

Markus Tschopp, Pascal Bourban, Klaus Hübner, Bernard Marti
Sportwissenschaftliches Institut, Bundesamt für Sport, Magglingen

Messgenauigkeit eines 4-teiligen, standardisierten dynamischen Rumpfkrafttests: Erfahrungen mit gesunden männlichen Spitzensportlern

Zusammenfassung

Zur Beurteilung der Reliabilität eines einfachen, mit Hilfe von leicht transportablen Geräten standardisierten Krafttests für die ventrale, laterale und dorsale Rumpfkette sowie für die isolierte Bauchmuskulatur wurden 30 Spitzensportler (Alter: 20.2 ± 0.8 Jahre) aus 16 verschiedenen Sportarten zweimal im Abstand von 48 Stunden getestet.

Beim Test der isolierten Bauchmuskulatur nahm der Mittelwert der Zeit bis zur Erschöpfung beim 2. Test um etwa 30% zu. Die Mittelwerte der drei anderen Tests unterschieden sich um maximal 3,5%. Der Variationskoeffizient für die ventrale Rumpfkette betrug 14,1%, für die laterale 14,6% und für die dorsale Rumpfkette 11,7%. Zur Beurteilung einer Mindestkraftanforderung bei gesunden männlichen Spitzensportlern scheint mit Ausnahme des Tests der isolierten Bauchmuskulatur die vorliegende Messgenauigkeit genügend hoch zu sein. Allerdings muss diese intra-individuelle Variation bei der Beurteilung der Testresultate berücksichtigt werden.

Summary

Reliability of a standardized, dynamic trunk muscle strength test: experiences with healthy male elite athletes

The aim of this investigation was to evaluate the reliability of the test performance of trunk muscle strength and endurance. In a sample of 30 healthy male elite athletes (age: 20.2 ± 0.8 years), 4 tests of the ventral, lateral and dorsal trunk muscle chains and of the abdominal muscles were performed on 2 days within 48 hours. The meantime until exhaustion of the abdominal muscle test increased (30%) highly significantly ($p < 0.01$) whereas it differed maximally 3.5% for the 3 other tests. Typical error was 14.1% for the test of the ventral chain, 14.6% and 11.7% for the test of the lateral and the dorsal muscle chain respectively. It was concluded that the tests of the ventral, lateral and dorsal muscle chain can be used to assess trunk muscle strength and endurance for healthy male elite athletes. Yet the within-subject variation should be taken into account when interpreting results or changes in results.

Schweizerische Zeitschrift für «Sportmedizin und Sporttraumatologie» 49 (2), 67–72, 2001

Eine gut ausgebildete Rumpfmuskulatur ist für Spitzensportler unterschiedlicher Sportarten nicht nur aus präventiver Sicht von Bedeutung [1]. Eine regelmässige Beurteilung der Rumpfmuskulatur in der routinemässigen sportmedizinischen/leistungsdiagnostischen Abklärung liegt daher auf der Hand. Es bestehen unterschiedliche Messmethoden zur Evaluation der Rumpfmuskulatur [2–7]. Apparative Messverfahren liefern zwar meist reliable Messdaten. Für einen verbreiteten Einsatz in der Praxis sind sie aber oft zu aufwendig und zu teuer. Zudem entsprechen isometrische oder isokinetische Bewegungsabläufe der apparativen Testformen nicht unbedingt der sportlichen funktionellen Bewegung. Nichtapparative Testformen, die prinzipiell die muskuläre Aktion gegen die Schwerkraft erfassen, haben den Vorteil, dass sie flexibel sowie kostengünstig sind und damit in einem breiteren Umfeld eingesetzt werden können. Neben statischen bestehen verschiedene dynamische – und damit eher den sportlichen Anforderungen gerecht werdende – nichtapparative Testformen im Rumpfbereich [2, 3, 4]. Die Reliabilität und Validität dieser zum Teil in der Schweiz entwickelten [2] und in der physiotherapeutischen Praxis [8] und im Fitnessbereich [9] angewendeten Tests der dynamischen Muskelleistung im Rumpfbereich ist untersucht [3, 4] und für einzelne Tests kritisiert worden [10].

Es bestand daher die Absicht, die Messgenauigkeit dieser bereits eingesetzten, einfachen dynamischen Testformen [2] durch flexible, kostengünstige Standardisierungsgeräte sowie durch eine exakte Testbeschreibung und Instruktionen zu verbessern und hinsichtlich der praktischen Anwendung zu überprüfen. Aus der Überlegung, dass im Sport Bewegungen des Rumpfes meist ganze Muskelketten umfassen, betrafen 3 der 4 ausgewählten Testfor-

men ganze Muskelketten (ventrale, laterale und dorsale). Zusätzlich wurde eine dynamische Testform der isolierten Bauchmuskulatur ausgewählt. Für die isolierte Rückenstreckmuskulatur konnte wegen der Schwierigkeiten, eine normierte Ausgangsstellung und eine Kontrolle des Flexions-/Extensionsgrades zu finden, keine geeignete Testform entwickelt werden.

30 Spitzensportler des Lehrganges der Spitzensport-Rekrutenschule im Jahr 2000 in Magglingen absolvierten im Abstand von 48 Stunden zweimal diesen vierteiligen Rumpfkrafttest. Ziel der vorliegenden Arbeit ist eine Beurteilung der Messgenauigkeit der vier dynamischen Testformen zur Evaluation der Rumpfkraft. Zudem sollen daraus Folgerungen für den praktischen Umgang mit Messresultaten gezeigt werden.

Methode

30 männliche Spitzensportler (Alter: 20.2 ± 0.8 Jahre, Grösse: 180 ± 6 cm, Gewicht: 73.2 ± 8.0 kg) aus 16 Sportarten nahmen an der Untersuchung teil. Alle Athleten waren Absolventen der Spitzensportler-Rekrutenschule in Magglingen, wofür sie sich gemäss bestimmten Selektionskriterien qualifizieren mussten. Im Abstand von 48 Stunden führten sie zweimal morgens zur selben Zeit in der gleichen Reihenfolge 4 Testformen der Grundkraft des Rumpfes aus. 1 Monat vor dem 1. Test wurde die Einstellung an den 4 Posten vorgenommen, und die Testbewegungen mit 5 Wiederholungen geübt. Die 8 Tester (Physiotherapeuten, Sportlehrer und Ärzte) waren mit dem Testprozedere vertraut. Sie testeten am gleichen Gerät die gleichen Athleten. Die Resultate aus dem

1. Test waren weder dem Athleten noch dem Tester bekannt. Die Athleten hatten vor jedem Test die Anleitung bekommen, die Arbeit gegen die Schwerkraft solange wie möglich aufrechtzuerhalten. Vor dem Test wurden sie zur maximalen Ausbelastung motiviert, während des Tests wurde dagegen keine verbale Motivation durch den Testleiter vorgenommen. Der Athlet wurde auch nicht über die verstrichene Zeit informiert.

Das Aufwärmen war mit 10 Minuten Velofahren und anschliessend 5 Minuten Gymnastik standardisiert. 4 verschiedene Prototypen* von Standardisierungsgeräten (BASPO, Magglingen, siehe Abbildungen 1–6) wurden zur Normierung der Ausgangsstellungen oder zur Begrenzung der Bewegungen gebraucht. Ausgangsstellung, Bewegungsausführung, Instruktion am Proband/Athlet, und Abbruchkriterien (bei mangelnder Bewegungsqualität werden zwei Verwarnungen ausgesprochen, bei der dritten wird abgebrochen) waren definiert (siehe Beschreibung der einzelnen Testformen). Die Bewegungsgeschwindigkeit wurde durch ein akustisches Signal im Sekundentakt vorgegeben. Aufgezeichnet wurde die Dauer (in Sekunden) bis zum Abbruch. Zwischen den einzelnen Testelementen wurde eine 10-minütige Erholungszeit eingehalten. Die Tests wurden in der folgenden Reihenfolge absolviert:

1. Ventrale Rumpfkette (Abbildungen 1 und 2)

Aus der abgebildeten Ausgangsstellung (Unterarmstütze, Oberarme vertikal, Unterarme parallel, Daumen nach oben, Beine gestreckt, Mitte Schultergelenk, Trochanter major und äusserer Knöchel bilden eine Gerade, Scheitelpunkt in Kontakt mit dem Polster, spinae iliacae posterior superior in Kontakt mit der Querstange) wurde ein wechselseitiges Abheben der Füße um 2 bis 5 cm, Knie gestreckt, im Einsekundenrhythmus pro Fuss ausgeführt (modifiziert nach Spring et al. [2]).

Das Kopfpolster verhinderte während des Tests eine relative Extension im Schultergelenk von proximal aus. Die Brückenlänge blieb somit isometrisch. Das Standardisierungsgerät auf dem Becken verhinderte, dass die Körperlängsachse nach oben gebogen und somit die Brückenaktivität ökonomisiert wurde. Beim dritten Verlust des Kontaktes des Beckens mit der Querstange wurde der Test abgebrochen.



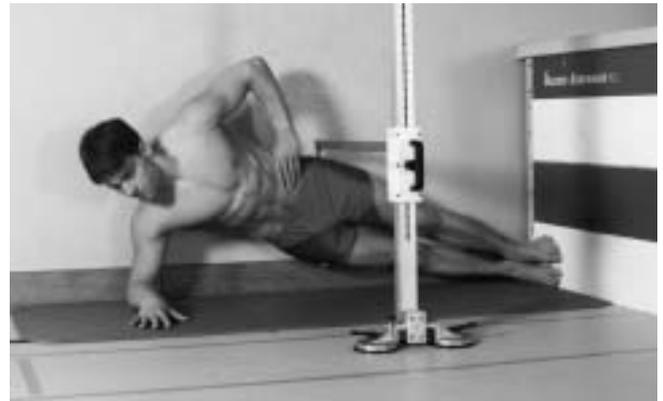
Abbildungen 1 und 2: Ventrale Rumpfkette.

* Bei genügend grosser Nachfrage besteht die Möglichkeit, die Standardisierungsgeräte am BASPO in einer kleinen Serie zu produzieren und für die Anwender unter Einbezug der Material- und Entwicklungskosten zur Verfügung zu stellen.

Auskunft: Bourban Pascal, Sportwissenschaftliches Institut, Bundesamt für Sport, CH-2532 Magglingen, Tel. 032 327 63 33, Fax 032 327 64 05, E-Mail: pascal.bourban@baspo.admin.ch

2. Laterale Rumpfkette (Abbildungen 3 und 4)

Aus der abgebildeten Ausgangsstellung (Ellbogenstütze in der bevorzugten Seitenlage, Füße aufeinander und gegen den Kasten gestützt, Sprunggelenke in Nullstellung, Fersen, Gesäss und Schultergürtel an der Wand, Oberarm vertikal, Ellbogen etwas von der Wand entfernt, die freie Hand auf dem Beckenkamm abgestützt) wurde im 2-Sekunden-Rhythmus pro Bewegungszyklus das Becken seitwärts abgehoben, bis der Rumpf die Nullstellung erreichte. Ohne das Körpergewicht abzusetzen, wurde wieder in die Ausgangsstellung zurückgegangen (modifiziert nach Spring et al. [2]). Die Wand diente der Überprüfung der korrekten Körperstellung, und die Bewegung wurde nach oben durch das Standardisierungsgerät limitiert. Der Test wurde abgebrochen, wenn zum dritten Mal entweder der Kontakt zum Standardisierungsgerät fehlte oder das Körpergewicht am Boden abgesetzt wurde.



Abbildungen 3 und 4: Laterale Rumpfkette.

3. Dorsale Rumpfkette (Abbildungen 5 und 6)

Aus der abgebildeten Ausgangsstellung (Füsse in der Sprossenwand, gepolsterte Sprosse auf Achillessehnen, Beine horizontal, Spinae iliacae anterior superior auf dem Schwedenkasten 2 cm von der Kante entfernt, Arme verschränkt, Finger liegen auf dem lateralen Drittel der Clavicula) wurde der Rumpf in 2 Sekunden pro Bewegungszyklus 30° nach unten und wieder zur Horizontale geführt (modifiziert nach Alaranta [3] und Moreland [4]). Die Endpositionen wurden mit einem auf einem Verlängerungsstab platzierten Schwerkraftgoniometer (Plurimeter, Dr. Rippstein, Switzerland) genau bestimmt. Zur Messung wurde der Verlängerungsstab auf den Dornfortsätzen der möglichst geraden Wirbelsäule angelegt. Das Standardisierungsgerät begrenzte die Bewegungen in der Sagittalebene nach unten bei 30° Flexion am angulus sterni und nach oben bei 0° (Körperlängsachse gerade) an dem vertikal liegenden Dornfortsatz der Brustwirbelsäule. Der Test wurde abgebrochen, wenn die obere Marke zum dritten Mal nicht mehr berührt wurde.



Abbildungen 5 und 6: Dorsale Rumpfkette.

4. Isolierte Bauchmuskulatur

Aus der Ausgangsstellung (Rückenlage, Gesäss am Kasten, 90° Hüftflexion, 90° Knieflexion, Unterschenkel auf einer Unterlage, Sprunggelenk in Nullstellung, Kopf liegt auf der Kopfstütze des Bauchmuskelttrainingsgerätes (Abdominal trainer, Fa. Cybex®), Arme gestreckt auf dem Bügel, Hände nebeneinander, Latte des

Standardisierungsgerätes vertikal über der Verbindungslinie der Spinae iliaca anterior superior) wurde dynamisch langsam der Oberkörper abgehoben und eingerollt, bis der Bügel die Latte berührte. Danach wurde in die Ausgangsstellung zurückgegangen, ohne dass der Kopf den Boden berührte. Pro Bewegungszyklus durften 2 Sekunden aufgewendet werden. Abgebrochen wurde, wenn der Bügel die Latte zum dritten Mal nicht mehr berührte.

Die statistische Analyse wurde mittels SPSS und Excel Software (Berechnungsformel: www.sportsci.org) durchgeführt. Die Mittelwerte wurden durch den gepaarten t-Test verglichen. Der Typical error wurde aus dem Quotienten der Standardabweichung der Differenzen aus Test 1 und Test 2 und der Quadratwurzel aus 2 berechnet [8]. Der Variationskoeffizient ist der Typical error dividiert durch den Mittelwert beider Tests. Als Korrelationskoeffizient wurde der Pearson-Koeffizient berechnet. Das Signifikanzniveau wurde bei $p = 0.05$ angesetzt.

Resultate

Tabelle 1 zeigt die Mittelwerte und die Standardabweichungen der 4 Grundkrafttests zum Zeitpunkt 1 und Zeitpunkt 2.

Auffallend ist die mittlere Zunahme von über 35% des Testresultates der isolierten Bauchmuskulatur zwischen dem ersten und zweiten Test, während dem für die anderen 3 Tests die Zubzw. Abnahme kleiner als 3,5% ist.

In den Diagrammen 1 bis 4 sind jeweils die Differenzen zwischen erstem und zweitem Testzeitpunkt (y-Achse) dem Mittelwert beider Testzeitpunkte (x-Achse) für jeden der 30 Probanden gegenübergestellt. Im Idealfall würde jeder Proband zweimal genau das gleiche Resultat erzielen, worauf die Datenpunkte in der gewählten Darstellung alle auf der x-Achse liegen würden. Je breiter die Punktwolke um die x-Achse gestreut ist, desto grösser sind die Differenzen der Ergebnisse der Probanden zwischen erstem und zweitem Test. Der Test der isolierten Bauchmuskulatur (Diagramm 4) zeigt die grösste Streuung, wogegen die Datenpunkte des Tests der dorsalen Rumpfkette (Diagramm 3) viel enger um die x-Achse verteilt sind. Auffallend sind im Diagramm 1 die beiden Extremwerte nach oben und unten, welche für 2 Probanden stehen, die zum zweiten Testzeitpunkt ein deutlich schlechteres bzw. besseres Resultat erzielten.

In Tabelle 2 sind die statistischen Kenngrössen zur Beurteilung der Reliabilität aufgeführt. Die Mittelwerte der Testergebnisse der isolierten Bauchmuskulatur zum Zeitpunkt 1 und 2 sind hochsignifikant verschieden (t-Test). Ebenfalls sind der Typical error und Variabilitätskoeffizient (CV) des Tests der isolierte Bauchmuskulatur deutlich grösser als diejenige der 3 anderen Testformen. Innerhalb dieser 3 Testformen hat der Test der dorsalen Rumpfkette den tiefsten Variabilitätskoeffizienten.

Testform	n	Test 1		Test 2	
		Mittelwert (sec)	SD (sec)	Mittelwert (sec)	SD (sec)
Ventrale Rumpfkette	30	175.3	63	174.3	69.7
Laterale Rumpfkette	30	87.2	29.1	89.5	26.6
Dorsale Rumpfkette	30	100.6	26.6	103.8	26.5
Isolierte Bauchmuskulatur	30	216.8	179.9	294.4	197.2

Tabelle 1: Mittelwert und Standardabweichung (SD) in Sekunden der 4 Testformen.

Testform	MW _{Diff(T1-2)}	SD _{Diff}	%SD _{Diff}	Typical error	CV %	p (t-Test)	r (Test1-2)
Ventrale Rumpfkette	1	35	20.0	24.7	14.1	0.88	0.87*
Laterale Rumpfkette	-2.3	18.2	20.6	12.9	14.6	0.5	0.81*
Dorsale Rumpfkette	-3.2	16.9	16.5	12	11.7	0.31	0.80*
Isolierte Bauchmuskulatur	-77.6	151.6	59.3	106.8	41.8	<0.01	0.68*

Tabelle 2: Statistische Kenngrössen der Reliabilität der 4 Testformen für Test 1 und Test 2 (Mittelwert MW und Standardabweichung SD der Differenzen in Sekunden und % Typical error, Variabilitätskoeffizient CV, p gepaarter t-Test, Pearson-Koeffizient r).

Testform	Referenzwert [1]	Wahrscheinlichkeit		
		66%	84%	95%
Ventrale Rumpfkette	101 sec	+/- 6 sec	+/- 15 sec	+/- 25 sec
Laterale Rumpfkette	50 sec	+/- 3 sec	+/- 7 sec	+/- 13 sec
Dorsale Rumpfkette	80 sec	+/- 4 sec	+/- 9 sec	+/- 17 sec

Tabelle 3: Wahrscheinlichkeit, dass tatsächlicher Wert grösser oder kleiner als Referenzwert [1] ist, bei einem erzielten Wert +/- sec (Sekunden).

Tabelle 3 zeigt die Wahrscheinlichkeiten, mit welcher das erzielte Resultat tatsächlich grösser oder kleiner ist als ein definierter Referenzwert [1]. Es zeigt sich, dass mit zunehmendem Abstand des erzielten Resultates vom Referenzwert diese Wahrscheinlichkeit zunimmt.

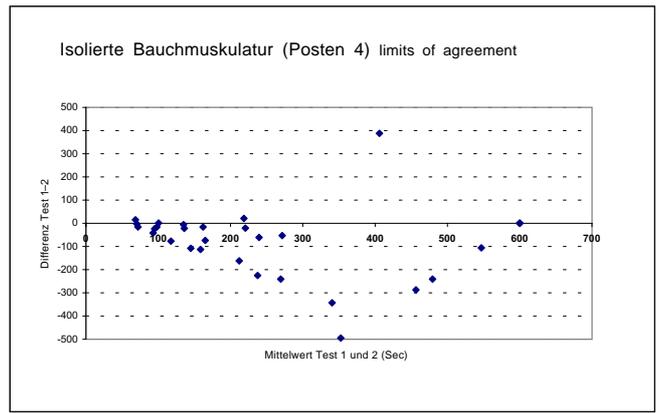


Diagramm 4: Mittelwerte (x-Achse) und Differenzen (y-Achse) aus Test 1 u. 2 der isolierten Bauchmuskulatur

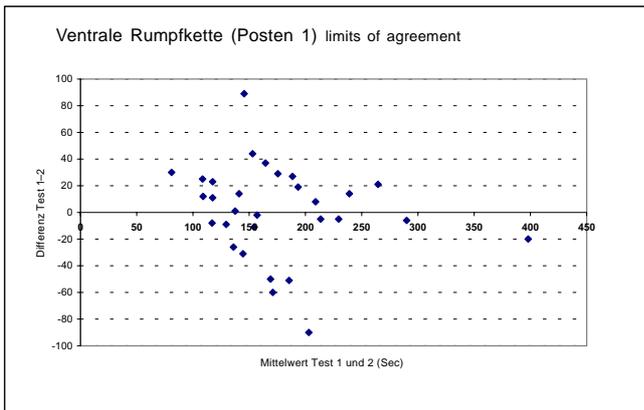


Diagramm 1: Mittelwerte (x-Achse) und Differenzen (y-Achse) aus Test 1 u. 2 der ventralen Rumpfkette

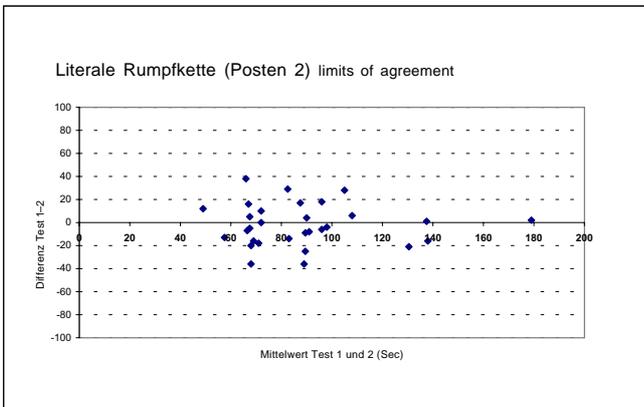


Diagramm 2: Mittelwerte (x-Achse) und Differenzen (y-Achse) aus Test 1 u. 2 der lateralen Rumpfkette

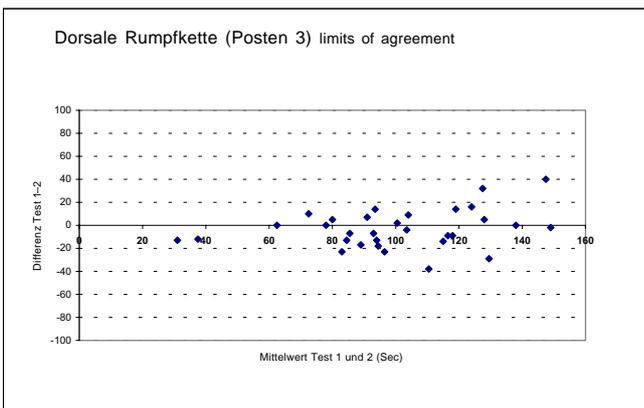


Diagramm 3: Mittelwerte (x-Achse) und Differenzen (y-Achse) aus Test 1 u. 2 der dorsalen Rumpfkette

Diskussion

Damit ein Test aussagekräftig ist und auch routinemässig angewendet werden kann, muss sichergestellt werden, dass die angewendete Messmethode genügend genau ist [11, 12]. Die Evaluation der Messgenauigkeit einer Testmethode beinhaltet im Wesentlichen zwei Komponenten: den systematischen Fehler (systematic bias) und den Zufallsfehler (random error) [11, 12].

Ersterer lässt sich aus dem Vergleich der Mittelwerte der Ergebnisse aus den verschiedenen Testzeitpunkten ableiten und gibt Rückschlüsse auf die Praxis der Testanwendung. Liegt ein signifikanter Unterschied zwischen den Mittelwerten der beiden Tests vor, muss davon ausgegangen werden, dass beispielsweise – bei einem tieferen Mittelwert am 2. Testzeitpunkt – ein Ermüdungseffekt oder – bei besserem 2. Testergebnis – ein Lerneffekt vorliegt. Möglicherweise ist auch die Motivation der Testathleten bei beiden Tests nicht gleich.

Wie ein Blick auf die Resultate zeigt, sind die Mittelwerte der beiden Tests der ventralen, lateralen und dorsalen Rumpfketten praktisch identisch, sodass oben genannte beeinflussende Gründe ausgeschlossen werden können. Die signifikante Zunahme der Ergebnisse des Tests der isolierten Bauchmuskulatur zwischen den beiden Testzeitpunkten deutet allerdings auf einen systematischen Fehler in der Testanordnung hin. Die gewählte Form der Messung der isolierten Bauchmuskulatur erlaubte den meisten Athleten eine Verbesserung im 2. Test, welche wahrscheinlich durch einen Lerneffekt aus dem 1. Test zustande kam. Das heisst, dass viele Athleten eine Bewegungsstrategie erlernten, die ihnen eine wechselnde Schonung und Belastung einzelner Muskelgruppen und somit eine deutlich höhere Wiederholungszahl ermöglichte. Möglicherweise verhalf auch ein Mobilisationseffekt der Hals- und Brustwirbelsäule während des 1. Tests zu besseren Resultaten in der Testwiederholung.

Die Messung der Kraft der isolierten Bauchmuskulatur ist mit der in dieser Studie angewendeten dynamischen Testform also nicht mit befriedigender Zuverlässigkeit möglich und erweist sich, wie auch andernorts [4, 13] gefunden wurde, als schwierig. Aufgrund der schlechten Test-Retest-Reliabilität ist der Test in dieser Form nicht valide und kann unseres Erachtens bei Spitzensportlern nicht zum praktischen Gebrauch empfohlen werden.

Die zweite Komponente der Variabilität der Testergebnisse – neben dem systematischen Fehler – ist der sogenannte Zufallsfehler (random error). Der Zufallsfehler beinhaltet die biologische und technische Variabilität. Die biologische Variabilität jedes einzelnen Probanden ist schwer zu kontrollieren und beinhaltet nach Abschluss grosser Änderungen in der Testvorbereitung zwischen den beiden Tests das, was man gemeinhin als «Tagesform» bezeichnet. Die technische Variabilität ist ein Ausdruck der Messmethodik, welche unter anderem durch die Testinstruktionen, Testgeräte oder Kontrolle der Bewegungsausführung gekennzeichnet wird.

Dieser Zufallsfehler wird durch den Variabilitätskoeffizienten ausgedrückt und liegt für alle drei Tests der ventralen, lateralen

und dorsalen Rumpfkette bei 12–15%. Ein Vergleich mit anderen Testmethoden [4–7, 10] in der Diagnostik der Rumpfmuskulatur ist wegen der unterschiedlichen Probandenkollektive (Spitzensportler, Untrainierte oder Patienten) kaum möglich. Ebenfalls sind die statistischen Kenngrößen vielfach nicht vergleichbar. Der in Reliabilitätsstudien oft verwendete Korrelationskoeffizient (r) lässt sich nur schwer zwischen verschiedenen Studien vergleichen, da er zu einem grossen Teil von der Heterogenität (Streubreite) der Resultate abhängt und somit zur Prüfung der Messgenauigkeit einer Methode ein eher irreführendes Verfahren darstellt [11, 12, 14]. Die Messgenauigkeit dürfte aber in etwa im Bereich derjenigen von anderen nicht apparativen [4, 5] und wie erwartet unter derjenigen von apparativen (isokinetischen) Rumpftests [6] liegen.

Es stellt sich nun die Frage, ob sich eine Messgenauigkeit mit einem Variabilitätskoeffizient von je nach Test 12–15% bei der Beurteilung der Testresultate als genügend gut erweist. Die Beurteilung der Messgenauigkeit kann nicht allein auf der reinen Höhe der statistischen Kriterien beruhen (Bsp. $r > 0.9$ oder $CV < 10\%$) [11, 12]. Vielmehr müssen die statistischen Kenngrößen der Reliabilitätsabklärung in den Zusammenhang mit einer bestimmten praktischen Fragestellung (analytic goal) gestellt werden [11].

Primäres Ziel eines Grundkrafttests ist wohl die Abklärung, ob ein Athlet einer Mindestkrafthanforderung genügt [1]. Eine Ausbildung der Rumpfmuskulatur darüber hinaus dürfte nicht mit einer proportionalen linearen Steigerung der sportlichen Leistung oder des präventiven Nutzens einhergehen [1]. Darin besteht ein wesentlicher Unterschied zur Diagnostik beispielsweise der Sprungkraft für einen Sprinter oder der Ausdauerleistungsfähigkeit für einen Ausdauerathleten. Während hier in der Bestimmung des Leistungsniveaus relativ geringfügig scheinende Unterschiede von einzelnen Zentimetern oder Stundenkilometern schlussendlich für das Wettkampfergebnis entscheidend sein können und deshalb durchaus sinnvoll zu bestimmen sind, ist ein solcher Anspruch an die Messgenauigkeit in der Bestimmung der Grundkraft des Rumpfes nicht erforderlich. Die Mindestkrafthanforderung wird nämlich – in Mangel eines in prospektiver Art ermittelten Zusammenhangs von Testresultaten und sportlichen Erfolgen oder Beschwerden – meist mehr oder weniger willkürlich festgelegt (Bsp. Mittelwert minus 1 Standardabweichung der Resultate eines Referenzkollektivs) [15]. Zudem dürfte ein solcher fixer Referenzwert, welcher in der Praxis bisweilen als (allzu) strenger Richtwert zwischen «genügend» und «ungenügend» diskriminiert, kaum eine scharfe Trennlinie von einer einzelnen Sekunde darstellen.

Was muss nun aufgrund dieser Relativierung bei der Messgenauigkeit der 3 vorliegend validierten Grundkrafttests in der praktischen Umsetzung berücksichtigt werden?

Erstens muss das vom Athleten erzielte Resultat in Bezug zur Messgenauigkeit gestellt werden. Je besser die Messgenauigkeit einer Methodik ist, desto exakter entspricht das erzielte Resultat dem tatsächlichen Leistungsvermögen. Bei der Beurteilung, ob ein Athlet eine Mindestkrafthanforderung erreicht oder nicht, muss dies berücksichtigt werden. Es ist klar, dass die Beurteilung für Athleten, welche ein sehr schlechtes bzw. sehr gutes Resultat erzielen, einfach und zuverlässig ist. Für Athleten jedoch, die ein Resultat nahe dem Referenzwert (Bsp. Mittelwert minus 1 Standardabweichung) erzielen, ist eine Beurteilung, ob das wirkliche Leistungsvermögen darüber oder darunter liegt, nicht ohne weiteres möglich. Bedingt durch die Messgenauigkeit der Testmethode ist eine Zuteilung bei einem Testresultat nahe dem Referenzwert rein zufällig. Eine Möglichkeit, um das wirkliche Niveau zu bestimmen, ergäbe sich aus einer mehrmaligen Testwiederholung, wobei das Ergebnis dem Mittelwert der gesamten Testreihe entsprechen würde. Jedoch ist ein solches Vorgehen in der alltäglichen Routine nicht durchführbar.

Es stellt sich also die Frage, wie viel besser oder schlechter ein Resultat als der Referenzwert sein muss, damit man die Leistung mit hoher Wahrscheinlichkeit richtig als «über» oder «unter dem Referenzwert liegend» bezeichnen kann.

Wie bereits erwähnt, steigt mit zunehmendem Abstand diese Wahrscheinlichkeit. Für den Test der ventralen Rumpfkette ist

ein Resultat, das 6 Sekunden über dem Referenzwert liegt, mit 66%iger Wahrscheinlichkeit, bei 15 Sekunden dagegen bereits mit 84%iger Wahrscheinlichkeit tatsächlich über dem Referenzwert. Hopkins [11] war in einem kürzlich erschienenen Artikel über die Berechnung der Reliabilität in der Sportmedizin der Meinung, dass das oft verwendete 95%ige Wahrscheinlichkeitsintervall für den praktischen Gebrauch bei Sportlern zu gross sei und dass bereits Intervalle mit Wahrscheinlichkeiten von $>84\%$ sinnvoll sind. Werden Resultate erzielt, welche innerhalb dieses Intervalls liegen, ist die Zuteilung in eine Kategorie nicht zuverlässig. Kann der Test in diesen Fällen aus praktischen Gründen nicht wiederholt werden, muss insbesondere die kritische Beurteilung der Leistung zurückhaltend erfolgen.

Da die absolute Grösse dieses Intervalls direkt von der Messgenauigkeit abhängt und für die Richtigkeit der Beurteilung entscheidend ist, muss unbedingt auf die exakte Durchführung der 3 Testformen in der beschriebenen Methode mit Hilfe der Standardisierungsgeräte geachtet werden.

Die Messgenauigkeit einer Testmethode ist auch bei der Beurteilung eines Kraftverlustes oder -zuwachses zu berücksichtigen. Ein Kraftzuwachs eventuell in Folge eines Krafttrainings kann erst ab einer entsprechenden Grösse mit Sicherheit als solcher bezeichnet werden, da kleinere Änderungen durch den Zufallsfehler bedingt sein können.

Nicht überprüft wurde die Intertester-Reliabilität. Durch die exakte Testbeschreibung, genaue Instruktion des Athleten [16] und durch die Standardisierungsgeräte wurde versucht, eine konstante Anwendung unabhängig vom Untersucher zu gewährleisten.

Bei der Anwendung dieser Testmethode muss berücksichtigt werden, dass es sich hier um eine homogene Population gesunder männlicher Spitzensportler handelt. Für die Anwendung bei einer anderen Population (Frauen, Kinder, Patienten) kann die hier aufgeführte Messgenauigkeit nicht ohne weiteres übernommen werden.

Zusammenfassend lässt sich sagen, dass mit Ausnahme des Tests der isolierten Bauchmuskulatur die beschriebenen Testmethoden der ventralen, lateralen und dorsalen Rumpfmuskulatur für die Evaluation einer Mindestkrafthanforderung bei gesunden männlichen Spitzensportlern angewendet werden können, dass jedoch bei der Durchführung die genaue Einhaltung der Testmethode (Standardisierungsgeräte, Instruktion) und bei der Beurteilung einmaliger Testergebnisse sowie deren Änderungen im Verlauf die Messgenauigkeit unbedingt berücksichtigt werden muss.

Literaturverzeichnis

- 1 Bourban P., Hübner K., Tschopp M., Marti B.: Grundkrafthanforderungen im Spitzensport: Ergebnisse eines 3-teiligen Rumpfkrafttests. Schweiz. Z. Sportmed. Sporttraumat., 49: 73–78, 2001.
- 2 Spring H., Kunz H.R., Schneider W., Tritschler T., Unold E.: Testung der Kraftausdauer; in «Kraft. Theorie und Praxis». Thieme Verlag, Stuttgart, New York: 26–47, 1990.
- 3 Alaranta H., Hurri H., Heliövaara M., Soukka A., Harju R.: Non-dynamometric trunk performance tests: reliability and normative data. Scand. J. Rehab. Med., 26: 211–215, 1994.
- 4 Moreland J., Finch E., Stratford P., Balsor B., Gill C.: Interrater Reliability of Six Tests of Trunk Muscle Function and Endurance. JOSP, 26: 200–208, 1997.
- 5 Oesch P., Kool J., Sloksnath U., Hasegawa T.: Die Zuverlässigkeit und Empfindlichkeit von vier isometrischen Muskeltests. Physiotherapie, 6: 4–12, 1994.
- 6 Delitto A., Rose S.J., Crandell C.E., Strube M.J.: Reliability of Isokinetic Measurements of Trunk Muscle Performance. Spine, 16: 800–803, 1991.
- 7 Moffroid M., Reid S., Henry S.H., Haugh L.D., Ricamato A.: Some Endurance Measures in Persons with Chronic Low Back Pain. JOSPT, 20: 81–87, 1994.
- 8 Bacq C., Braun M., Fleury M.P., Gillis G., Nuttin F.: Entraînement musculaire du patient lombalgique. Physiotherapie, Annexe, 10: 1–15, 1996.
- 9 Weber A.: Empirische Studie als Grundlage für einen praxisnahen Test der Kraft und Beweglichkeit im Gesundheitssport. Diplomarbeit, Eidgenössische Sportschule Magglingen: 19–20, 1999.

- 10 Gerber I., Lüthi S., Radlinger L.: Kraftausdauerstest der Rumpfmuskulatur (Globaltest) – Überprüfung der Hauptgütekriterien. *Manuelle Therapie*, 4: 3–9, 2000.
- 11 Hopkins W.G.: Measure of Reliability in Sports Medicine and Science. *Sports Med.*, 30: 1–15, 2000.
- 12 Atkinson G., Nevill A.M.: Statistical Methods For Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. *Sports Med.*, 26: 217–238, 1998.
- 13 Abernethy P., Wilson G.: Introduction to the Assessment of Strength and Power; in «Physiological Tests for Elite Athletes», Gore C. Human Kinetics, Champaign: 147–150, 2000.
- 14 Bland J.M., Altman D.G.: Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *The Lancet I*: 307–310, 1986.
- 15 Denner A.: Referenzdaten; in «Analyse und Training der wirbelsäulenstabilisierenden Muskulatur». Springer, Berlin, Heidelberg, New York: 89–114, 1998.
- 16 Matheson L., Mooney V., Caiozzo V., Jarvis G., Pottinger J., DeBerry C., Backlund K., Klein K., Antoni J.: Effect of Instructions on Isokinetic Trunk Strength Testing Variability, Reliability, Absolute Value, and Predictive Validity. *Spin* 17: 914–921, 1991.

Korrespondenzadresse:

Dr. med. Markus Tschopp, Sportwissenschaftliches Institut, Bundesamt für Sport, CH-2532 Magglingen, Tel. 032 327 61 22, E-Mail: markus.tschopp@baspo.admin.ch