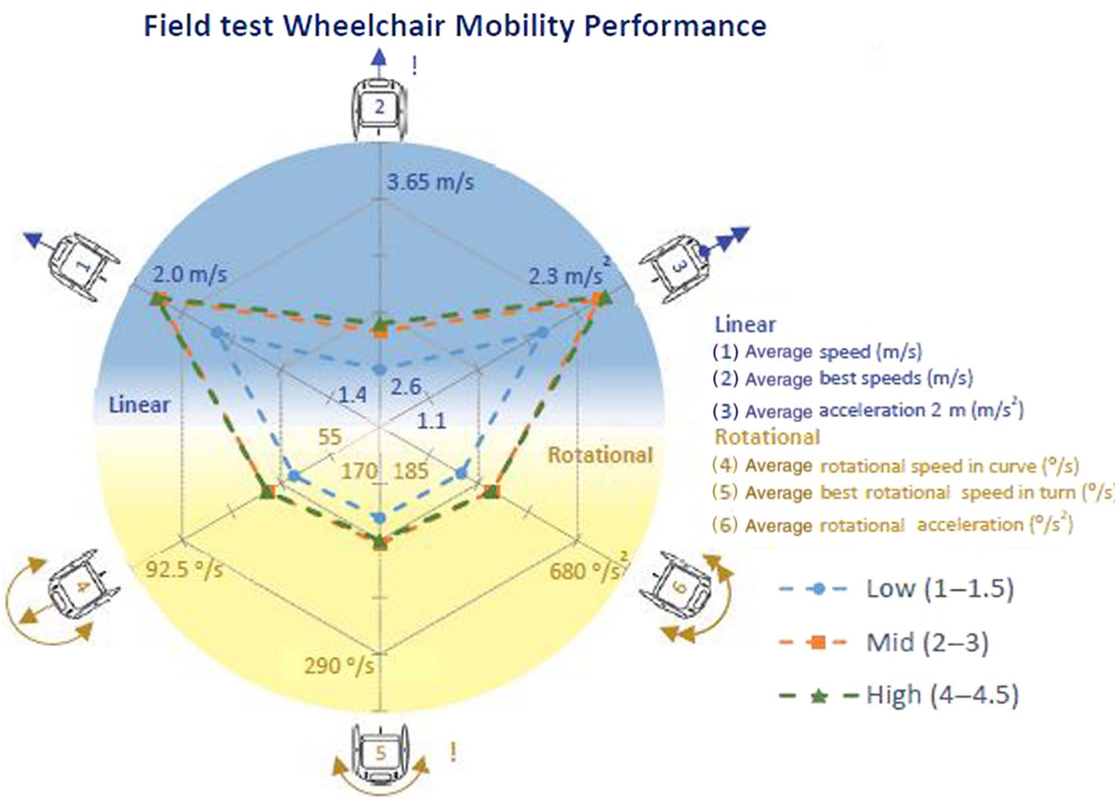

Figure 3 — Best possible wheelchair-mobility performance as measured in the field test for the 3 classification groups. The axis scaling of the wheelchair-mobility performance plot is based on match performance; it clearly shows that average field-test speed and acceleration (axes 1 and 3) exceed match performance. Furthermore, there is hardly any performance difference noticeable between the mid-classified (2–3) and high-classified (4–4.5) athletes.

Table 2 Classification-Group Differences in the Field Test Expressed as a Factor of the SDD (see Appendix I)

	SDD 95%	Low–mid	Low–high	Mid–high
Forward speed avg., m/s	0.038	**6.2**	**6.5**	0.3
Forward speed best, m/s	0.046	**5.2**	**6.2**	1.0
Forward acceleration avg., m/s^2	0.085	**5.3**	**6.0**	0.6
Rotational speed curve avg., °/s	3.409	**2.0**	**2.0**	0.0
Rotational speed turn best, °/s	12.065	**1.5**	**1.3**	0.2
Rotational acceleration avg., $°/s^2$	18.740	**5.5**	**5.5**	0.0

Abbreviations: avg., average; SDD, smallest detectable difference. Note: Factors of SDDs over 1 are marked bold.

Table 3 Classification-Group Statistics in the Match and Field Test Data

	Match Bonferroni post hoc				Field-test Bonferroni post hoc			
	ANOVA	Low–high	Low–mid	Mid–high	ANOVA	Low–high	Low–mid	Mid–high
Forward speed avg., m/s	**0.000**	**0.000**	**0.021**	0.214	**0.000**	**0.000**	**0.000**	1.000
Forward speed best, m/s	**0.000**	**0.000**	0.993	**0.003**	**0.000**	**0.000**	**0.003**	1.000
Forward acceleration avg., m/s^2	**0.001**	**0.001**	0.139	0.105	**0.003**	**0.003**	**0.010**	1.000
Rotational speed curve avg., °/s	**0.002**	**0.004**	**0.007**	1.000	**0.009**	**0.012**	**0.016**	1.000
Rotational speed turn best, °/s	**0.003**	**0.004**	**0.013**	1.000	0.068	0.146	0.078	1.000
Rotational acceleration avg., $°/s^2$	**0.006**	**0.005**	0.115	0.443	**0.002**	**0.003**	**0.004**	1.000

Abbreviations: ANOVA, analysis of variance; avg., average. Note: Significance levels are shown, with all levels $P < .05$ marked bold. Result description is based on adjacent class groups, that is, between low–mid and between mid–high. Differences between the low- and high-classified athletes are obvious and not used in further interpretation of results.

Table 4 Pearson Correlation and Mean Differences Between Match and Field-Test Performance (n = 12)

	Pearson correlation	Mean difference	*t*-test *P* value
Forward speed avg., m/s	**.735**	0.42	**.000**
Forward speed best, m/s	**.756**	–0.19	**.001**
Forward acceleration avg., m/s^2	**.702**	0.92	**.000**
Rotational speed curve avg., °/s	**.721**	1.70	.221
Rotational speed turn best, °/s	.616	0.60	.936
Rotational acceleration avg., $°/s^2$	**.745**	64.0	**.002**

Abbreviation: avg., average. Note: All Pearson correlations were significant ($P < .05$), >.7 marked bold; if match performance exceeded test outcomes, a negative value is shown in the mean difference; significance levels <.05 in the *t* test are marked bold.

6 athletes of the mid/high-classified groups are included as well. The second cluster (B) corresponds very well to the mid/high-classified groups, with only 1 athlete of the low-classified group included. The differences in performance outcomes between clusters, as expressed in the factor of SDD 95%, are quite similar to those shown between classification groups (low–mid and low–high, Table 2).

Discussion

This study was aimed at exploring the relationship between match and best wheelchair mobility performance and to what extend that relationship is affected by impairment level as expressed in the current classification. In general, it is clear that wheelchair mobility performance is clearly affected by the athlete's impairment level. This effect is shown in the match results, with increased performance outcomes for each successive classification group. Of the 6 wheelchair mobility performance outcomes, 3 differ significantly between the low- and mid-classified groups and 1 between the mid- and high-classified groups. Once the match-related factors are expelled, a different pattern emerges, as shown by the best results (field test measurements). Rather than a gradual incline of performance with classification (Figure 2), a clear performance separation is evident, with the most prominent difference between low- and mid-classified group outcomes. The wheelchair mobility plot (Figure 3) neatly shows that, in the field test, only the low-classified group deviates from the performance of the other athletes. Five of the wheelchair mobility performance outcomes differed significantly between these class groups, whereas no significant differences showed between mid- and high-classified athletes.

A relationship between classification and wheelchair mobility performance was anticipated in match and best conditions. Indeed, low-classified athletes show the lowest performance outcomes and high-classified athletes the highest wheelchair mobility performance values in both conditions, but the patterns of mid-classified athletes differ between conditions. Only moderate correlations between match and best performance were expected due to those differences in the mid-classified group. Moderate to high

Table 5 The TwoStep Clustering Method Applied to the Data Set of the 47 Athletes Measured in the Field Test Revealed 2 Clusters (A and B)

	Cluster			Factor	
Class	A	B	Mean difference	SDD 95%	*t*-test *P* value
Low	8	**1**			
Mid	**4**	14			
High	**2**	18			
Total	14	33			
Forward speed avg., m/s	1.87	2.13	0.26	6.83	.000
Forward speed best, m/s	2.60	2.90	0.30	6.51	.000
Forward acceleration avg., m/s^2	1.97	2.60	0.63	7.37	.000
Rotational speed curve avg., $°/s^2$	64.5	71.9	7.4	2.16	.000
Rotational speed turn best, $°/s^2$	193.9	213.9	20.0	1.66	.001
Rotational acceleration avg., $°/s^2$	307.3	404.7	97.4	5.20	.000

Abbreviations: avg., average; SDD, smallest detectable difference. Note: If optimized for group size (most athletes per class in each cluster), there is a clear split (dashed line) between the low- and mid- to high-classification groups. Bold indicates that the athletes were assigned to a cluster that did not meet their current classification. The lower part of the table shows the wheelchair-mobility performance outcomes per cluster and their difference, also expressed as a factor of the SDD 95% (Appendix I). The table shows the distribution of athlete classification over the 2 clusters, cluster performance characteristics, and their differences.

correlations (.62–.76) showed for the performance of the 12 athletes measured in both conditions. Given the unrestrained nature of the field test (no opponent or other obstructions), it was anticipated that wheelchair mobility outcomes would equal or exceed those of match conditions. Indeed, 3 out of 6 outcomes were significantly higher in the field test. Only average best speed appeared to score significantly lower in the field test. In the field test, the longest continuous run is 12 m, but in a match—although not frequent—longer continuous runs occur with corresponding higher speeds.

The impairment effect on performance should shape the classification system, so the International Paralympic Committee (IPC) is committed to the development of selective classification systems, not performance classification systems.[1] It is vital that athletes who improved their performance by training are not competitively disadvantaged by being placed into a less impaired class. Nevertheless, as performance level seems more dominated by impairment level than athlete training status or competition level,[11] performance clusters could be used to outline the number of classes needed in a particular system.

Once extracted from the match-specific confounders, field test wheelchair mobility performance data could be enforced to argue for a reduced number of classifications. Based on TwoStep clustering, only 2 performance clusters appeared. In clustering, outcomes related to forward speed and acceleration showed to be dominant factors. The 2 clusters show much similarity with the current classification of athletes, with only 1 athlete of the low-classified group assigned to cluster B. The remaining athletes of the low-classified group were assigned to cluster A, but this cluster also included 4 athletes of the mid-classified group and 2 of the high-classified group. In the population measured, athletes from both international and national competition level were included. The mid- and high-classified athletes assigned to cluster A were national males (n = 4) and international females (n = 2). In future research, a more homogenic group of athletes regarding competition level might slightly alter TwoStep cluster analysis outcomes.

Regarding wheelchair mobility performance, a single separation between the current class 1 to 1.5 athletes and the rest would be adequate. Subsequently, the 2+ class athletes could be divided into 2 groups, given the effect of their impairment, regarding ball handling. Such a reduced number of classes is in line with the conclusion of Vanlandewijck et al[5] and Molik et al,[7] pinpointing the viability of a reduction in the number of classes. A reduction in classes is also in line with the idea that the range of activity limitation within a class should also be as large as possible without disadvantaging those most severely impaired.[1] The wheelchair basketball–specific field test used is more closely related to match mobility performance than general performance measures (such as a physical fitness test or Wingate Anaerobic Test) frequently used in earlier research, so it provides more match-specific functional outcomes.

The aim of this study was to provide insight into the relationship between impairment and mobility performance in both best and match conditions, and to demonstrate the additional value of objective measures as provided by new technologies. Although the current classification system functions, with athletes and coaches generally satisfied,[23] there remains some controversy about the best approach to determine function level. The International Wheelchair Basketball Federation does not want to discard a reasonably well-functioning classification system based on years of gradual improvement, whereas the IPC seeks unity in systems over all sports, with selective classification based on “physical and technical assessment” off-court. Given that aspiration, the wheelchair mobility performance method used in this research seems unsuitable as a direct classification tool. Still, the need for sport-specific test batteries to aid the classifiers in objective decision making is emphasized by Tweedy and Vanlandewijck.[1] They state that current classification systems are still based on the judgment of a small number of experienced classifiers, rather than on empirical evidence, making the validity of the systems often questionable. In wheelchair basketball, the classification method is also time consuming and complicated. The use of objective measurement methods and sport-specific field tests can aid classifiers in their decision making. Results of the present study show the significance of on-court mobility performance measurements, whereas the ease of use of the inertial sensor-based method enables large-scale measurements in the future. By using the same method in both conditions, results of continued measurements in match play will approximate best performance (field test), reducing the effect of random factors typical to the observation of only a few matches, as in the current classification system. Indeed, it also brings to light whether athletes intentionally misrepresent their abilities in the classification tests, a major issue in Paralympic sports.

Practical Applications

The wheelchair basketball–specific field test used in this study[12] proved to be reliable combined with the inertial sensor-based method for measuring wheelchair mobility performance. In that sense, it complies with the IPC appeal to develop sport-specific test batteries for classification support. In addition to use for classification support, the field test is also a useful tool for individual athletes and coaches. Given the magnitude of the SDDs for all 6 outcomes, the field test is expected to be sensitive enough to detect performance changes as a result of training or interventions in wheelchair settings. Additional body-fixed inertial sensors could be used for more profound insight into the relationship between body movement ("volume of action") and wheelchair mobility performance.

Conclusions

Technological advancement, especially application of inertial sensors, allows for easy-to-use, large-scale, objective, and increasingly precise measurement of performance. Those benefits enable data science in adapted sports research that is traditionally characterized by small participant numbers. Such a big data approach with continued measurements in all conditions might offer an alternative point of view for classification outlining in Paralympic sports. Future research with additional body-fixed inertial sensors might provide more insight into the relationship between impairment and performance, bridging the gap to the selective classification envisioned by the IPC.

Acknowledgments

The authors would like to thank all athletes participating in this research project, all coaches involved in measurement planning, and Y. Vanlandewijck, K.van Breukelen, and F. van der Meulen for their contribution to this research and manuscript.

References

1. Tweedy SM, Vanlandewijck YC. International Paralympic Committee position stand—background and scientific principles of classification in Paralympic sport. *Br J Sports Med*. 2011;45(4): 259–269. PubMed ID: 19850575 doi:10.1136/bjsm.2009.065060
2. Vanlandewijck YC, Evaggelinou C, Daly DJ, et al. The relationship between functional potential and field performance in elite female wheelchair basketball players. *J Sports Sci*. 2004;22(7):668–675. PubMed ID: 15370498 doi:10.1080/02640410310001655750
3. Pickering Francis L. Competitive sports, disability, and problems of justice in sports. *J Philos Sport*. 2005;32(2):127–132. doi:10.1080/00948705.2005.9714677
4. DePauw KP, Gavron SJ. *Disability and Sport*. Champaign, IL: Human Kinetics; 1995.
5. Vanlandewijck YC, Spaepen AJ, Lysens RJ. Relationship between the level of physical impairment and sports performance in elite wheelchair basketball athletes. *Adapt Phys Activ Q*. 1995;12(2): 139–150. doi:10.1123/apaq.12.2.139
6. Altmann VC, Hart AL, Vanlandewijck YC, van Limbeek J, van Hooff ML. The impact of trunk impairment on performance of wheelchair activities with a focus on wheelchair court sports: a systematic review. *Sports Med Open*. 2015;1(1):22. doi:10.1186/s40798-015-0013-0
7. Molik B, Laskin JJ, Kosmol A, Marszałek J, Morgulec-Adamowicz N, Frick T. Relationships between anaerobic performance, field tests, and functional level of elite female wheelchair basketball athletes. *Hum Mov*. 2013;14(4):366–371. doi:10.2478/humo-2013-0045
8. van der Slikke R, Berger M, Bregman D, Lagerberg A, Veeger H. Opportunities for measuring wheelchair kinematics in match settings; reliability of a three inertial sensor configuration. *J Biomech*. 2015;48(12):3398–3405. PubMed ID: 26141162 doi:10.1016/j.jbiomech.2015.06.001
9. Mason BS, van der Woude LH, Goosey-Tolfrey VL. The ergonomics of wheelchair configuration for optimal performance in the wheelchair court sports. *Sports Med*. 2013;43(1):23–38. PubMed ID: 23315754 doi:10.1007/s40279-012-0005-x
10. Byrnes D, Hedrick B, Byrnes D, Shaver L. Comprehensive basketball grading system. In: Hedrick B, Byrnes D, Shaver L, eds. *Wheelchair Basketball*. 2nd ed. Washington, DC: Paralyzed Veterans of America; 1994:79.
11. van der Slikke C, Berger M, Bregman D, Veeger H. From big data to rich data: the key features of athlete wheelchair mobility performance. *J Biomech*. 2016;49(14):3340–3346.
12. de Witte AMH, Hoozemans MJM, Berger MAM, van der Slikke RMA, van der Woude LHV, Veeger DHEJ. Development, construct validity and test–retest reliability of a field-based wheelchair mobility performance test for wheelchair basketball. *J Sports Sci*. 2018;36: 23–32. doi:10.1080/02640414.2016.1276613
13. Pansiot J, Zhang Z, Lo B, Yang G. WISDOM: wheelchair inertial sensors for displacement and orientation monitoring. *Meas Sci Technol*. 2011;22(10):105801. doi:10.1088/0957-0233/22/10/105801
14. Fuss FK. Speed measurements in wheelchair sports—theory and application. *Sports Technol*. 2012;5(1–2):29–42. doi:10.1080/19346182.2012.754895
15. van der Slikke R, Berger M, Bregman D, Veeger H. Wheel skid correction is a prerequisite to reliably measure wheelchair sports kinematics based on inertial sensors. *Procedia Eng*. 2015;112:207–212. doi:10.1016/j.proeng.2015.07.201
16. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. *Int J Endocrinol Metab*. 2012;10(2):486–489. doi:10.5812/ijem.3505.
17. Field A. *Discovering Statistics Using IBM SPSS Statistics*. London: Sage; 2013.
18. Fraley C, Raftery AE. How many clusters? Which clustering method? Answers via model-based cluster analysis. *Comput J*. 1998;41(8): 578–588. doi:10.1093/comjnl/41.8.578
19. Bacher J, Wenzig K, Vogler M. *SPSS TwoStep Cluster—A First Evaluation*. Arbeits- und Diskussionspapiere. 2004-2. http://nbn-resolving.de/urn:nbn:de:0168-ssoar-327153
20. Mooi E, Sarstedt M. *Cluster Analysis*. Berlin Heidelberg: Springer; 2010.
21. Chiu T, Fang D, Chen J, Wang Y, Jeris C. A robust and scalable clustering algorithm for mixed type attributes in large database environment. Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '01). New York, NY; 2001:263–268. doi:10.1145/502512.502549
22. Schuetz T. A concise guide to market research: the process, data and methods using IBM SPSS Statistics. *Int J Mark Res*. 2011;53(4):563–564.
23. Molik B, Laskin JJ, Golbeck AL, et al. The International Wheelchair Basketball Federation's classification system: the participants' perspective. *Kinesiol Int J Fundam Appl Kinesiol*. 2017;49(1):21–22.
24. Bloxham L, Bell G, Bhambhani Y, Steadward R. Time motion analysis and physiological profile of Canadian World Cup wheelchair basketball players. *Sports Med Train Rehabil*. 2001;10(3):183–198. doi:10.1080/10578310210398

Appendix I

The athlete's performance can be divided into physical performance, mobility performance, and game performance. Physical performance only concerns the athlete,[24] whereas mobility performance is the measure for the combined wheelchair–athlete combination.[9] Therefore, although mobility performance is established by athlete exertion, it is often expressed in terms of wheelchair kinematics. van der Slikke et al[11] used a set of 3 inertial sensors to measure the wheelchair kinematics of 29 athletes in wheelchair basketball match play. To reduce the vast number of kinematic outcomes that could be measured with this configuration, principal component analysis was used to extract a set of 6 key features describing wheelchair mobility performance characteristics. Three of these outcomes describe forward motion and 3 describe the rotational aspect (maneuverability). All outcomes are plotted in a radar plot, with a scale relative to the group average and standard deviation.

Appendix II

Reproducibility of wheelchair mobility performance outcomes in the field test was tested by measuring 23 male athletes twice.[12] Retests were performed 1 week after the initial field test, under the same conditions (same timeframe, day of the week, and location). For each of the 6 performance outcomes, the intraclass correlation coefficient for consistency (ICC_c) between test and retest was calculated (Table AII1). Based on the ICC_c value and standard deviation, the standard error of mean for consistency (SEM_c) and the smallest detectable difference (SDD 95%) were calculated using

$$SEM_c = SD \times \sqrt{(1 - ICC_c)}$$

$$SDD\,95\% = SEM_c \times \sqrt{2} \times 1.96$$

The SDD 95% for each of the 6 performance outcomes is used to describe the differences between average performance of classification groups. For each outcome, the difference is divided by the SDD 95%, resulting in a dimensionless factor.

Table AII1 ICC, SEM, and SDD 95% of Wheelchair-Mobility Performance Outcomes Measured Twice in the Standardized Field Test

	ICC	SD	SEM	SDD 95%
Forward speed avg., m/s	**.947**	0.059	0.014	0.038
Forward speed best, m/s	**.947**	0.072	0.016	0.046
Forward acceleration avg., m/s^2	**.950**	0.138	0.031	0.085
Rotational speed curve avg., °/s	.870	3.41	1.23	3.41
Rotational speed turn best, °/s	.837	10.78	4.35	12.07
Rotational acceleration avg., $°/s^2$	**.944**	28.57	6.76	18.74

Note: All ICCs significant ($P < .001$), with ICCs over .9 marked bold. Abbreviations: ICC, intraclass correlation coefficient; SDD, smallest detectable difference; SEM, standard error of the mean.

Appendix III

The TwoStep cluster analysis is an exploratory tool designed to reveal natural groupings (or clusters) within a data set that would otherwise not be apparent. It has several unique features that make it very versatile. The most important feature for application in this study is the fact that it is capable of automatic selection of the number of natural clusters.

The 2 steps can be summarized as follows. Step 1: The procedure begins with the construction of a Cluster Features tree. The first case is placed at the root of the tree in a leaf node that contains variable information about that case. Each successive case is then added to an existing node or forms a new node based upon its similarity to existing nodes and using the distance measure as the similarity criterion. A node that contains multiple cases contains a summary of variable information about those cases. Thus, the Cluster Features tree provides a capsule summary of the data file. Step 2: The leaf nodes of the Cluster Features tree are then grouped using an agglomerative clustering algorithm. The agglomerative clustering can be used to produce a range of solutions. To determine which number of clusters is "best," each of these cluster solutions is compared using the Schwarz's Bayesian criterion.

In this study, for each of the 47 athletes, 6 wheelchair mobility performance outcomes are included in the data set for clustering. The TwoStep clustering procedure reveals the number of natural clusters and the assignment of each athlete to a cluster. To quantify the "goodness" of a cluster solution, the silhouette coefficient is used. This coefficient indicates how well the elements within a cluster are similar to one (cohesive) while the clusters themselves are different (separated). The TwoStep analysis also indicates which of the data (6 wheelchair mobility performance outcomes) were of most importance for clustering. The factor for importance to the model prediction can range from 0 (unimportant) to 1 (most important). This information helps to gain insight into the bases for the clustering model and the contribution of each performance outcome.