

5. Ergebnisse

5.1. Radiologische Analyse

5.1.1. Zusammenfassung

Eine Beurteilung von seitlichen Funktionsröntgenbildern nach verschiedenen Verfahren führt zu unterschiedlichen Ergebnissen. Bis auf zwei Ausnahmen (Dupius-Translation und Morgan-Translation sowie Dupius-Angulation und Panjabi-Angulation) sind alle Einzelvergleiche zwischen den Messmethoden signifikant unterschiedlich (Chi-2-Test, $p < 0,01$). Die relativ hohe Streuung der Messwerte (mittlere Standardabweichung 3,93 mm bzw. $2,29^\circ$) führte bei einer erlaubten Abweichung von ± 1 mm bzw. $\pm 1^\circ$ zu Übereinstimmungen bei Mehrfachuntersuchungen von 38,7 % für angulatorische Verfahren und 49 % für translatorische Verfahren. Für die translatorische Beurteilung von Funktionsröntgenbildern sollte wegen der besten Übereinstimmung (54,4 %) das Verfahren nach PANJABI genutzt werden. Für die angulatorische Beurteilung von Funktionsröntgenbildern sollte das Verfahren nach PANJABI (41,1 % Übereinstimmung) oder nach DUPIUS (41,7 % Übereinstimmung) genutzt werden. Trotz relativ geringer Übereinstimmungen korrelierten alle Verfahren gut bis sehr gut ($r > 0,65$). Die Korrelationsanalyse sagt bei einem Vergleich zweier klinischer Methoden nichts über deren tatsächliche Übereinstimmung aus und ist daher für derartige Beurteilungen nicht geeignet.

5.1.2. Korrelationsanalyse

Bei der Korrelationsanalyse nach PEARSON zeigten sich für alle untersuchten Paarungen Korrelationswerte von 0,65 oder höher. Nach der Bewertungsskala von LADIS und KOCH bedeutet dies für alle untersuchten Zusammenhänge eine gute ($r=0,61...0,8$) bis sehr gute ($r=0,81...1,0$) Korrelation.

Sowohl für die intraindividuellen Paarungen [RK1-RK2 ($r=0,736...0,909$); TS1-TS2 ($r=0,793...0,890$); CH1-CH2 ($r=0,791...0,879$)] als auch für die interindividuellen Paarungen zum Untersuchungszeitpunkt 1 [RK1-TS1 ($r=0,767...0,911$); RK1-CH1 ($r=0,756...0,948$) und TS1-CH1 ($r=0,855...0,922$)] und zum Untersuchungszeitpunkt 2 [RK2-TS2 ($r=0,669...0,901$); RK2-CH2 ($r=0,693...0,939$) und TS2-CH2 ($r=0,754...0,912$)] ergaben sich Korrelationskoeffizienten von $r=0,668...0,949$. In Abb. 16 und 17 sind exemplarisch die graphischen Verläufe einer intraindividuellen und einer interindividuellen Analyse dargestellt.

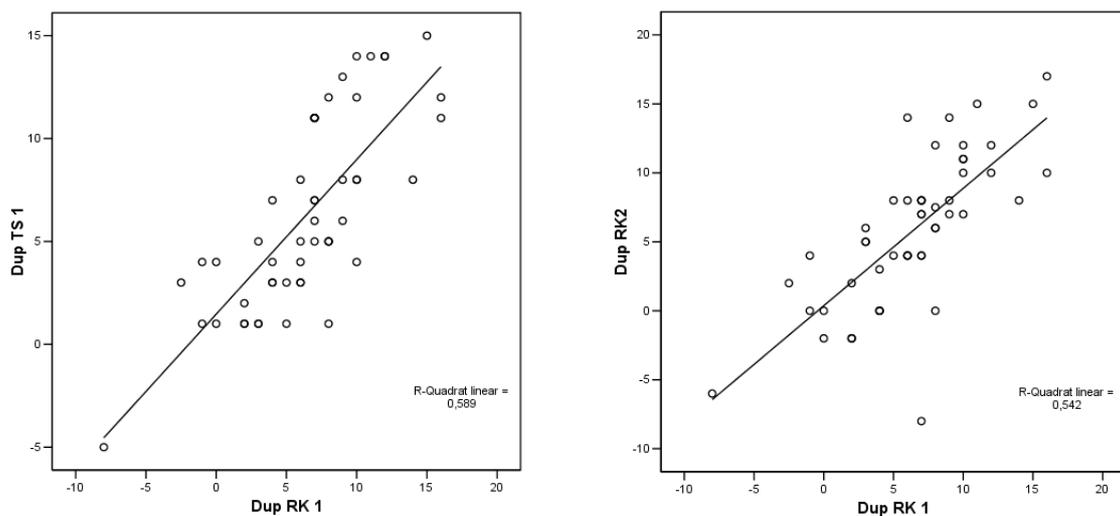


Abb. 16 (links): Interindividuelle Korrelation des Verfahrens Dupius-Angulation, Untersucher RK und Untersucher TS zum Zeitpunkt 1 ($r=0,767$)

Abb. 17 (rechts): Intraindividuelle Korrelation des Verfahrens Dupius-Angulation, Untersucher RK zum Zeitpunkt 1 und 2 ($r=0,736$)

5.1.3. Kennwerte und Übereinstimmungsberechnungen

Parameter 1: Dupius Angulation

Vergleich	Mittlere Differenz	Signifikanz	Std. Abweichung	Korrelation	% Übereinstimmung	% Nicht-Übereinstimmung
RK1 - RK2	0,59	0,266	3,73	0,736	35,3	64,7
TS 1 - TS 2	0,51	0,313	2,89	0,793	43,1	56,9
CH 1 - CH 2	0,02	0,960	2,79	0,797	35,3	64,7
<i>MW Intra-Beurt.</i>	<i>0,37</i>		<i>3,14</i>	<i>0,777</i>	<i>37,9</i>	<i>62,1</i>
RK1 - TS 1	0,13	0,775	3,17	0,767	25,5	74,5
RK 1 - TS 2	0,54	0,240	3,24	0,754	43,1	56,9
RK 1 - CH 1	0,50	0,186	2,66	0,832	58,8	41,2
RK 1 - CH 2	0,52	0,214	2,95	0,787	39,2	60,8
RK 2 - TS 1	0,46	0,395	3,84	0,718	47,1	52,9
RK 2 - TS 2	0,05	0,920	3,46	0,771	41,2	58,8
RK 2 - CH 1	1,09	0,047	3,81	0,719	35,3	64,7
RK 2 - CH 2	1,11	0,031	3,55	0,756	39,2	60,8
TS 1 - CH 1	0,63	0,042	2,14	0,888	45,1	54,9
TS 1 - CH 2	0,65	0,081	2,60	0,831	45,1	54,9
TS 2 - CH 1	1,04	0,007	2,64	0,821	52,9	47,1
TS 2 - CH 2	1,06	0,013	2,94	0,767	39,2	60,8
<i>MW Inter-Beurt.</i>	<i>0,65</i>		<i>3,08</i>	<i>0,790</i>	<i>42,6</i>	<i>57,4</i>
Durchschnitt (insges.)	0,59		3,09	0,787	41,7	58,3

Tab. 9: Kennwerte für das Verfahren Dupius-Angulation

Das Verfahren nach Dupius-Angulation zeigt bei einer durchschnittlichen Intra-Beurteiler-Abweichung von $0,37^\circ$ und einer Inter-Beurteiler-Abweichung von $0,65^\circ$ eine durchschnittliche Standardabweichung von $3,09^\circ$. Die Gesamtübereinstimmung beträgt $41,7\%$ (Toleranz $\pm 1^\circ$). Bei $p < 0,05$ (Tab. 9, Spalte 3), misst einer der beiden Untersucher systematisch höhere Werte.

Das Verfahren nach Stokes-Angulation zeigt bei einer durchschnittlichen Intra-Beurteiler-Abweichung von $0,80^\circ$ und einer Inter-Beurteiler-Abweichung von $0,93^\circ$ eine durchschnittliche Standardabweichung von $5,51^\circ$. Die Gesamtübereinstimmung beträgt $33,2\%$ (Toleranz $\pm 1^\circ$). Bei $p < 0,05$ (Tab. I im Anhang, Spalte 3), misst einer der beiden Untersucher systematisch höhere Werte.

Das Verfahren nach Panjabi-Angulation zeigt bei einer durchschnittlichen Intra-Beurteiler-Abweichung von $0,62^\circ$ und einer Inter-Beurteiler-Abweichung von $0,58^\circ$ eine durchschnittliche Standardabweichung von $3,19^\circ$. Die Gesamtübereinstimmung

berägt 41,1 % (Toleranz +/-1°). Bei $p < 0,05$ (Tab. II im Anhang, Spalte 3), misst einer der beiden Untersucher systematisch höhere Werte.

Das Verfahren nach Dupius-Translation zeigt bei einer durchschnittlichen Intra-Beurteiler-Abweichung von 0,64 mm und einer Inter-Beurteiler-Abweichung von 0,73 mm eine durchschnittliche Standardabweichung von 2,35 mm. Die Gesamtübereinstimmung beträgt 47,1 % (Toleranz +/-1 mm). Bei $p < 0,05$ (Tab. III im Anhang, Spalte 3), misst einer der beiden Untersucher systematisch höhere Werte.

Das Verfahren nach Morgan-Translation zeigt bei einer durchschnittlichen Intra-Beurteiler-Abweichung von 0,59 mm und einer Inter-Beurteiler-Abweichung von 0,60 mm eine durchschnittliche Standardabweichung von 2,58 mm. Die Gesamtübereinstimmung beträgt 45,4 % (Toleranz +/-1 mm). Bei $p < 0,05$ (Tab. IV im Anhang, Spalte 3), misst einer der beiden Untersucher systematisch höhere Werte.

Das Verfahren nach Panjabi-Translation zeigt bei einer durchschnittlichen Intra-Beurteiler-Abweichung von 0,76 mm und einer Inter-Beurteiler-Abweichung von 0,54 mm eine durchschnittliche Standardabweichung von 1,93 mm. Die Gesamtübereinstimmung beträgt 54,4 % (Toleranz +/-1 mm). Bei $p < 0,05$ (Tab. VI im Anhang, Spalte 3), misst einer der beiden Untersucher systematisch höhere Werte als der andere.

Kennwert	Angulation	Translation
durchschnittliche Intra- Beurteiler-Variabilitäten	Dup = 0,37°	Dup = 0,64 mm
	Sto = 0,80°	Morg = 0,59 mm
	Panj = 0,62°	Panj = 0,76 mm
durchschnittliche Inter- Beurteiler-Variabilitäten	Dup = 0,65°	Dup = 0,73 mm
	Sto = 0,93°	Morg = 0,60 mm
	Panj = 0,58°	Panj = 0,54 mm
durchschnittliche Standardabweichungen	Dup = 3,09°	Dup = 2,35 mm
	Sto = 5,51°	Morg = 2,58 mm
	Panj = 3,19°	Panj = 1,93 mm
Korrelationen	Dup = 0,787	Dup = 0,880
	Sto = 0,788	Morg = 0,810
	Panj = 0,851	Panj = 0,900
durchschnittliche Übereinstimmungen	Dup = 41,1 %	Dup = 47,1 %
	Sto = 33,2 %	Morg = 45,4 %
	Panj = 41,1 %	Panj = 54,4 %
durchschnittliche Nicht- Übereinstimmungen	Dup = 58,3 %	Dup = 52,9 %
	Sto = 66,8 %	Morg = 54,6 %
	Panj = 58,9 %	Panj = 45,6 %

Tab. 10: Zusammenfassung der Kennwerte getrennt nach angulatorischen und translatorischen Verfahren

5.1.4. Bewertung der Messmethoden hinsichtlich signifikanter Unterschiede in der Beurteilerübereinstimmung

5.1.4.1. Unterschiede zwischen Messmethode Dupius-Angulation und Stokes-Angulation

	Beobachtetes N	Erwartete Anzahl	Residuum
Nicht Übereinstimmung	583	668,0	-85,0
Übereinstimmung	417	332,0	85,0
Gesamt	1000		

	Dupius Angulation
Chi-Quadrat ^a	32,578
df	1
Asymptotische Signifikanz	,000

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 332,0.

Abb. 18: Statistische Auswertung des Vergleiches Dupius vs. Stokes in Angulation
Das Verfahren Dupius-Angulation unterscheidet sich in seinen Übereinstimmungen signifikant vom Verfahren Stokes-Angulation zugunsten einer höheren Trefferrate beim Verfahren nach Dupius-Angulation ($p < 0,001$).

5.1.4.2. Unterschiede zwischen Messmethode Dupius-Angulation und Panjabi-Angulation

	Beobachtetes N	Erwartete Anzahl	Residuum
Nicht Übereinstimmung	583	589,0	-6,0
Übereinstimmung	417	411,0	6,0
Gesamt	1000		

Statistik für Test	
	Dupius Angulation
Chi-Quadrat ^a	,149
df	1
Asymptotische Signifikanz	,700

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 411,0.

Abb. 19: Statistische Auswertung des Vergleiches Dupius vs. Panjabi in Angulation

Das Verfahren Dupius-Angulation unterscheidet sich in seinen Übereinstimmungen nicht signifikant vom Verfahren Panjabi-Angulation ($p=0,7$).

5.1.4.3. Unterschiede zwischen Messmethode Stokes-Angulation und Panjabi-Angulation

Stokes Angulation vs. Panjabi Angulation			
	Beobachtetes N	Erwartete Anzahl	Residuum
Nicht Übereinstimmung	668	589,0	79,0
Übereinstimmung	332	411,0	-79,0
Gesamt	1000		

Statistik für Test	
	Stokes Angulation
Chi-Quadrat ^a	25,781
df	1
Asymptotische Signifikanz	,000

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 411,0.

Abb. 20: Statistische Auswertung des Vergleiches Stokes vs. Panjabi in Angulation

Das Verfahren Stokes-Angulation unterscheidet sich in seinen Übereinstimmungen signifikant vom Verfahren Panjabi-Angulation zugunsten einer höheren Trefferrate beim Verfahren nach Panjabi-Angulation ($p<0,001$).

5.1.4.4. Unterschiede zwischen Messmethode Dupius-Translation und Morgan-Translation

	Beobachtetes N	Erwartete Anzahl	Residuum
Nicht Übereinstimmung	529	546,0	-17,0
Übereinstimmung	471	454,0	17,0
Gesamt	1000		

	Dupius Translation
Chi-Quadrat ^a	1,166
df	1
Asymptotische Signifikanz	,280

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 454,0.

Abb. 21: Statistische Auswertung des Vergleiches Dupius vs. Morgan in Translation

Das Verfahren Dupius-Translation unterscheidet sich in seinen Übereinstimmungen nicht signifikant vom Verfahren Morgan-Translation ($p=0,28$).

5.4.1.5. Unterschiede zwischen Messmethode Dupius-Translation und Panjabi-Translation

	Beobachtetes N	Erwartete Anzahl	Residuum
Nicht Übereinstimmung	529	456,0	73,0
Übereinstimmung	471	544,0	-73,0
Gesamt	1000		

	Dupius Translation
Chi-Quadrat ^a	21,482
df	1
Asymptotische Signifikanz	,000

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 456,0.

Abb. 22: Statistische Auswertung des Vergleiches Dupius vs. Panjabi in Translation

Das Verfahren Dupius-Translation unterscheidet sich in seinen Übereinstimmungen signifikant vom Verfahren Panjabi-Translation zugunsten einer höheren Trefferrate beim Verfahren nach Panjabi-Translation ($p < 0,001$).

5.1.4.6. Unterschiede zwischen Messmethode Morgan-Translation und Panjabi-Translation

	Beobachtetes N	Erwartete Anzahl	Residuum
Nicht Übereinstimmung	546	456,0	90,0
Übereinstimmung	454	544,0	-90,0
Gesamt	1000		

	Morgan Translation
Chi-Quadrat ^a	32,653
df	1
Asymptotische Signifikanz	,000

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 456,0.

Abb. 23: Statistische Auswertung des Vergleiches Morgan vs. Panjabi in Translation

Das Verfahren Morgan-Translation unterscheidet sich in seinen Übereinstimmungen signifikant vom Verfahren Panjabi-Translation zugunsten einer höheren Trefferrate beim Verfahren nach Panjabi-Translation ($p < 0,001$).

5.1.5. Beispiele der Einzelergebnisse – Vergleich der Beurteiler im Detail

Die Ergebnisse t-Test, Histogramm der Differenzen und Streudiagramm nach BLANT und ALTMAN wurden für alle Einzelpaarungen erstellt. Aus Gründen der Übersichtlichkeit werden die Resultate der Berechnungen lediglich zusammenfassend dargestellt.

Nachfolgend ist ein Beispiel des Ergebnisses einer intraindividuellen (RK1 – RK2) und einer interindividuellen (RK1 – TS1) Analyse anhand des Messverfahrens nach DUPIUS in Angulationsrichtung.

5.1.5.1. Beispiel Dupius – Angulation

5.1.5.1.1. Vergleich RK1 – RK2

T-Test

		Mittelwert	N	Standardabweichung	Standardfehler des Mittelwertes
Paaren	RK1	6,42	51	4,69	,66
1	RK2	5,83	51	5,42	,76

		N	Korrelation	Signifikanz
Paaren 1	RK1 & RK2	51	,736	,000

		Gepaarte Differenzen							
		Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	RK1 - RK2	,59	3,73	,52	-,46	1,64	1,125	50	,266

Abb. 24: Statistische Auswertung des intraindividuellen Paarvergleiches RK1 – RK2. Gute Korrelation der beiden Messungen ($r=0,736$).

Histogramm der Differenzen

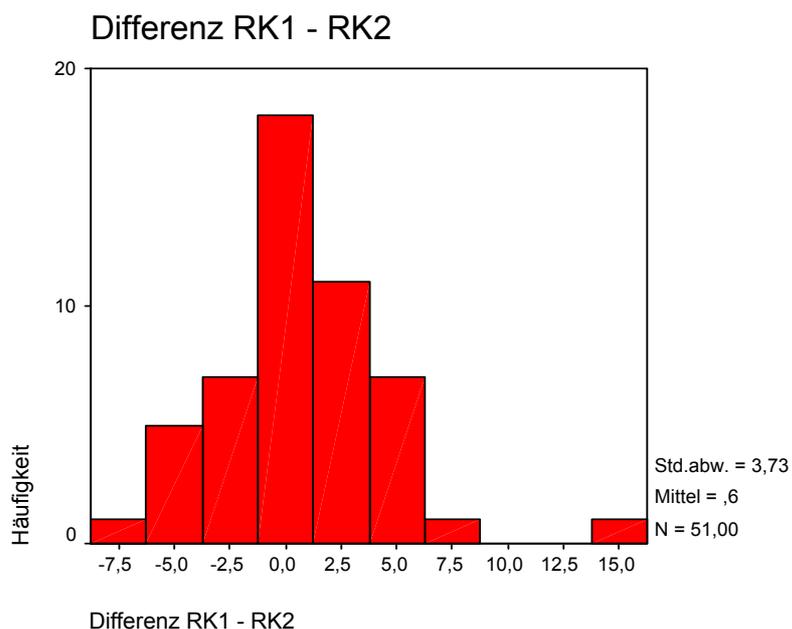
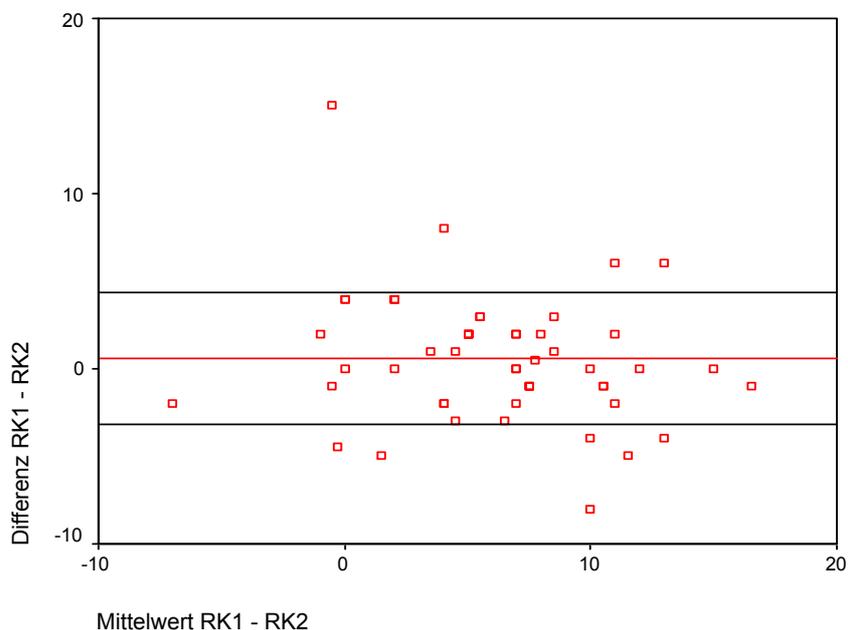


Abb. 25: Graphische Darstellung der intraindividuellen Differenzen RK1 – RK2

Plot: Differenzen gegen Mittelwerte der Messungen

Abb. 26: Graphische Darstellung der Mittelwerte gegen die Differenzen nach BLANT und ALTMAN. Die Standardabweichung ($3,37^\circ$) entspricht dem intraindividuellen Fehler dieser Messung.

Prozentuale Übereinstimmung

Übereinstimmung RK1 - RK2

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Keine Übereinstimmung	33	64,7	64,7	64,7
	Übereinstimmung	18	35,3	35,3	100,0
Gesamt		51	100,0	100,0	

Abb. 27: Darstellung der prozentualen Übereinstimmung der Messung RK1 – RK2. Trotz guter Korrelation ($r=0,736$; siehe Abb. 24) findet sich bei den gewählten Entscheidungsgrenzen ($\pm 1^\circ$) lediglich eine Übereinstimmung von 35,3 %.

5.1.5.1.2. Vergleich RK1 – TS1

T-Test

Statistik bei gepaarten Stichproben

		Mittelwert	N	Standardab- weichung	Standardfehler des Mittelwertes
Paaren	RK1	6,42	51	4,69	,66
1	TS1	6,29	51	4,59	,64

Korrelationen bei gepaarten Stichproben

		N	Korrelation	Signifikanz
Paaren 1	RK1 & TS1	51	,767	,000

Test bei gepaarten Stichproben

		Gepaarte Differenzen							
		Mittelwert	Standardab- weichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz		T	df	Sig. (2-seitig)
					Untere	Obere			
Paaren 1	RK1 - TS1	,13	3,17	,44	-,76	1,02	,287	50	,775

Abb. 28: Statistische Auswertung des interindividuellen Paarvergleiches RK1 – TS1. Gute Korrelation der beiden Messungen ($r=0,767$).

Histogramm der Differenzen

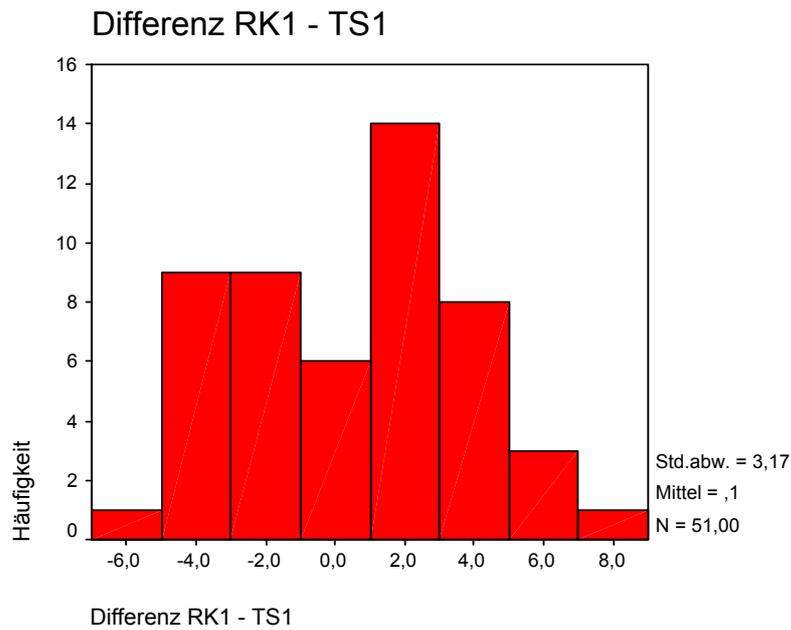
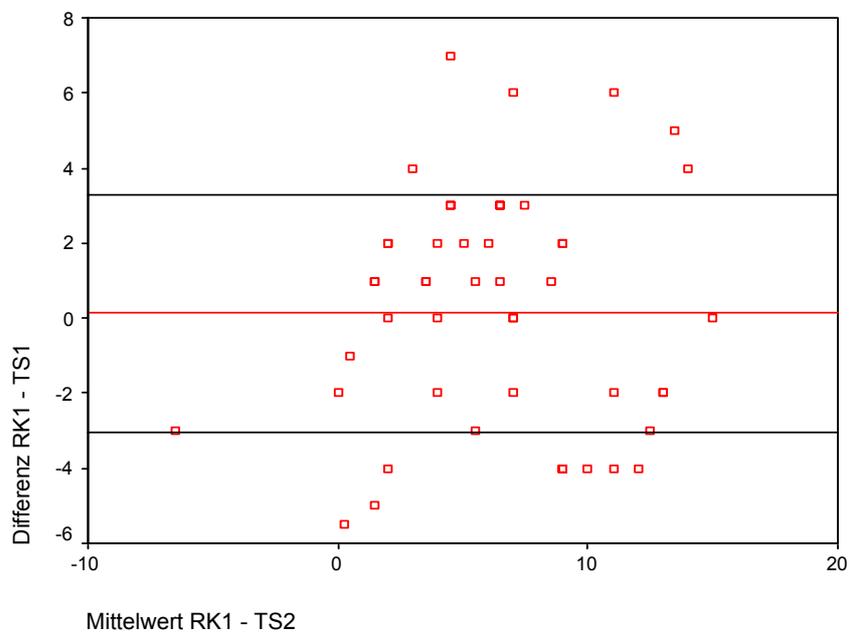


Abb. 29: Graphische Darstellung der interindividuellen Differenzen RK1 – TS1

Plot: Differenzen gegen Mittelwerte der Messungen

Abb. 30: Graphische Darstellung der Mittelwerte gegen die Differenzen nach BLANT und ALTMAN. Die Standardabweichung ($3,17^\circ$) entspricht dem interindividuellen Fehler dieser Messung.

Prozentuale Übereinstimmung

Übereinstimmung RK1 - TS 1

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Keine Übereinstimmung	38	74,5	74,5	74,5
	Übereinstimmung	13	25,5	25,5	100,0
Gesamt		51	100,0	100,0	

Abb. 31: Darstellung der prozentualen Übereinstimmung der Messung RK1 – TS1. Trotz guter Korrelation ($r=0,767$; siehe Abb. 28) findet sich bei den gewählten Entscheidungsgrenzen ($\pm 1^\circ$) lediglich eine Übereinstimmung von 25,5 %.

5.1.6. Zusammenfassung hinsichtlich der Übereinstimmungen der Messmethoden

Die deutlich beste Übereinstimmung über alle Einzelübereinstimmungen hinweg (Intra- und Inter-Beurteilerübereinstimmung) ergibt sich mit 54,4% Übereinstimmung in den Urteilen für die Messmethode Panjabi-Translation (Tab. 11). Die schlechteste Übereinstimmung findet man bei Strokes-Angulation (Tab. 11). Bis auf zwei Ausnahmen (Dupius-Translation und Morgan-Translation sowie Dupius-Angulation und Panjabi-Angulation) sind alle Einzelvergleiche zwischen den Messmethoden statistisch signifikant unterschiedlich. Die Unterschiede sind jeweils auf dem 1%-Niveau signifikant (Chi-2-Test), d. h. die Methoden sind jeweils dann eindeutig besser bzw. schlechter in der Übereinstimmung.

Messmethode	Ø Übereinstimmung (über alle Einzelvergleiche hinweg)
Panjabi – Translation	54,4%
Dupius – Translation	47,1%
Morgan – Translation	45,4%
Dupius – Angulation	41,7%
Panjabi – Angulation	41,1%
Strokes – Angulation	33,2%

Tab. 11: Zusammenfassung der prozentualen Übereinstimmungen der einzelnen Messverfahren

Es fällt auf, dass die translatorischen Messmethoden generell besser in ihrer Übereinstimmung sind als die angulatorischen Verfahren. Zusammenfassung der vier unterschiedlichen Niveaus in der Übereinstimmung bei den sechs untersuchten Messverfahren. Dupius-Translation und Morgan-Translation, sowie Dupius-Angulation und Panjabi-Angulation sind als ein Niveau in der Übereinstimmung zu

werten. Zwischen den Verfahren in unterschiedlichen Zellen der Tabelle bestehen jeweils mindestens auf dem 1%-Niveau signifikante Unterschiede.

5.1.7. Diskussion

Änderungen der kinematischen Verhältnisse an der Lendenwirbelsäule können Ursache für behandlungsbedürftige Beschwerden sein [28, 34, 67, 68, 94]. Funktionsaufnahmen im seitlichen Strahlengang sind auch heute noch das bildgebende Diagnostikum der Wahl [24, 45, 147], wenn die Frage der Indikationsstellung zur Fusionsoperation bei einem Verdacht auf klinisch relevante Instabilität beantwortet werden soll [9, 47, 64, 73, 119, 142, 163, 176]. Seit der Erstbeschreibung der Technik durch BAKKE (1931) [5] hat die Methode eine Vielzahl von Modifikationen erfahren [2, 5, 6, 41, 49, 54, 80, 81, 89, 115, 133, 150, 151, 168]. In den letzten Jahrzehnten sind zahlreiche, teilweise auch sehr komplizierte, Anwendungsmöglichkeiten beschrieben worden [15, 147, 155]. Das Verfahren erlebt derzeit durch die steigenden Zahlen von Bandscheibenprothesen-Implantationen unweigerlich eine Renaissance [50, 51, 79]. Andere Autoren haben versucht, sich über Verfahren wie Videofluoroskopie dem Problem der lumbalen Instabilität zu nähern [43, 61, 62, 98, 149].

Es herrscht wohl Einigkeit darüber, dass eine sagittale Translation von ≥ 3 mm und eine segmentale Angulation von $10-12^\circ$ pathologischen Wert haben können [9, 53, 176]. Höhere Werte (5-6 mm bzw. 10-15 % der Wirbelkörperlänge in der Translationsebene und $15-25^\circ$ in der Angulationsebene) werden jedoch ebenfalls angegeben [22, 23, 45, 68, 124, 138]]. IGUCHI und Mitarbeiter fanden heraus, dass der sagittalen Translation größere Bedeutung zukommt als der angulären Abweichung. Eine Kombination von beidem prädisponiert besonders zu klinischer Symptomatik [53]. Alters- und Geschlechtseinflüsse sind bekannt [52], so dass die Interpretation der Aufnahmen nicht leicht ist. Sehr wesentlich ist auch das Verfahren der Auswertung, welches für die Interpretation herangezogen wird. SHAFFER et al. [138] untersuchten in ihrer preisgekrönten Arbeit experimentell die Genauigkeit von verschiedenen Methoden der Röntgenbildauswertung. Sie zeigten, dass die Rate falsch-positiver und falsch-negativer Resultate von der Qualität der

Röntgenaufnahmen abhängt, aber insgesamt sehr gut ist (Übereinstimmungen 0,952...0,995). Bei der Beurteilung der verschiedenen Auswerteverfahren wurden deutliche Differenzen in der Genauigkeit festgestellt. Für alle untersuchten Verfahren wurden falsch-positive Raten von 25 %, 17 % bzw. 13 % gefunden, je nach Ausmaß der zu beurteilenden Strecke (+/-3mm, +/-3,5mm bzw. +/-4mm). Falsch-negative Raten fanden sich in 29 %, 37 % bzw. 49 %. Die daraufhin auch in unserer Studie verwendete Methode nach MORGAN und KING [91] erzielte die besten Ergebnisse. Bei der Untersuchung klinischer Röntgenbilder wurde geurteilt, dass ihre Heranziehung zur Diagnosestellung einer Instabilität zu falsch-positiven Resultaten führen kann. Dies ist weniger häufig, wenn 5 mm Translationsstrecke oder mehr zu beurteilen sind. Trotz der exakten experimentellen Untersuchungstechnik wird die Problematik der Fehlbeurteilung bereits deutlich. Hinzu kommen die unterschiedlichen Urteile bei mehrfacher Bewertung durch einen Untersucher (Intra-Beurteiler-Variabilität) und die Abweichung bei der Untersuchung eines Falles durch mehrere Beurteiler (Inter-Beurteiler-Variabilität). WOOD und Mitarbeiter [176] fanden in ihrer Studie eine Interobserver-Variabilität von 0,92 mm in der Translationsbeurteilung und 1,7° bei der Beurteilung der Angulation. Die Intraobserver-Variabilitäten betragen 1 mm und 2,27°. MAIGNE [85] gab in seiner Studie keinen signifikanten Intraobserverfehler und einen Interobserverfehler von 0,13 % an. An Röntgenbildern in Funktion nach Implantation einer künstlichen Bandscheibe wurden von LIM et al. [79] für angulatorische Beurteilungen Intraobserver-Variabilitäten von 1,4° bis 3,3° – je nach Verfahren – angegeben. Die Interobserver-Variabilitäten betragen hier 1,8° bis 3,3°. In unserer Studie zeigte sich eine relativ geringe Intraobserver-Variabilität von durchschnittlich 0,60° (0,37...0,80) für die angewendeten angulatorischen Verfahren und von durchschnittlich 0,66 mm (0,59...0,76) für die translatorischen Verfahren. Es fand sich weiterhin eine relativ geringe Interobserver-Variabilität von durchschnittlich 0,72° (0,58...0,93) für die angewendeten angulatorischen Verfahren und von durchschnittlich 0,62 mm (0,54...0,73) für die translatorischen Verfahren. Eine Analyse der Streuungen in Form der errechneten Standardabweichungen zeigt die Problematik jedoch deutlicher. In unserer Studie fanden sich durchschnittliche Werte über alle Einzelpaarvergleiche von 3,93° (3,09...5,51) für die angulatorischen Verfahren und 2,29 mm (1,93...2,58) für die translatorischen Verfahren. Es zeigte sich weiterhin sehr deutlich, dass eine alleinige Bestimmung der Korrelationskoeffizienten nicht aussagekräftig genug ist. Es

fanden sich für alle untersuchten Parameter gute bis sehr gute Korrelationen ($r=0,787\dots0,900$). Noch deutlicher zeigt sich dieses Problem bei der Analyse der tatsächlich gefundenen Übereinstimmungen im klinischen Alltag. Angulatorische Verfahren stimmen bei einer Toleranz von 1° in 38,7 (33,2...41,7) % der Fälle überein. Translatorische Verfahren stimmen bei einer Toleranz von 1 mm in 49 (45,4...54,4) % der Fälle überein. Diese Werte sind natürlich durch die Festlegung der akzeptierten Toleranzen beeinflussbar. Unsere Grenzen von 1 mm bzw. 1° ergaben sich aus der Genauigkeit der verwendeten handelsüblichen Messutensilien (Lineal, Winkelmesser) und aus den Werten der in der Literatur beschriebenen pathologischen Kennwerte.

Die Indikationsstellung zur Spondylodese sollte von klinischen, radiologischen und biomechanischen Parametern getragen werden. Hierzu sind exakte Techniken der Datenerhebung unerlässlich. Bei der täglichen Arbeit sollte man sich des unvermeidlichen Fehlers der einzelnen Verfahren der Beurteilung von Funktionsaufnahmen bewusst sein. Alle Verfahren sind in nicht geringem Maße untersucherabhängig. Translatorische Messverfahren sind genauer als angulatorische Verfahren. Insgesamt erscheinen nach unseren Daten die Interpretation der Messergebnisse und mehr noch die Ableitung therapeutischer Konsequenzen aufgrund der hohen Streubreite schwierig. Die Einbeziehung weiterer objektiver Entscheidungshilfen ist deshalb bei der Indikationsstellung zur operativen Therapie, die ja immer eine erhebliche Konsequenz für den betroffenen Patienten darstellt, unbedingt notwendig.