

Aus der Klinik für Innere Medizin  
der Philipps-Universität Marburg  
Geschäftsführender Direktor: Prof. Dr. med. T. Gress  
Medizinische Klinik, S.P. Pneumologie  
Direktor: Prof. Dr. med. C. Vogelmeier

---

in Zusammenarbeit mit dem Universitätsklinikum Gießen und Marburg GmbH,  
Standort Marburg

**Reliabilitätsuntersuchung  
zur standardisierten audio-visuellen  
Bewertung von Wheezing**

INAUGURAL – DISSERTATION  
zur Erlangung des Doktorgrades der gesamten Medizin  
dem Fachbereich Medizin der  
Philipps-Universität Marburg  
vorgelegt

von  
**JULIA WILMA GUNDA BENZ**  
aus Bochum

Marburg 2008

Angenommen vom Fachbereich Humanmedizin  
der Philipps-Universität Marburg am 17.03.2008.

Gedruckt mit Genehmigung des Fachbereichs.

Dekan:	Prof. Dr. med. M. Rothmund
Referent:	Prof. Dr. med. U. Koehler
Korreferent:	PD Dr. med. A. Torossian

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Die Lungengeräusche	1
1.2	Wheezing	3
1.3	Obstruktive Lungenerkrankungen	6
1.4	Die Lungengeräuschanalyse	8
<b>2</b>	<b>Zielsetzung und Fragestellungen der Arbeit</b>	<b>11</b>
<b>3</b>	<b>Patienten und Methoden</b>	<b>12</b>
3.1	Das Untersuchungskollektiv	12
3.2	Methoden	14
3.2.1	Standardisierte Atmung	14
3.2.2	Das Messgerät	15
3.2.3	Die audio-visuelle Bewertung der Lungengeräuschaufzeichnungen	16
3.2.4	Das Lernset	18
3.2.5	Das Testset	19
3.2.6	Statistische Verfahren der Auswertung	21
<b>4</b>	<b>Ergebnisse</b>	<b>22</b>
4.1	Statistische Datenanalyse	22
4.1.1	Reliabilität der Untersucher	22
4.1.2	Re-Test Reliabilität der Untersucher	28
4.1.3	Zusammenhang der Beurteilungsaspekte	32
<b>5</b>	<b>Diskussion</b>	<b>33</b>
5.1	Diskussion der Methode	33
5.1.1	Die Messungen	33
5.1.2	Die Rater	33
5.1.3	Lernset und Testset	34
5.2	Diskussion der Ergebnisse	36
5.2.1	Reliabilität der Untersucher	36
5.2.2	Re-Test Reliabilität der Untersucher	38
5.2.3	Zusammenhang der Beurteilungsaspekte	40
5.3	Ausblick	41
<b>6</b>	<b>Zusammenfassung</b>	<b>42</b>
6.1	Summary	44
<b>7</b>	<b>Literaturverzeichnis</b>	<b>46</b>
<b>8</b>	<b>Anhang</b>	<b>54</b>
8.1	Abbildungsverzeichnis	54
8.2	Tabellenverzeichnis	55
8.3	Abkürzungsverzeichnis	56
8.4	Graphische Darstellungen	57
8.5	Verzeichnis der akademischen Lehrer	67
8.6	Danksagung	68

# 1 Einleitung

Die Auskultation der Lunge ist heute ein fester Bestandteil der körperlichen Untersuchung und das Stethoskop seit seiner Erfindung durch den französischen Mediziner Laennec (1819) ein effektives Diagnosewerkzeug [Laennec R, 1819]. Bei der Auskultation wird über die Beurteilung der Lungengeräusche, die über die Thoraxwand übertragen werden, eine Einschätzung in Hinblick auf physiologische und pathologische Prozesse möglich.

## 1.1 Die Lungengeräusche

Die Lungengeräusche werden in physiologische Atem- und pathologische Nebengeräusche unterteilt. Das Atemgeräusch eines Gesunden wird als *peripheres* oder *normales Atemgeräusch* bezeichnet. Es ist als nicht-musikalisches Geräusch definiert und diverse Adjektive, wie „normal“, „verschärft“ oder „vermindert“ können zur genaueren Beschreibung herangezogen werden. Das periphere Atemgeräusch ist ein niederfrequentes Geräusch, dessen Hauptfrequenz unter 100 Hz bleibt [Forgacs P, 1978a+b; Gavriely N, 1981]. Es gibt jedoch Unterschiede in der In- und Expirationsphase. Nach den gegenwärtigen „Computerized Respiratory Sound Analysis“ (CORSA) - Richtlinien für die Beurteilung von Lungengeräuschen, gilt ein Geräusch über 600 Hz bei der Inspiration und über 400 Hz bei der Expiration, nicht mehr als peripheres Lungengeräusch [Sovijärvi ARA, 2000b].

Das *tracheale Atemgeräusch* kann in der suprasternalen Vertiefung über der Trachea oder am seitlichen Nacken gehört werden. Es wird als Breitspektrumgeräusch bezeichnet, da es eine Frequenzbreite von ca. 100 bis 1.500 Hz umfasst [Charbonneau G, 1987; Gavriely N, 1996].

Das *bronchio-vesiculäre Atemgeräusch* wird bei gesunden Patienten über dem Brustbein gehört und als „hohl“ oder „tubulär“ bezeichnet. Unter pathologischen Bedingungen, z. B. bei einer Infiltration, kann es als *bronchiales Atemgeräusch* auch an peripheren Auskultationspunkten auftreten. In diesem Fall ist durch die Infiltration, im Vergleich zur gesunden Lunge, die Schalleitung verbessert [Loudon RG, 1984].

Die Nebengeräusche werden unterteilt in *kontinuierliche* und *diskontinuierliche* Geräusche. Zu den kontinuierlichen Nebengeräuschen werden neben dem Wheezing, hinweisend auf eine obstruktive Atemwegserkrankung oder einen Bronchospasmus, der Stridor (Verengung der oberen Atemwege) und die Rhonchi (Sekretansammlung) gezählt. Unter die diskontinuierlichen Nebengeräusche fallen die Crackles. Sie treten auf bei bronchialer Flüssigkeitsansammlung und werden unterteilt in „grob“ und „fein“ [American Thoracic Society, 1977]. Zusätzlich sei als Nebengeräusch noch der Husten genannt. Es ist ein in Intensität und Tonhöhe hoch variables Geräusch, erzeugt durch den Hustenreflex [Kelemen SA, 1987]. Husten, einhergehend mit Wheezing, ist typisch für das Asthma bronchiale.

<b><i>Einteilung der Lungengeräusche</i></b>		
<b><i>Physiologische Atemgeräusche</i></b>	<b><i>Ursprungsort</i></b>	<b><i>Bedeutung</i></b>
Peripheres AG	Zentrale Atemwege	Regionale Belüftung
Tracheales AG	Obere Atemwege	Konfiguration der oberen Atemwege
Broncho-vesikuläres AG	Zentrale + untere Atemwege	Konfiguration der unteren Atemwege
<b><i>Pathologische Nebengeräusche</i></b>		
<b><i>„kontinuierlich“</i></b>		
Wheezing	Zentrale + untere Atemwege	Obstruktion, Flusslimitierung
Stridor	Obere Atemwege	Verengung obere Atemwege
Rhonchi	Zentrale Atemwege	Sekret, Kaliberschwankungen
<b><i>„diskontinuierlich“</i></b>		
Crackles	Zentrale + untere Atemwege	Sekret, Ödem, Stauung
Husten	Zentrale + obere Atemwege	Hustenreflex, multiple Ursachen

Tabelle 1: Übersicht über die Einteilung der Lungengeräusche nach der aktuellen Nomenklatur der „International Lung Sound Association“.

## 1.2 Wheezing

Das zu den kontinuierlichen Nebengeräuschen zählende Wheezing wird als Begriff in der Pneumologie sehr häufig in Zusammenhang mit obstruktiven Lungenerkrankungen verwendet. Es wird als musikalisches Geräusch bezeichnet und kann anhand seiner Lokalisation, Intensität, Tonhöhe, Dauer und Beziehung zur Respirationsphase charakterisiert werden [Meslier N, 1995].

Nach der aktuellen Definition der CORSA-Richtlinien erstreckt sich die Frequenzbreite von Wheezing von  $\geq 100$  Hz bis  $\leq 1$  kHz mit einer spezifischen Dauer von mehr als 100 ms [Sovijärvi ARA, 2000b]. Wheezing kann bei einem Patienten über der Lunge, aber auch über der Trachea detektiert werden.

In einer Lungengeräuschaufzeichnung - wie in Abb. 1 dargestellt - kann man Wheezing als bandförmige, quer verlaufende dunkle Linien erkennen, die auf ein Ereignis von hoher Intensität hinweisen. Während der Ausatmung zeigt sich eher monophones Wheezing, wohingegen es in der Einatmung polyphon, also mit mehreren Frequenzen gleichzeitig, auftritt. Die senkrechte Achse gibt die Frequenz, die horizontale die Zeit wieder. Im unteren Abschnitt ist parallel zu den 4 Atemzyklen der gemessene Atemfluss in Litern pro Sekunde dargestellt.

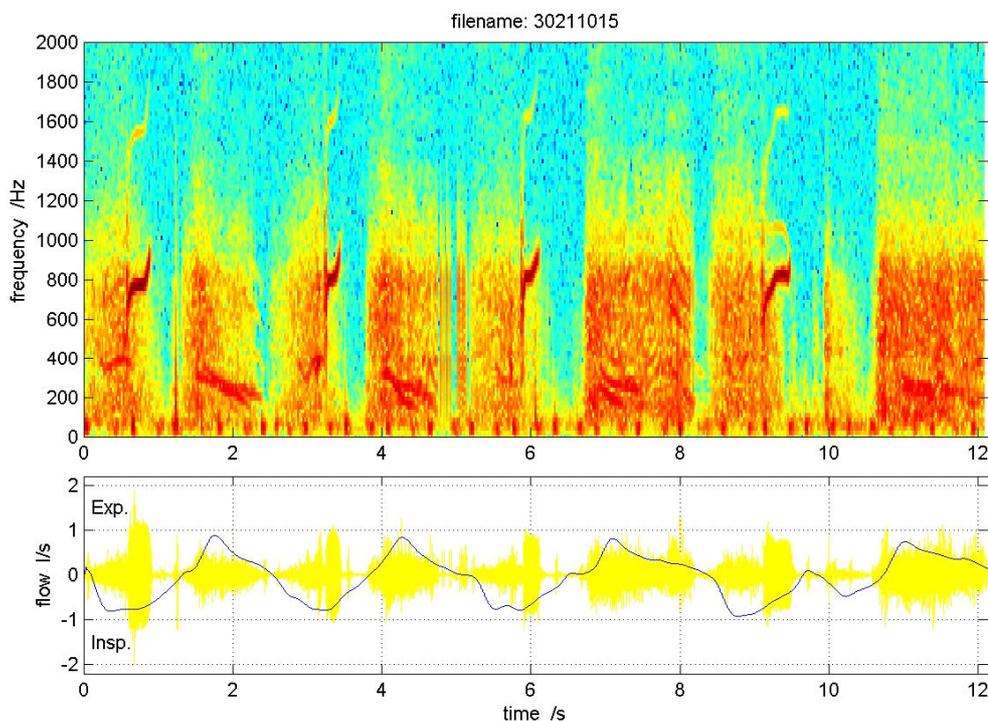


Abb. 1: Beispiel einer Lungengeräuschaufzeichnung mit Wheezing (tracheales Mikrophon) in den vier Ventilationsphasen nach Anwendung der Fast Fourier Transformation.

Die Mechanismen, die bei der Produktion von Wheezing eine Rolle spielen, sind noch nicht vollends geklärt. Sehr wahrscheinlich liegt die Ursache in einer Interaktion der Atemwegswände mit dem Luftstrom. Forgacs führte die Entstehung von Geräuschen mit hoher Tonhöhe auf Kaliberschwankungen in den Luftwegen zurück, bei denen sich gegenüberliegende Wände berühren [Forgacs P, 1967]. Aufgrund der Beschleunigung der Luft durch das verschmälerte Lumen und die Druckschwankungen werden die Wände der Luftwege in Schwingungen versetzt.

Einen entscheidenden Beitrag zur Entstehungstheorie von Wheezing lieferte die Arbeitsgruppe um Grotberg. Sie entwickelte ein theoretisch-mathematisches Modell zu Atemwegsflattern. In ihrer Arbeit konnte sie Wheezing mit Flusslimitierung in kollabierbaren Röhren in Verbindung bringen und nahm damit Bezug zu Forgacs Studien [Grotberg JB, 1980 u. 1984]. Nach ihrer Definition geht Wheezing immer mit Flusslimitierung einher, Flusslimitierung aber nicht notwendigerweise immer mit Wheezing. Untermuert von weiteren Studien wird Atemwegsflattern in Kombination mit Sekretbewegungen heute als ursächlich für die Entstehung von Wheezing angenommen [Gavriely N, 1989a]. Des Weiteren konnten eine „kritische Gasflussgeschwindigkeit“ und ein „kritischer interpleuraler Druck“ definiert werden, ohne dessen Erreichen Wheezing nicht möglich ist [Ploysongsang Y, 1988; Gavriely N, 1989b].

Verschiedene Erkrankungen der Luftwege, nicht nur das Asthma bronchiale und die chronisch obstruktive Bronchitis, können mit Wheezing einhergehen. Auch Tumore, Ödembildung oder Fremdkörperwirkung können zu einer Verengung der Luftwege führen, wodurch Wheezing entstehen kann.

<i>Einige Lungenerkrankungen und ihre assoziierten Geräusche</i>		
<b>Erkrankung</b>	<b>Lungengeräusche</b>	<b>Nebengeräusche</b>
Gesunde Lunge	Normal	Keine
Asthma	Normal, verlängertes Expirium	Wheezing, speziell expiratorisch
Chronische Bronchitis	Normal, verlängertes Expirium	Wheezing, Krepitationen, Rhonchi
Lungenemphysem	Vermindertes AG, verlängertes Expirium	Crackles
Pneumonie	Bronchialatmung	Fine Crackles
Interstitielle Lungenerkrankung	Bronchialatmung, verlängertes Expirium	Krepitationen
Bronchiektasen	Bronchialatmung	Coarse Crackles

Tabelle 2: Einige Lungenerkrankungen und ihre assoziierten Geräusche [Sankur B, 1994].

Hinsichtlich des Zusammenhangs zwischen Wheezing und dem Schweregrad der Obstruktion wurden Untersuchungen von diversen Arbeitsgruppen durchgeführt.

Marini und Mitarbeiter untersuchten 83 Patienten mit chronischer Atemwegs-obstruktion und konnten einen eindeutigen Zusammenhang zwischen Wheezing und einer mittleren bis schweren Obstruktion feststellen ( $r=0,42$ ) [Marini JJ, 1979].

Shim und Williams fanden ebenfalls einen Zusammenhang zwischen Wheezing und dem Schweregrad der Obstruktion. In ihrer Studie zeigte sich die bronchiale Obstruktion bei denjenigen Patienten, die sowohl inspiratorisches als auch expiratorisches Wheezing aufwiesen, stärker ausgeprägt als bei Patienten mit nur expiratorischem Wheezing. Zusätzlich konnte ein deutlicher Abfall des expiratorischen Spitzenflusses (PEF) mit der Zunahme von Wheezing verzeichnet werden [Shim H, 1983].

Baughman und Loudon berechneten bei Patienten mit akuter Exazerbation eines Asthma bronchiale den Anteil des Respirationszyklus in dem Wheezing auftrat. Sie fanden einen signifikanten Zusammenhang zwischen dem Prozentanteil von Wheezing pro Zeiteinheit  $T_w/T_{tot}$  (Wheezing Rate) und dem Grad der Obstruktion, definiert über den FEV1-Wert [Baughman R & Loudon RG, 1984 und 1985].

Pasterkamp und Mitarbeiter untersuchten Mitte der achtziger Jahre Lungengeräusche von Patienten mit Anstrengungs- bzw. Belastungs-Asthma. Die Wheezing Rate korrelierte hier ebenfalls mit der Obstruktion, festgelegt durch den FEV1-Wert. Wheezing konnte bei allen Patienten nachgewiesen werden, wenn der FEV1-Wert auf  $<45\%$  des Soll-Wertes fiel. Bei FEV1-Werten  $>84\%$  war kein Wheezing mehr nachweisbar [Pasterkamp H, 1985].

Wheezing wurde in verschiedenen Arbeitsgruppen auch unter Provokation untersucht [Sanchez I, 1993; Rietveld S, 1995]. Bronchiale Provokationstests mit Methacholin führte Noviski an 15 Kindern mit bronchialer Obstruktion durch. Es konnte ein Zusammenhang zwischen der Methacholinkonzentration, die einen Abfall des FEV1 um 20% verursachte, und der Methacholinkonzentration, bei welcher Wheezing erstmalig über der Trachea hörbar war, festgestellt werden. Die Konzentration bei der Wheezing ausgelöst wurde, war hier im Vergleich bei allen Kindern durchschnittlich um 52% höher [Noviski N, 1991].

### 1.3 Obstruktive Lungenerkrankungen

Asthma bronchiale ist eine multifaktorielle Erkrankung mit einer anfallsweise auftretenden reversiblen Bronchialobstruktion infolge entzündlicher und hyperreaktiver Veränderungen der Atemwege. Die Prävalenz in Deutschland liegt aktuell bei ca. 6,9% [Masoli M, 2004]. Die Definition von Asthma bronchiale ist jedoch nur eine Syndromdefinition. Ihre Ätiologie ist bis heute nicht vollständig geklärt. Erbfaktoren, Umwelteinflüsse sowie eine überschießende Immunreaktion und eine gesteigerte Sensitivität gegenüber unspezifischen Reizen kommen ursächlich in Betracht.

Aufgrund der Vielzahl von auslösenden Faktoren werden heute verschiedene Asthmaformen unterschieden: Exogen allergisches Asthma, endogenes (nicht-allergisches) Asthma, Belastungs-Asthma und Analgetika-Asthma. Die noch heute übliche Klassifizierung geht auf die historischen Untersuchungen von Rackemann zurück [Gross R, 1999].

Das Leitsymptom der Erkrankung ist die anfallsweise auftretende Atemnot, einhergehend mit Wheezing (Giemen und Brummen). Die klassische Vorstellung geht davon aus, dass Asthma primär mit expiratorischem Wheezing einhergeht [Loudon RG, 1982]. Es konnte jedoch nachgewiesen werden, dass inspiratorisches Wheezing bei Asthma ebenso häufig oder sogar häufiger vorkommen kann [Fenton FR, 1985; Bentur L, 2003].

Asthma bronchiale kann mit oder ohne Hustenreiz und Produktion eines zähglasigen Sputums einhergehen und in der Ausprägung von wenig beeinträchtigend bis hin zum „Status Asthmaticus“ reichen. In der Lungenfunktionsprüfung zeigen sich charakteristischerweise eine verminderte 1-Sekunden-Kapazität (FEV1) sowie eine Reduktion des expiratorischen Spitzenflusses (PEF = peak expiratory flow) und der maximalen expiratorischen Flussrate (MEF50). Der Nachweis einer Reversibilität nach Inhalation eines Beta2-Agonisten, mit einem Anstieg des FEV1 um mindestens 20%, spricht in Abgrenzung zu einer chronischen für das Vorliegen einer akuten Bronchialobstruktion. Bei einer ausgeprägten Obstruktion kann aufgrund der intrathorakal gefangenen Luft auch eine verminderte forcierte Vitalkapazität (FVC) bei erhöhter Residualkapazität festgestellt werden [Dantzer DR (ATS), 1986]. Die Prognose der Erkrankung ist jedoch gut; es gibt keine Hinweise auf eine Übersterblichkeit bei Asthma.

Differenzialdiagnostisch muss das Asthma bronchiale zu anderen Erkrankungen, die zu einer Atemwegsobstruktion führen können (z.B. Tumore, Fremdkörper, Asthma cardiale), abgegrenzt werden. Am schwierigsten erscheint jedoch die Abgrenzung zu der chronisch obstruktiven Lungenerkrankung (COPD).

Die chronisch obstruktive Lungenerkrankung (COPD) ist eine Erkrankung, die primär mit Atemnot, Husten und Auswurf einhergeht. Gemäß der World Health Organization (WHO) wird sie definiert über einen produktiven Husten, der an den meisten Tagen in einem Zeitraum von mindestens drei Monaten in zwei aufeinander folgenden Jahren vorhanden ist. Für die COPD liegt die geschätzte Häufigkeit in Deutschland und den USA in der 5. und 6. Lebensdekade bei 3,5% und in der 7. Dekade bei ca. 5 % und höher. Ihre Entstehung wird auf das Versagen mechanischer wie immunologischer Abwehrsysteme im Bronchialbaum zurückgeführt. Ursächlich hierfür ist eine chronisch inhalative Belastung, z.B. durch Zigarettenrauch oder Staub [Man SF, 2003]. Des Weiteren werden rezidivierende bronchopulmonale Infekte, Antikörpermangelsyndrome, Alpha-1-Proteaseinhibitormangel und genetische Faktoren als Ursachen diskutiert.

Die COPD entwickelt sich langsam über eine vorausgehende chronische nicht-obstruktive Bronchitis mit Husten und primär morgendlichem Auswurf, zu einer manifesten Atemwegsobstruktion. Klinisch leidet der Patient zunehmend unter Belastungsdyspnoe und Leistungsabfall. Auskultatorisch können ein abgeschwächtes Atemgeräusch und pathologische Nebengeräusche (wie z.B. Wheezing) festgestellt werden. Die Lungenfunktionsprüfung zeigt einen erniedrigten FEV1-Wert mit erhöhtem Atemwegswiderstand, und die Reversibilität nach Broncholyse ist deutlich eingeschränkt [Medbo A, 2007]. Als Spätkomplikationen können das obstruktive Lungenemphysem, die respiratorische Insuffizienz und das Cor pulmonale auftreten. Die COPD ist in den USA und Deutschland eine häufige Todesursache mit steigender Tendenz. Die Zehn-Jahres-Überlebensrate nach Diagnosestellung liegt bei unter 50 % [Gross R, 1999].

## 1.4 Die Lungengeräuschanalyse

Die Entstehungsmechanismen der Lungengeräusche sind seit langem von ärztlichem und wissenschaftlichem Interesse. Erste Untersuchungen hierzu gehen auf die Zeit von Hippokrates zurück. Die Auskultation anhand des Stethoskops wurde schließlich durch den französischen Mediziner Laennec zu Beginn des 19. Jahrhunderts möglich. Sie ist seit jeher eine altbewährte diagnostische Methode, die jedoch limitiert wird durch ihre Subjektivität und die technische Beschaffenheit des Stethoskops.

Der Zugang zu einer größeren Informationsbreite und einer objektiven Beurteilung wird mit der computergestützten Analyse möglich. Sie stellt ein schnelles, kostengünstiges und nicht-invasives Verfahren zur Erfassung und Speicherung von Lungengeräuschen dar. Sowohl Früherkennung als auch Verlaufsbeobachtungen von Lungenerkrankungen werden erleichtert.

Gegen Ende der zwanziger Jahre des letzten Jahrhunderts wurde der Grundstein für die Lungengeräuschanalyse gelegt [Hannon RR, 1929]. Von da an begannen Wissenschaftler die Geräusche mit verschiedenen Medien aufzuzeichnen. Die ersten Methoden basierten auf einer mikrophonegekoppelten Aufnahme, an die Filter, Bandoszilloskop und ein Aufzeichnungsgerät mit graphischer Darstellung angeschlossen waren [Loudon RG, 1984]. Bis in die Neunziger Jahre hinein war es üblich, die Geräusch- und Flowsignale (Atemflusssignale) auf einer analogen, magnetischen Aufnahmekassette zu speichern [Earis JE, 2000]. Mit der computergestützten Lungengeräuschanalyse, zu deren Einführung Forgacs einen wichtigen Beitrag leistete, wurde die digitale Signalgewinnung und Verarbeitung möglich [Forgacs P, 1967 und 1969].

Auf dem medizintechnischen Markt sind bis heute noch keine Geräte zur Lungengeräuschanalyse erhältlich. Das mag daran liegen, dass es z.B. im Vergleich zu Herzgeräuschen schwieriger ist, eine adäquate Aufnahme- und Analysemethode für Lungengeräusche zu entwickeln. Wie bei anderen diagnostischen Methoden auch, basiert eine objektive Beurteilung von Lungengeräuschen darauf, reproduzierbare und analysierbare Parameter zu liefern. Das grundlegende Problem bei Lungengeräuschen besteht in der niedrigen Geräuschintensität mit einem komplexen Muster [Druzgalski CK, 1980]. Veränderungen im Atemverhalten, wie Atemfluss und Volumen, nehmen einen Einfluss auf die Lungengeräusche und ihre Sig-

nale [Kraman SS, 1986; Schreur H, 1992]. Speziell die Atemflussrate stellt einen wichtigen Faktor dar und sollte konstant gehalten werden. In früheren Studien konnte gezeigt werden, dass Lungengeräuschk спектра oberhalb einer Atemflussrate von 1,0 l/s relativ stabil sind, wohingegen sie unterhalb dieser Grenze eine deutliche Abhängigkeit hierzu aufweisen [Charbonneau G, 1983; Kraman SS, 1986]. Auch ein Zusammenhang zwischen Wheezing und Atemfluss konnte nachgewiesen werden. Es wurde eine kritische Flussrate ( $< 0,1$  l/s) beschrieben, unterhalb derer Wheezing nicht auftritt, unabhängig von dem Vorhandensein einer schweren Atemwegsobstruktion [Forgacs P, 1978a; Grotberg JB, 1980; Fenton FR, 1985]. Seit der Einführung elektronischer Aufnahmesysteme und der Entwicklung der Computeranalyse im Laufe der letzten Jahrzehnte haben sich verschiedene Arbeitsgruppen mit der Methodik der Lungengeräuschanalyse beschäftigt. Die „International Lung Sounds Association“ (ILSA), als Zusammenschluss aus Medizinern, Ingenieuren und Wissenschaftlern, hält ein Forum bereit, in dem die neuesten Erkenntnisse in der Erforschung von Lungengeräuschen veröffentlicht werden. Arbeitsgruppen um Pasterkamp, Gavriely und Sovijärvi haben entscheidend zu der Weiterentwicklung auf dem Gebiet der Lungengeräuschanalyse beigetragen.

Die meisten Studien in diesem Bereich folgen dabei dem Muster: Aufnahme, Datenverarbeitung und anschließende Geräuscherkennung. Es existieren Bemühungen um Standardisierung. Das CORSA-Projekt veröffentlichte im Jahr 2000 einen Katalog über Mindestanforderungen bei der digitalen Geräuschanalyse. Trotzdem variieren Technik und Ausstattung erheblich in den diversen Arbeitsgruppen. Es gibt keinen einheitlichen Standard hinsichtlich der Auswertung, was eine Vergleichbarkeit der Daten deutlich erschwert [Gavriely N, 1995; Dalmay F, 1995; Earis JE, 2000].

In den letzten Jahrzehnten haben sich verschiedenen Arbeitsgruppen auch mit der automatischen Mustererkennung von Lungengeräuschen durch Computeralgorithmen befasst. Sie basiert auf der Erkennung von spektralen Veränderungen im Vergleich zu normalen Lungengeräuschen [Baughman RP, 1984; Fenton R, 1985; Shabtai-Musih Y, 1992].

Die Entwicklung eines computergestützten Algorithmus setzt jedoch die reproduzierbare und vergleichbare Beurteilbarkeit durch den Untersucher voraus. Einige Arbeiten haben sich bereits mit dem Vergleich von Untersucherübereinstimmung

und automatischer Mustererkennung durch einen Algorithmus beschäftigt [Pasterkamp H, 1987a; Mussell MJ, 1992; Rietveld S, 1999; Elphick HE, 2004]. Hier zeigten sich die Ergebnisse jedoch uneinheitlich.

Die Vergleichbarkeit der existierenden Studien wird dadurch erschwert, dass bis zum heutigen Tag ein Goldstandard für die Auswertung und Beurteilung von Lungengeräuschen fehlt. Aufgrund dessen bedarf es weiterer Untersuchungen, um eine Standardisierung voranzutreiben.

## 2 Zielsetzung und Fragestellungen der Arbeit

Das Ziel der vorliegenden Arbeit bestand darin, anhand einer standardisierten Analyseverfahren zu untersuchen, ob klinisch erfahrene Untersucher audio-visuell präsentierte Lungengeräuschzeichnungen übereinstimmend erkennen und beurteilen können.

Fünf Ärzten wurde die Aufgabe gestellt, das Nebengeräusch „Wheezing“ anhand einer Skaleneinteilung von 1-7 wiederholt zu bewerten. Es sollte eine Beurteilung anhand der **Wheezing Rate** und zusätzlich eine Einschätzung der **Klinischen Relevanz** vorgenommen werden.

Anschließend wurden folgende Fragestellungen untersucht:

- 1) Gibt es eine signifikante Korrelation (Reliabilität) unter den Untersuchern bei der Beurteilung von Wheezing anhand der Aspekte:
  - *Subjektive Wheezing Rate und*
  - *Klinische Relevanz?*
  
- 2) Gibt es eine signifikante Zuverlässigkeit (Re-Test Reliabilität) unter den Untersuchern bei wiederholter Beurteilung (zwei Durchgänge) von Wheezing anhand der Aspekte:
  - *Subjektive Wheezing Rate und*
  - *Klinische Relevanz?*
  
- 3) Gibt es einen signifikanten Zusammenhang zwischen den Beurteilungen der subjektiven Wheezing Rate und der Klinischen Relevanz?
  
- 4) Welche Empfehlungen ergeben sich aus den Untersuchungen hinsichtlich einer standardisierten Bewertung?

### 3 Patienten und Methoden

#### 3.1 Das Untersuchungskollektiv

Das Untersuchungskollektiv umfasste insgesamt 98 Personen, die in einem Zeitraum von einem Jahr von der Versuchsleiterin gemessen wurden. Hierunter waren 73 Patienten und 25 Probanden. Dieses Kollektiv setzte sich zusammen aus stationären und ambulanten Patienten der internistisch-pneumologischen Abteilung des Universitätsklinikums Marburg. In die Studie wurden ausschließlich Patienten mit Asthma bronchiale und chronisch obstruktiver Bronchitis eingeschlossen, da diese mit hoher Wahrscheinlichkeit Wheezing aufweisen würden. Daneben wurde ein Kollektiv an lungengesunden Probanden gemessen. Beide Gruppen befanden sich im Alter zwischen 15 und 82 Jahren. Aufgenommen wurden sowohl Männer und Frauen, als auch Raucher und Nichtraucher. Alle Patienten und Probanden erhielten vor der Messung eine Lungenfunktionsanalyse und erteilten vor Beginn der Studie ihr schriftliches Einverständnis.

In Tabelle 3 sind die anthropometrischen und klinischen Parameter des ausgewählten Untersuchungskollektivs (24 Patienten mit Asthma bronchiale, 18 mit COPD und 25 lungengesunde Probanden), aus dem die 67 Lungengeräuschaufzeichnungen für das zu beurteilende „Testset“ entnommen wurden, aufgeführt.

KLINISCHE PARAMETER DES UNTERSUCHUNGS- KOLLEKTIVS IM TESTSET:			
	ASTHMA BRONCHIALE	COPD	PROBANDEN
Anzahl Patienten	24	18	25
Alter (MW + STD)	41 ± 18	64 ± 13	34 ± 12
Geschlecht m/w	13 / 10	14 / 5	16 / 9
FEV1% (MW + STD)	83 ± 21	56 ± 19	106 ± 17

Tabelle 3: Klinische Parameter des Untersuchungskollektivs im Testset.

Folgende Ausschlusskriterien wurden vorab für das Untersuchungskollektiv festgelegt:

- andere Lungenerkrankungen, wie Lungenemphysem oder Pneumonie,
- systemische Erkrankungen, wie Sarkoidose oder Lymphome,
- neoplastische Geschehen in den Atemwegen oder der Lunge,
- sonstige metastasierende, intra- oder extrathorakal gelegene Tumore,
- inspiratorischer Stridor der extrathorakalen Luftwege aufgrund von Fremdkörperaspiration, Glottisödem oder größeren Schilddrüsenstrumata.

## 3.2 Methoden

### 3.2.1 Standardisierte Atmung

In der vorliegenden Studie erfolgten die Messungen unter standardisierter Atmung, da die Atemflussrate einen wichtigen Einfluss auf das Ergebnis der Lungen-geräuschanalyse und das Auftreten von Wheezing hat (vgl. Abschnitt 1.4).

Der Patient atmete durch den Mund in den Pneumotachographen, während eine Nasenklammer die Nasenatmung unterband. Es wurde versucht, die individuelle Fluss-Volumen-Kurve des Patienten innerhalb der Grenzwerte von 1,0 bis 1,3 l/s zu halten. Damit dies gewährleistet werden konnte, erhielt jeder Patient im Vorfeld genaue Instruktionen und machte vorab einige Übungsdurchläufe.

Der Atemfluss wurde zeitgleich zur Geräuschemessung registriert und für Versuchsleiterin und Patienten gleichermaßen sichtbar auf den Computermonitor übertragen. Auf diese Weise konnte in In- und Expiration unterschieden werden und der Patient konnte unter visueller Kontrolle versuchen, seine maximalen Atemexkursionen so konstant wie möglich zu halten. Die Messung umfasste ca. 8 bis 10 Atemzyklen bei normaler ruhiger Atmung.

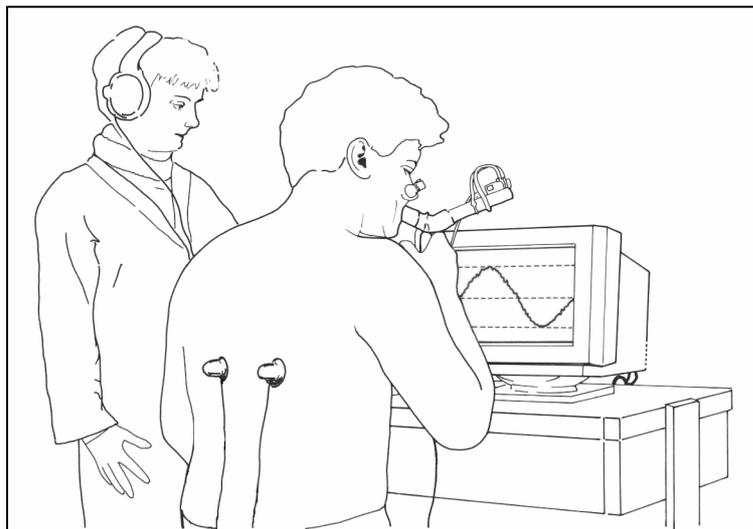


Abb. 2: Messung eines Patienten bei standardisierter Atmung unter Online-Kontrolle.

### 3.2.2 Das Messgerät

In der „Arbeitsgruppe Lung-Sound“ der Philipps-Universität Marburg (Abteilung Pneumologie/ Innere Medizin) wurde speziell für die Lungengeräuschanalyse ein Gerät nach den aktuellen CORSA-Standards entwickelt [Sovijärvi ARA, 2000a+b]. Dieses Gerät ist ein Prototyp und existiert in seiner derzeitigen Form und Ausführung nur im Universitätsklinikum Marburg. Es zeichnet Lungengeräusche auf, die anhand von luftgekoppelten Mikrofonen detektiert und im Anschluss digitalisiert werden.

Für diese Studie wurden fünf luftgekoppelte Elektret-Mikrophone (Sony ECM77), sowie Aluminium-Koppler verwendet. Koppler und Mikrophone wurden auf ihre Vergleichbarkeit hin im Fachlaboratorium der Physikalisch-Technischen Bundesanstalt in Braunschweig geprüft (Zertifikat Nr. 1385, PTB 1997). Die Wahl der Positionen für die einzelnen Mikrophone basierte auf Vorerfahrungen aus vorangegangenen Studien der „Arbeitsgruppe Lung-Sound“ [Gross V, 2000 u. 2003; Sulzer J, 2000; Dittmar A, 2002].

Die Positionen wurden wie folgt festgelegt:

- a) 3. Interkostalraum paravertebral links und rechts,
- b) 7. Interkostalraum auf der Linea scapularis links und rechts,
- c) paratracheal neben dem Cartilago cricoidea links oder rechts.

Zum Ausschluss systematischer Fehler wurde die Zuteilung der jeweils fünf Mikrophone zu den oben genannten Positionen bei jeder Messung randomisiert. Die Signale wurden analog verstärkt und gefiltert. Es wurde ein Anti-Aliasing-Filter zur Tiefpassfilterung oberhalb von 2.200 Hz und ein Hochpassfilter unterhalb von 60 Hz verwendet (TP-MF-01-48-B und HP-MF-01-48-B, GEPAmbH-München, 48 dB/Oktave Filterdämpfung). Im Anschluss an die Filterung und Verstärkung wurden die Daten mittels eines A/D-Wandlers (SORCUS M4/486) mit einer Abtastrate von 5.512 Hz digitalisiert und auf der Festplatte des portablen Messrechners abgespeichert. Die Auswertung und Verarbeitung der aufgezeichneten Signale wurde über eine unter Matlab 5.3 erstellte Programmroutine mit Hilfe der Fast-Fourier-Transformation (FFT) vorgenommen.

### 3.2.3 Die audio-visuelle Bewertung der Lungengeräuschaufzeichnungen

Die qualitative Bewertung der Lungengeräuschaufzeichnungen sollte im Sinne einer audio-visuellen Beurteilung erfolgen und somit sowohl eine visuelle als auch akustische Beurteilung der Geräusche beinhalten. Die Länge eines vorgespielten Frequenzbandes umfasste vier Atemzüge und entsprach ungefähr der geschätzten Zeit, die die Auskultation einer Position auf dem Thorax im klinischen Alltag einnehmen würde.

Für diese Aufgabe wurden fünf Ärzte aus dem medizinischen Personal der Abteilung Pneumologie des Universitätsklinikums Marburg als Untersucher (Rater) ausgewählt. Alle Fünf waren mit dem Gebiet der Pneumologie im klinischen Alltag gut vertraut und wiesen eine gleichwertige klinische Erfahrung auf. Als „klinische Erfahrung“ wurde der Nachweis von mindestens einem Jahr Tätigkeit auf dem Gebiet der Inneren Medizin mit dem Schwerpunkt Pneumologie festgelegt.

Es sollte das Nebengeräusch „Wheezing“ anhand von Lungengeräuschaufzeichnungen beurteilt werden. Für diese Beurteilung wurde die Wheezing Rate gewählt, da sie sich bereits in vorangegangenen Studien als gut einsetzbarer Parameter zur objektiven Beurteilung des Obstruktionsgrades bewährt hat [Baughman R 1982; Pasterkamp H, 1985].

Die Untersucher wurden zudem gebeten, den Gesamteindruck der Geräuschaufzeichnungen mit einer eigenen Einschätzung der Klinischen Relevanz zu bewerten. Dies sollte im Sinne einer subjektiven Beurteilung erfolgen. Dadurch wurde es möglich, die klinische Erfahrung des Einzelnen sowie einen subjektiv wahrgenommenen Zusammenhang zwischen beiden Parametern abzubilden.

Den Untersuchern waren die Kriterien zur Einschätzung des Schweregrades bekannt. Sie wussten, dass z.B. ein gleichzeitiges Auftreten von in- und expiratorischem Wheezing, wie auch eine größere prozentuale Häufigkeit von Wheezing pro Zeiteinheit, auf eine stärkere Ausprägung der Obstruktion hinweisen.

Die Vorgaben für die Untersucher waren folgende:

- 1) Subjektive Beurteilung des Vorhandenseins und der Dauer von „Wheezing“ anhand der **Wheezing Rate**. (Die Wheezing Rate entspricht der Dauer des Auftretens von Wheezing pro Zeiteinheit, hier im Zeitfenster von vier Atemzügen).
- 2) Subjektive Beurteilung der **Klinischen Relevanz**. (Die Einschätzung der Klinischen Relevanz soll den Gesamteindruck der Atem- und Nebengeräusche in der jeweiligen Lungengeräuschaufzeichnung widerspiegeln.)

Für die Beurteilung der Geräuschaufzeichnungen wurde sowohl für die Wheezing Rate, als auch für die Klinische Relevanz eine 7er-Skala zur Graduierung des Schweregrades gewählt. Die Unterteilung in sieben Unterpunkte entspricht einer Standardskaleneinteilung und lässt eine gute Differenzierung zu.

Es wurde jeweils nur die Definition für die Eckpunkte der 7er-Skala vorgegeben. Die Zwischenwerte sollten anhand subjektiver Einschätzung vergeben werden.

Folgende Einteilung wurde für die Wheezing Rate und die Klinische Relevanz festgelegt:

1 = kein Wheezing vorhanden.

7= sehr hohe Wheezing Rate.

1 = keine Klinische Relevanz.

7 = sehr hohe Klinische Relevanz.

### 3.2.4 Das Lernset

Als Vorbereitung auf die bevorstehende Beurteilung wurde vorab von drei unabhängigen Experten ein Übungs- bzw. Lernset aus 50 Lungengeräuschaufzeichnungen erstellt. Die Untersucher sollten auf diese Weise mit der Beurteilung von Geräuschaufzeichnungen vertraut gemacht werden.

Diese 50 Aufzeichnungen wurden nach denselben Kriterien, jedoch völlig unabhängig vom Testset, zusammengestellt. Die Lungengeräuschaufzeichnungen stammten von Patienten mit Asthma bronchiale, COPD sowie einem lungengesunden Probandenkollektiv und wurden der Multimedia Database „Marburg Respiratory Sounds“ (MARS) entnommen. Die Präsentation erfolgte randomisiert und ohne Hintergrundinformationen, wie ID, Geschlecht, Alter, Erkrankung und Mikrofonposition. Den Untersuchern wurde eine Beurteilung von „Wheezing“ anhand der Parameter Wheezing Rate und Klinische Relevanz beispielhaft dargestellt, wie sie anschließend im Testset von ihnen selbst durchgeführt werden sollte.

### 3.2.5 Das Testset

Das Testset umfasste 67 Lungengeräuschaufzeichnungen, die aus dem Gesamtkollektiv von 98 gemessenen Patienten und Probanden zusammengestellt wurden (vgl. Abschnitt 3.1). Die Zahl 67 wurde gewählt, um den zeitlichen Aufwand für die Beurteilung auf ca. 30 min begrenzt zu halten.

Insgesamt wurden 42 Aufzeichnungen aus der Patienten-Gruppe mit Asthma bronchiale und COPD, und 25 aus der Gruppe der lungengesunden Probanden verwendet. Alle 42 Geräuschaufzeichnungen aus der Patienten-Gruppe wiesen Phasen mit Wheezing auf.

Die Parameter der Lungengeräuschaufzeichnungen im Testset sind untenstehend in Tabelle 4 einzusehen.

PARAMETER DER GERÄUSCHAUFZEICHNUNGEN IM TESTSET:			
Kollektiv	Asthma bronchiale	COPD	Probanden
Anzahl Aufzeichnungen	24	18	25
Wheezing	+	+	-

*Tabelle 4: Parameter der Geräuschaufzeichnungen im Testset.*

Den fünf Untersuchern wurde das Testset getrennt voneinander vorgelegt. Das Abspielen der Geräuschaufzeichnungen erfolgte nacheinander am Computer und konnte von den Untersuchern selbst gesteuert werden.

Nach dem ersten Durchlauf des Testsets wurde jedem Untersucher in einem zweiten Durchgang dieselbe Anzahl Aufzeichnungen in neuer randomisierter Reihenfolge noch einmal vorgespielt. Für den zeitlichen Abstand zwischen beiden Durchgängen wurde als Vorgabe mindestens ein Tag Pause festgelegt. Weitere Vorgaben wurden nicht gemacht. Anschließend konnte somit beurteilt werden, ob bei Wiederholung des Testsets unter gleichen Bedingungen das Ergebnis stabil bleibt und eine Zuverlässigkeit der Untersucher vorliegt.

Nachfolgend sind zur Verdeutlichung der Methode zwei Beispiele (vgl. Abb. 3 u. 4) aus dem ausgewählten Kollektiv an Lungengeräuschaufzeichnungen dargestellt. Es wurde jeweils eine Aufzeichnung mit ausgeprägtem und eine ohne Wheezing ausgewählt.

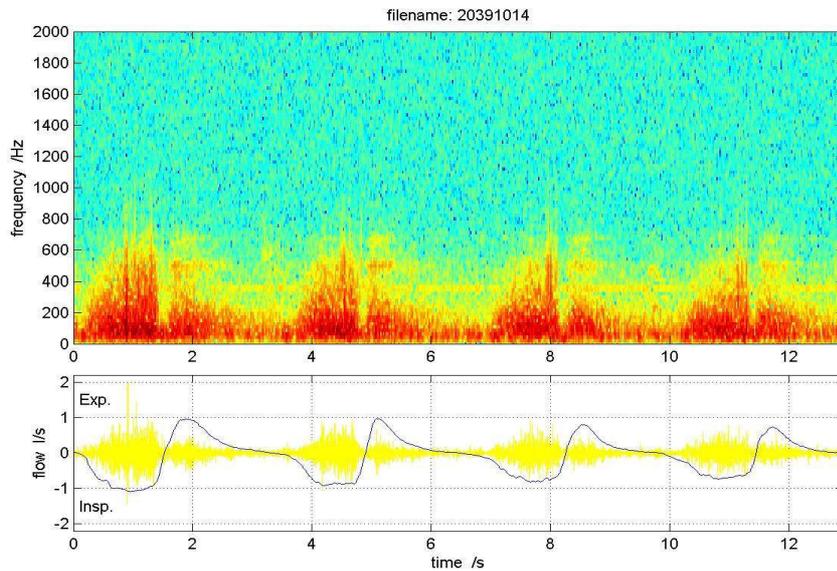


Abb. 3: Beispiel einer Lungengeräuschaufzeichnung eines lungengesunden Probanden (ohne Wheezing) inklusive Atemfluss-Kurve.

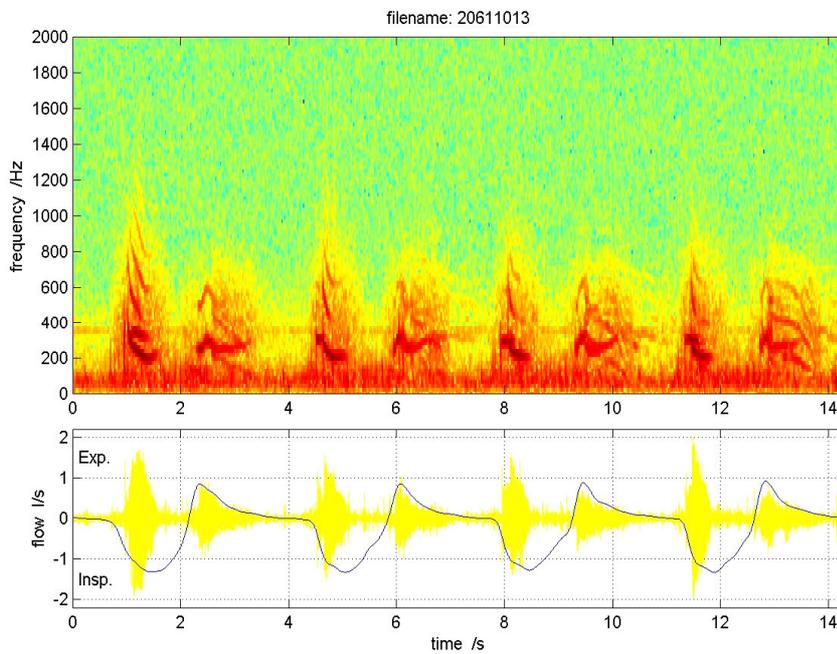


Abb. 4: Beispiel einer Lungengeräuschaufzeichnung eines Patienten mit obstruktiver Lungenerkrankung (mit deutlichem in- und expiratorischen Wheezing).

### 3.2.6 Statistische Verfahren der Auswertung

Als deskriptive Maße wurden Mittelwert, Standardabweichung, Minimal- und Maximalwerte bestimmt.

Wenn die standardisierte Schiefe und Kurtosis einer Variablen im Bereich  $\pm 2$  lag, wurde von Normalverteilung ausgegangen.

Da die Beurteilungen (Ratings) der einzelnen Untersucher (Rater) nicht normalverteilt sind, wurde als Maß für den Zusammenhang zweier Messwertreihen die Rangkorrelation nach Spearman (Spearman's Rho) berechnet und auf Signifikanz geprüft. Dies fand Anwendung bei beiden Beurteilungskategorien, Wheezing Rate und Klinische Relevanz.

Für die Korrelation der Beurteilungen aus subjektiver Wheezing Rate und Klinischer Relevanz wurde die Produkt-Moment-Korrelation verwendet.

## 4 Ergebnisse

### 4.1 Statistische Datenanalyse

#### 4.1.1 Reliabilität der Untersucher

##### Wheezing Rate

Tabelle 5 zeigt die Auswertung der Korrelationen (Reliabilität) der 5 Rater hinsichtlich des Beurteilungsaspektes Wheezing Rate bei 67 Lungengeräuschaufzeichnungen.

Die Korrelation der Beurteilungswerte kann als Maß der Stärke des linearen Zusammenhangs zwischen den Ratern angesehen werden. Für den Aspekt Wheezing Rate konnte eine sehr hohe Korrelation (Reliabilität) unter den 5 Ratern gefunden werden. Alle errechneten Korrelationen waren signifikant ( $p < 0,01$ ) (Einklassifizierung nach der Tabelle für den Spearman Rho Korrelationstest).

Die Korrelationskoeffizienten der einzelnen Untersucher reichten von  $r = 0,868$  bis  $r = 0,952$ .

<b>Wheezing Rate</b>					
Spearman-Rho Korrelation					
	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5
Rater 1		0,868	0,884	0,923	0,898
Rater 2			0,949	0,878	0,943
Rater 3				0,878	0,952
Rater 4					0,905
Rater 5					

Tabelle 5: Korrelation aller 5 Untersucher nach Spearman-Rho für die Wheezing Rate.

Zur Auswertung der Abweichungen (Differenzen) wurden die Beurteilungsreihen aller Rater für jede Lungengeräuschaufzeichnung einander paarweise gegenübergestellt. Hieraus wurden die absoluten Differenzwerte ermittelt und zur Verdeutlichung Graphiken erstellt. Ein Beispiel (Rater 3 versus Rater 4) ist unter Abb. 5 exemplarisch für den Beurteilungsaspekt Wheezing Rate aufgeführt.

Die graphischen Gegenüberstellungen aller Rater sind im Anhang unter 8.4. (Differenzwerte) zu finden.

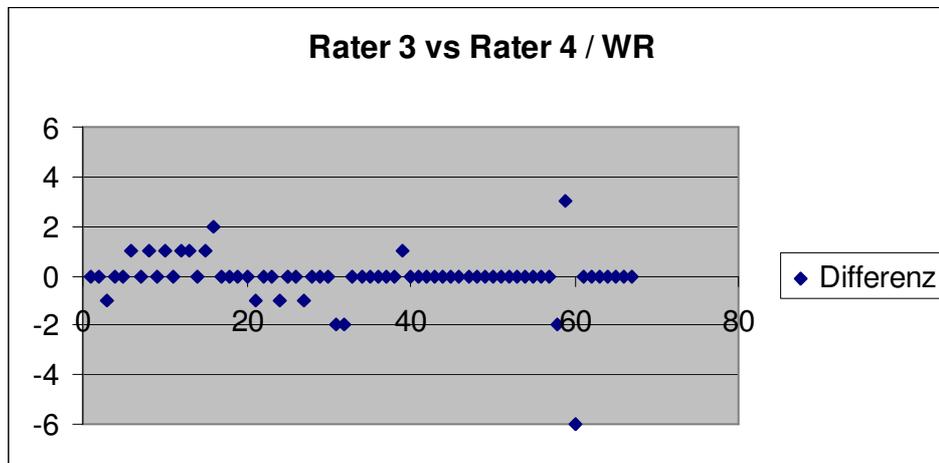


Abb.5: Graphische Darstellung der Differenzwerte Rater 3 vs. Rater 4/ Wheezing Rate bei insgesamt 67 Beurteilungen.

Tabelle 6 zeigt eine Übersicht über das Ausmaß der Abweichungen, dargestellt als Mittelwert und Standardabweichung, bzw. Anteil an der Gesamtmenge in Prozent. Die geringste Abweichung bei dem Aspekt Wheezing Rate war eine Abweichung um 1, die maximale um 6 Skalenwerte. Eine Abweichung um 0 ist gleichzusetzen mit Übereinstimmung, und wurde mit einem Anteil von  $70,1 \pm 3,5\%$  an der Gesamtmenge am häufigsten gefunden. Die Abweichung um 1 Skalenwert lag bei einem Anteil von  $25,4 \pm 4,6\%$  und um 2 bei  $3,1 \pm 2,1\%$ , mit einer stark sinkenden Häufigkeit für höhere Abweichungen. Insgesamt kann festgestellt werden, dass eine Abweichung um  $\leq 1$  Wert zu  $95,5\%$  vorlag.

Abschließend zeigt Abb. 6 die Häufigkeitsverteilung der Abweichungen Rater 1 zu den Ratern 2-5 als graphische Darstellung.

**Abweichungstabelle Wheezing Rate (n = 67)**

Abweichung (in Skalenwerten)	0	1	2	3	4	5	6
Rater 1 zu 2	43	21	2	0	0	0	1
Rater 1 zu 3	49	16	1	0	0	0	1
Rater 1 zu 4	50	14	2	0	1	0	0
Rater 1 zu 5	44	21	1	0	1	0	0
Rater 2 zu 3	48	19	0	0	0	0	0
Rater 2 zu 4	45	18	2	1	0	0	1
Rater 2 zu 5	48	14	5	0	0	0	0
Rater 3 zu 4	50	11	4	1	0	0	1
Rater 3 zu 5	46	19	2	0	0	0	0
Rater 4 zu 5	47	17	2	0	1	0	0
Mittelwert	47,0	17,0	2,1	0,2	0,3	0,0	0,4
±STD	±2,3	±3,1	±1,4	±0,4	±0,5	±0,0	±0,5
Anteil in %	70,1	25,4	3,1	0,3	0,9	0,0	0,6
±STD in %	±3,5	±4,6	±2,1	±0,6	±0,7	±0,0	±0,7

Tabelle 6: Tabellarische Darstellung der Abweichungen mit Mittelwert und Standardabweichung für die Wheezing Rate.

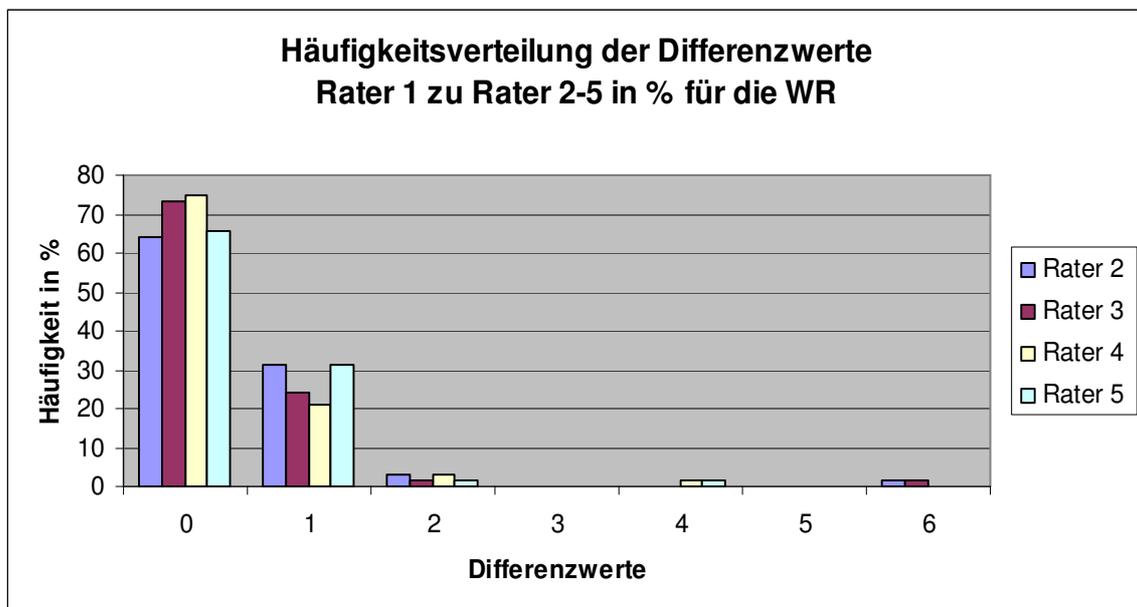


Abb. 6: Häufigkeitsverteilung der Differenzwerte Rater 1 zu den Ratern 2-5 in % für die Wheezing Rate.

## Klinische Relevanz

Tabelle 7 zeigt die Auswertung des Beurteilungsaspektes Klinische Relevanz. Die Auswertung ergab auch für diesen Beurteilungsaspekt sehr hohe Korrelationen (Reliabilität) unter den 5 Ratern. Die errechneten Korrelationen für die Klinische Relevanz waren ebenfalls signifikant ( $p < 0,01$ ) (Einordnung nach der Tabelle für den Spearman Rho Korrelationstest). Die Korrelationskoeffizienten reichten von  $r = 0,882$  bis  $r = 0,935$ .

<b>Klinische Relevanz</b>					
Spearman-Rho Korrelation					
	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5
Rater 1		0,882	0,889	0,893	0,904
Rater 2			0,924	0,933	0,911
Rater 3				0,935	0,919
Rater 4					0,916
Rater 5					

Tabelle 7: Korrelation aller 5 Untersucher nach Spearman-Rho für die Klinische Relevanz.

Die Beurteilungsreihen aller Rater für die Klinische Relevanz wurden einander paarweise gegenübergestellt. Daraus wurden die absoluten Differenzwerte ermittelt und graphisch dargestellt. Ein Beispiel für die paarweise Gegenüberstellung (Rater 3 versus Rater 4) bei der Klinischen Relevanz wurde unter Abb. 7 aufgeführt. Die graphischen Gegenüberstellungen aller Rater sind im Anhang unter 8.4. (Differenzwerte) zu finden.

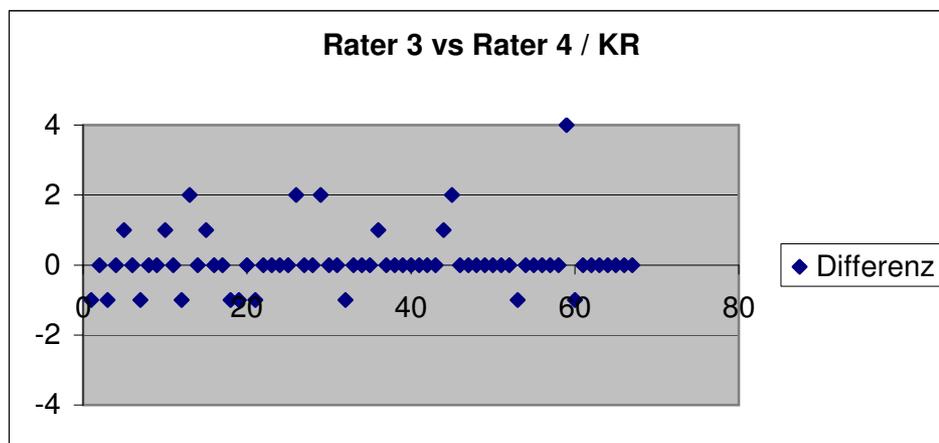


Abb. 7: Graphische Darstellung der Differenzwerte Rater 3 vs. Rater 4/  
Klinische Relevanz bei insgesamt 67 Beurteilungen.

In Tabelle 8 folgt eine Darstellung der Abweichungen für die Klinische Relevanz. Die kleinste Abweichung lag bei 1, die maximale bei 4.

Auch hier wurde die Abweichung um 0 Skalenwerte (= Übereinstimmung) am häufigsten gefunden (Anteil an der Gesamtmenge  $61,5 \pm 4,6\%$ ), gefolgt von der Abweichung um 1 Wert (Anteil von  $28,2 \pm 3,9\%$ ). Zusammengefasst betrug der Anteil der Abweichungen um  $\leq 1$  Wert insgesamt  $89,7\%$ . Abweichungen um 2 Werte lagen bei  $8,1 \pm 4,0\%$ , Abweichungen um 3 oder 4 Skalenwerte kamen deutlich seltener vor.

Die Häufigkeitsverteilung der Differenzwerte am Beispiel Rater 1 zu Rater 2-5 für die Klinische Relevanz wurde nachfolgend graphisch dargestellt (vgl. Abb. 8).

<b>Abweichungstabelle Klinische Relevanz (n = 67)</b>							
<b>Abweichung (in Skalenwerten)</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
<b>Rater 1 zu 2</b>	38	16	12	1	0	0	0
<b>Rater 1 zu 3</b>	40	22	2	3	0	0	0
<b>Rater 1 zu 4</b>	42	17	6	1	1	0	0
<b>Rater 1 zu 5</b>	37	23	6	1	0	0	0
<b>Rater 2 zu 3</b>	45	19	2	1	0	0	0
<b>Rater 2 zu 4</b>	38	22	6	1	0	0	0
<b>Rater 2 zu 5</b>	43	17	5	2	0	0	0
<b>Rater 3 zu 4</b>	47	15	4	0	1	0	0
<b>Rater 3 zu 5</b>	40	20	6	1	0	0	0
<b>Rater 4 zu 5</b>	42	18	5	2	0	0	0
<b>Mittelwert</b>	<b>41,2</b>	<b>18,9</b>	<b>5,4</b>	<b>1,3</b>	<b>0,2</b>	<b>0,0</b>	<b>0,0</b>
<b>±STD</b>	<b>±3,1</b>	<b>±2,6</b>	<b>±2,7</b>	<b>±0,8</b>	<b>±0,4</b>	<b>±0,0</b>	<b>±0,0</b>
<b>Anteil in %</b>	<b>61,5</b>	<b>28,2</b>	<b>8,1</b>	<b>1,9</b>	<b>0,3</b>	<b>0,0</b>	<b>0,0</b>
<b>±STD in %</b>	<b>±4,6</b>	<b>±3,9</b>	<b>±4,0</b>	<b>±1,2</b>	<b>±0,6</b>	<b>±0,0</b>	<b>±0,0</b>

*Tabelle 8: Tabellarische Darstellung der Abweichungen mit Mittelwert und Standardabweichung für die Klinische Relevanz.*

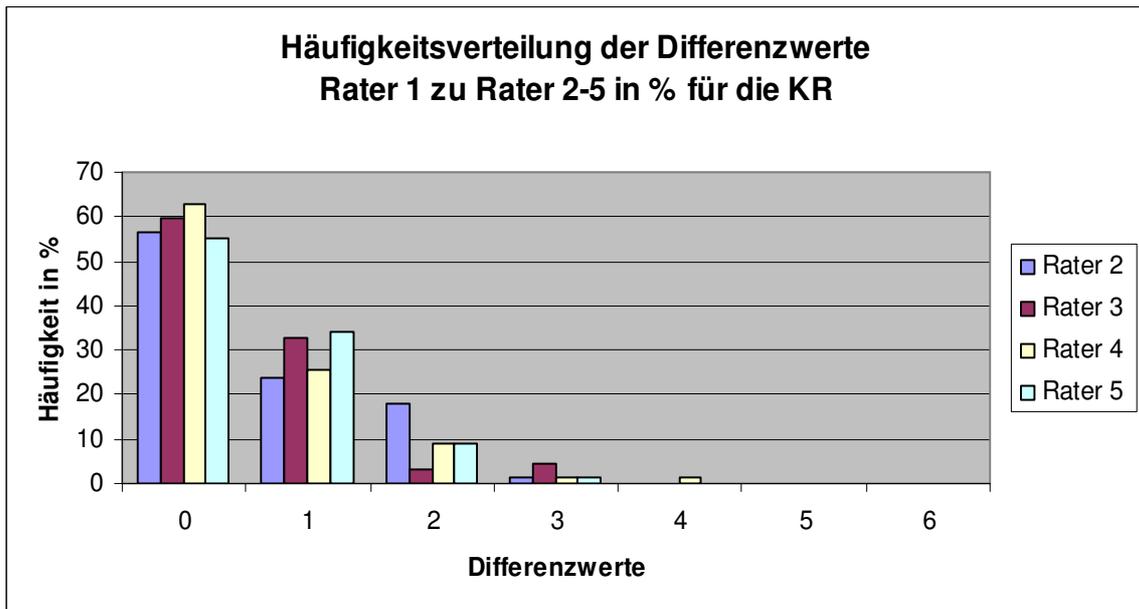


Abb. 8: Häufigkeitsverteilung der Differenzwerte Rater 1 zu den Ratern 2-5 in % für die Klinische Relevanz.

#### 4.1.2 Re-Test Reliabilität der Untersucher

##### Wheezing Rate

Die Ergebnisse des Re-Tests zeigen, dass die Beurteilung über beide Durchgänge stabil blieb und somit bezüglich des Beurteilungsaspektes Wheezing Rate eine hohe Zuverlässigkeit vorlag.

Die Reliabilitätskoeffizienten der 5 Rater reichten bei der Wheezing Rate von  $r=0,857$  bis  $r=0,976$  und waren signifikant ( $p<0,01$ ). Die Reliabilitätskoeffizienten sind in Tabelle 9 aufgeführt.

RE-TEST RELIABILITÄT	
	<i>Wheezing Rate</i>
Rater 1	0,857
Rater 2	0,976
Rater 3	0,960
Rater 4	0,910
Rater 5	0,955

Tabelle 9: Re-Test Reliabilität aller 5 Untersucher durch den Vergleich 1. und 2. Durchgang für die Wheezing Rate.

Für den Re-Test erfolgte ebenfalls eine paarweise Gegenüberstellung der Beurteilungen aller Rater mit Ermittlung der Differenzwerte. Abb. 9 zeigt das Beispiel Rater 3 im Vergleich 1. und 2. Durchgang für die Wheezing Rate.

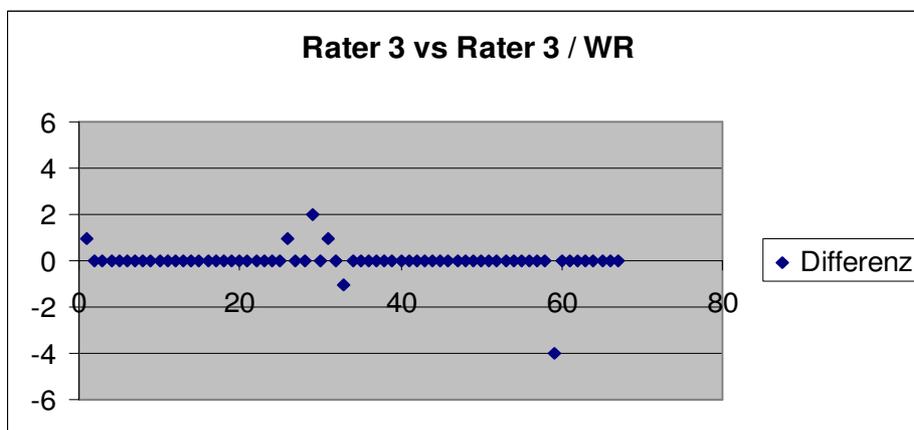


Abb. 9: Graphische Darstellung der Differenzwerte Rater 3 vs. Rater 3 (2. Dg)/ Wheezing Rate bei insgesamt 67 Beurteilungen.

Eine tabellarische Darstellung der Abweichungen für die Wheezing Rate im Re-Test zeigt Tabelle 10. Die Abweichung um 0 Skalenwerte (= Übereinstimmung) kam mit einem Anteil an der Gesamtmenge von  $74,0 \pm 10,3\%$  am häufigsten vor. Für die Abweichung um 1 Wert lag der Anteil bei  $22,7 \pm 10,1\%$ , um 2 Werte bei  $2,1 \pm 0,7\%$ . Größere Abweichungen lagen bei  $0,3-0,6\%$ . Abweichungen um mehr als 5 Skalenwerte konnten für die Wheezing Rate im Re-Test nicht gefunden werden. Deutlich wird, dass sich mit  $96,7\%$  die Häufigkeit der Abweichungen auf  $\leq 1$  Skalenwert konzentriert.

<b>Abweichungstabelle Wheezing Rate (n = 67) im Re-Test</b>							
<b>Abweichung (in Skalenwerten)</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
Rater 1 zu 1 (2.Dg)	43	20	2	1	0	1	0
Rater 2 zu 2 (2.Dg)	51	14	2	0	0	0	0
Rater 3 zu 3 (2.Dg)	61	4	1	0	1	0	0
Rater 4 zu 4 (2.Dg)	51	14	1	0	1	0	0
Rater 5 zu 5 (2.Dg)	42	24	1	0	0	0	0
<b>Mittelwert</b>	<b>49,6</b>	<b>15,2</b>	<b>1,4</b>	<b>0,2</b>	<b>0,4</b>	<b>0,2</b>	<b>0,0</b>
<b>±STD</b>	<b>±6,9</b>	<b>±6,8</b>	<b>±0,5</b>	<b>±0,4</b>	<b>±0,5</b>	<b>±0,4</b>	<b>±0,0</b>
<b>Anteil in %</b>	<b>74,0</b>	<b>22,7</b>	<b>2,1</b>	<b>0,3</b>	<b>0,6</b>	<b>0,3</b>	<b>0,0</b>
<b>±STD in %</b>	<b>±10,3</b>	<b>±10,1</b>	<b>±0,7</b>	<b>±0,6</b>	<b>±0,7</b>	<b>±0,6</b>	<b>±0,0</b>

Tabelle 10: Tabellarische Darstellung der Abweichungen für die Wheezing Rate im Re-Test.

## Klinische Relevanz

Die Reliabilitätskoeffizienten der 5 Rater im Re-Test können in Tabelle 11 eingesehen werden. Sie reichten von  $r=0,879$  bis  $r=0,962$  und waren signifikant ( $p<0,01$ ). Auch hier zeigte sich eine stabile Beurteilung über beide Durchgänge, was auf eine hohe Zuverlässigkeit hinweist.

RE-TEST RELIABILITÄT	
	<i>Klinische Relevanz</i>
Rater 1	0,879
Rater 2	0,942
Rater 3	0,928
Rater 4	0,962
Rater 5	0,958

Tabelle 11: Re-Test Reliabilität aller 5 Untersucher durch den Vergleich 1. und 2. Durchgang für die Klinische Relevanz.

Die Abb. 10 gibt exemplarisch die graphische Darstellung der Differenzwerte von Rater 2 durch den Vergleich 1. und 2. Durchgang wieder.

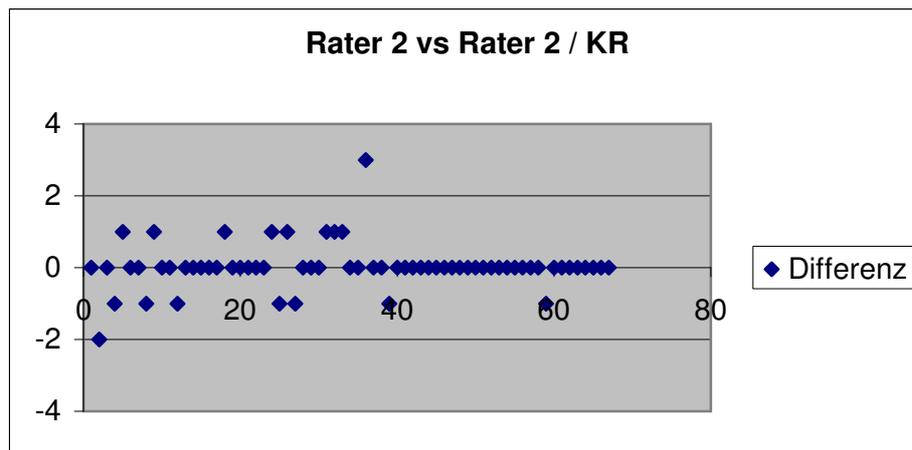


Abb. 10: Graphische Darstellung der Differenzwerte Rater 2 vs. Rater 2 (2. Dg)/ Klinische Relevanz bei insgesamt 67 Beurteilungen.

Betrachtet man die unten stehende tabellarische Darstellung (vgl. Tabelle 12), wird deutlich, dass die Abweichung um 0 Skalenwerte (= Übereinstimmung) am häufigsten vorkam. Sie lag bei einem Anteil von  $70,1 \pm 3,8\%$  an der Gesamtmenge. Die Abweichung um 1 Wert lag bei einem Anteil von  $24,8 \pm 5,2\%$  und um 2 bei  $3,0 \pm 3,9\%$ . Die Abweichungen um 3 oder 4 Skalenwerte zeigten dagegen eine deutlich abnehmende Häufigkeit. Abweichungen um mehr als 4 Werte kamen bei der Klinischen Relevanz im Re-Test nicht vor. Der Anteil der Abweichungen um  $\leq 1$  Skalenwert lag damit insgesamt bei  $94,9\%$ .

#### Abweichungstabelle Klinische Relevanz (n = 67) im Re-Test

Abweichung (in Skalenwerten)	0	1	2	3	4	5	6
Rater 1 zu 1 (2.Dg)	45	12	7	3	0	0	0
Rater 2 zu 2 (2.Dg)	50	15	1	1	0	0	0
Rater 3 zu 3 (2.Dg)	50	15	0	1	1	0	0
Rater 4 zu 4 (2.Dg)	44	22	0	1	0	0	0
Rater 5 zu 5 (2.Dg)	46	19	2	0	0	0	0
<b>Mittelwert</b>	<b>47,0</b>	<b>16,6</b>	<b>2,0</b>	<b>1,2</b>	<b>0,2</b>	<b>0,0</b>	<b>0,0</b>
<b>±STD</b>	<b>±2,5</b>	<b>±3,5</b>	<b>±2,6</b>	<b>±1,0</b>	<b>±0,4</b>	<b>±0,0</b>	<b>±0,0</b>
<b>Anteil in %</b>	<b>70,1</b>	<b>24,8</b>	<b>3,0</b>	<b>1,8</b>	<b>0,3</b>	<b>0,0</b>	<b>0,0</b>
<b>±STD in %</b>	<b>±3,8</b>	<b>±5,2</b>	<b>±3,9</b>	<b>±1,5</b>	<b>±0,6</b>	<b>±0,0</b>	<b>±0,0</b>

Tabelle 12: Tabellarische Darstellung der Abweichungen für die Klinische Relevanz im Re-Test.

### 4.1.3 Zusammenhang der Beurteilungsaspekte

In der unten stehenden Tabelle 13 wurde die Korrelation der Beurteilungen aus subjektiver Wheezing Rate und Klinischer Relevanz dargestellt. Die Korrelationskoeffizienten reichten von  $r=0,8307$  bis  $r=0,9532$ .

Die Ergebnisse waren signifikant unter Verwendung der 0,05 Signifikanzgrenze für die Produkt-Moment-Korrelation [Glass GV, 1970].

Dies weist auf einen eindeutigen linearen Zusammenhang zwischen der Beurteilung der Wheezing Rate und der Klinischen Relevanz hin, konstant über alle Rater.

KORRELATION WHEEZING RATE ZU KLINISCHE RELEVANZ	
WR/ KR Rater 1	0,8419
WR/ KR Rater 2	0,9377
WR/ KR Rater 3	0,9348
WR/ KR Rater 4	0,8307
WR/ KR Rater 5	0,9532

*Tabelle 13: Korrelation der Wheezing Rate  
zur Klinischen Relevanz.*

## 5 Diskussion

### 5.1 Diskussion der Methode

#### 5.1.1 Die Messungen

Da auf dem medizintechnischen Markt bis heute kein Gerät zur Lungengeräuschanalyse erhältlich ist, wurde in der vorliegenden Studie ein von der „Arbeitsgruppe Lung-Sound“ der Universität Marburg entwickelter Prototyp von der Versuchsleiterin für die Messungen verwendet [Schüttler F, 1996; Sulzer J, 1997 und 2000; Gross V, 2003]. Hinsichtlich der technischen Anforderungen und der Durchführung der Messungen wurden die aktuellen CORSA-Richtlinien berücksichtigt [Sovijärvi ARA, 2000a]. Da bereits in früheren Studien deutlich wurde, dass die Atemflussrate nicht nur einen Einfluss auf das Ergebnis der Lungengeräuschanalyse, sondern auch auf das Auftreten von Wheezing hat, erfolgten die Messungen zu dieser Arbeit unter standardisierter Atmung [Charbonneau G, 1983; Fenton FR, 1985; Kraman SS, 1986].

Die Gruppe des zu messenden Untersuchungskollektivs wurde mit 98 Personen groß gewählt, um für die anschließende Beurteilung der Lungengeräuschaufzeichnungen repräsentative Daten zu erhalten (67 Aufzeichnungen, davon 42 mit Wheezing).

#### 5.1.2 Die Rater

Für die Beurteilung der Lungengeräusche im Testset wurden in der vorliegenden Arbeit fünf Ärzte aus dem medizinischen Personal der Abteilung für Pneumologie des Universitätsklinikums Marburg als Untersucher (Rater) ausgewählt. Die Anzahl der Untersucher wurde für die vorliegende Studie als adäquat angenommen. Die Untersucher wiesen eine vergleichbar lange Erfahrung im Fachgebiet Pneumologie auf und für alle fünf war die Art der Beurteilung gleichermaßen neu. Untergruppen wurden in der vorliegenden Studie nicht gebildet. Pasterkamp wählte in einer Studie zur Beurteilung von Wheezing Untersucher aus verschiedenen medizinischen Berufen aus und teilte sie hiernach in Gruppen ein [Pasterkamp H, 1987b]. In Folgestudien könnte diese Idee aufgegriffen und im Hinblick auf die Reliabilität näher untersucht werden.

### 5.1.3 Lernset und Testset

Es gibt bis heute keinen internationalen Standard zur Beurteilung von Lungengeräuschen, weder in vivo durch die Auskultation, noch computergestützt anhand der Lungengeräuschanalyse. In anderen Studien mit in vivo Auskultation hatte sich gezeigt, dass sich im Verlauf des Beurteilungszeitraums die Lungengeräusche verändern, wenn Patienten für eine größere Anzahl Untersuchungen eingesetzt werden [Brooks D, 1995]. In der vorliegenden Arbeit wurden Lungengeräusche in vivo gemessen und über den Computer als Lungengeräuschaufzeichnungen wiedergegeben. Durch diese Art der Präsentation wurde zwar im Vergleich zur in vivo Auskultation eine künstliche Situation geschaffen, jedoch auch garantiert, dass allen Untersuchern exakt dieselben Geräusche zur Beurteilung vorlagen.

Des Weiteren müssen bei der Präsentation von Geräuschaufzeichnungen am Computer Untersucher-Fehler, entstanden durch unterschiedliche Handhabung des Stethoskops oder Zeitaufwendung bei der Auskultation, nicht berücksichtigt werden. Bei dieser Methode sind dagegen Fehler, die im Rahmen der Beurteilung am Computer auftreten, zu bedenken. Diese umfassen z.B. Eingabefehler auf der Computertastatur oder Ermüdungserscheinungen im Rahmen der konzentrierten Beurteilungsarbeit am Bildschirm. Um den Ermüdungsfaktor so gering wie möglich zu halten, wurde in der vorliegenden Arbeit die Anzahl der zu beurteilenden Lungengeräuschaufzeichnungen begrenzt. Die Gesamtmenge wurde auf 67 festgelegt, um den Zeitfaktor von 30 min. für die Beurteilung eines Durchgangs ohne Pause und bei mittlerer Beurteilungsfrequenz nicht zu überschreiten. Dieser Zeitraum wurde hinsichtlich der Ermüdungsbelastung und der Durchführbarkeit im klinischen Alltag als adäquat angenommen. Ein weiterer Faktor, der ebenfalls berücksichtigt werden muss, ist die gleichzeitige Präsentation sowohl eines visuellen als auch eines auditiven Reizes, welche parallel wahrgenommen und beurteilt werden müssen.

Es ist noch nicht abschließend geklärt, ob die Art der Geräuschpräsentation in Form einer Lungengeräuschaufzeichnung einen Einfluss auf die Beurteilung haben kann. Pasterkamp befasste sich mit dieser Frage und führte hier so genannte „psychoakustische Einflüsse“ an. Er gab zu bedenken, dass die selektiv auf einzelne Nebengeräusche gerichtete Aufmerksamkeit die Wahrnehmung verändern kann [Pasterkamp H, 1987a]. Die Nebengeräusche könnten z.B. lauter, länger oder aus-

geprägter erscheinen und die Beurteilung würde somit im Vergleich zur in vivo Beurteilung verändert ausfallen.

Schon in früheren Studien wurden Vergleichstests zu der Beurteilerübereinstimmung von in vivo auskultierten gegenüber durch einen Kassettenrekorder oder Computer wiedergegebenen Lungengeräuschen durchgeführt. Workum und Pasterkamp, aber auch andere Arbeitsgruppen, kamen zu dem Ergebnis, dass die Geräuschpräsentation keinen großen Einfluss auf die Übereinstimmung der Beurteiler zu haben scheint. Für die Untersucher waren die durch den Kassettenrekorder oder Computer erzeugten Geräusche zwar zunächst gewöhnungsbedürftig, wurden jedoch von den meisten als weitgehend realistisch bezeichnet und gut angenommen [Workum P, 1986; Pasterkamp H, 1987a; Kompis M, 1997].

## 5.2 Diskussion der Ergebnisse

### 5.2.1 Reliabilität der Untersucher

Bezüglich der Untersucherübereinstimmung in der subjektiven Beurteilung von Lungengeräuschen oder klinischen Symptomen fielen bei einigen Studien die Ergebnisse nicht eindeutig aus.

Pasterkamp ließ 40 klinisch tätige Personen aus unterschiedlichen Berufsgruppen eine subjektive Einschätzung von Wheezing (Dauer und Schweregrad) anhand von aufgezeichneten Lungengeräuschen abgeben. Die Beurteilungen der Untersucher wiesen große Abweichungen auf, sowohl im Vergleich zueinander, als auch zu sich selbst bei Wiederholung [Pasterkamp H, 1987b].

Smyllie untersuchte die Übereinstimmung von 9 Ärzten in der Beurteilung von 20 klinischen Symptomen bei Patienten mit verschiedenen Lungenerkrankungen und kam zu dem gleichen Ergebnis [Smyllie HC, 1965].

Andere Arbeiten wiesen dagegen eindeutigere Ergebnisse auf.

In einer Studie von Kiyokawa zur Crackle-Erkennung fand sich eine Übereinstimmung der Untersucher deutlich oberhalb der Zufallsgrenze ( $\kappa > 0,6$ ) [Kiyokawa H, 2001].

Angelilli fand ebenfalls eine starke Übereinstimmung bei drei von fünf Aspekten unter vier Untersuchern, die ein Asthma-Bewertungsschema an 17 Kindern mit Asthma bronchiale testeten ( $\kappa = 0,698-0,759$ ) [Angelilli ML, 2002].

Hinsichtlich der Genauigkeit und Reliabilität der Untersucher bei der Beurteilung von klinischen Symptomen bei Pleuraerguss fand die Arbeitsgruppe um Kalantri eine sehr gute Übereinstimmung unter den Untersuchern ( $\kappa = 0,84-0,89$ ) [Kalantri S, 2007].

In der vorliegenden Studie konnte für beide Beurteilungsaspekte ebenfalls eine sehr hohe Untersucherübereinstimmung gefunden werden. Die errechneten Korrelationskoeffizienten bei der Wheezing Rate reichten von  $r = 0,868$  bis  $r = 0,952$ , und bei der Klinischen Relevanz von  $r = 0,882$  bis  $r = 0,924$ . Es zeigte sich, dass die Geräuschaufzeichnungen der lungengesunden Probanden zu 100% von den Untersuchern korrekt spezifiziert wurden. Daraus lässt sich ableiten, dass die An- und Abwesenheit von Wheezing mit der verwendeten Methode zuverlässig erkannt werden konnte und nicht auf Zufall basiert. In anderen Arbeiten konnte ebenso gezeigt werden, dass es wichtig ist, eine feste Anzahl Gesunder zu integrieren, um

die Übereinstimmung im Fall „gesund“ (= Abwesenheit der zu beurteilenden Symptome) und „pathologisch“ vergleichen zu können [Meade TW, 1968; Workum P, 1986; Ehrhardt EE, 1990].

Insgesamt zeigte sich, dass die Abweichungen der Untersucher von  $\leq 1$  Skalenwert bei der Wheezing Rate mit einem Anteil von 95,5% und bei der Klinischen Relevanz mit 89,7% bezogen auf die Gesamtmenge vorkamen. Prozentual am häufigsten zeigte sich hiervon die Übereinstimmung (= keine Abweichung) der Untersucher. Als häufigste Abweichung war die Differenz um 1 Wert zu verzeichnen.

Dieses Phänomen lässt sich unter anderem auch dadurch erklären, dass jeder Untersucher gezwungen war, eine Entscheidung zu treffen. Hierdurch kann schnell eine Abweichung um einen Skalenwert auftreten. Eine absolute Übereinstimmung der fünf Untersucher bei einer Beurteilungsskala von 1 – 7 ist sehr unwahrscheinlich. Dies begründet sich darin, dass es schwierig ist, den subjektiven Eindruck von Wheezing in einen Zahlenwert zu übersetzen und anschließend Rückschlüsse auf die Klinische Relevanz zu ziehen.

Bei näherer Betrachtung der Beurteilungen in der vorliegenden Arbeit zeigte sich, dass Rater 1 bei beiden Beurteilungsaspekten, trotz sehr guter Übereinstimmung insgesamt, am meisten von den anderen abweicht. Als Ursache hierfür können zwei Möglichkeiten in Betracht gezogen werden. Entweder hat dieser Untersucher einfach nur unsauber gearbeitet (PC-Eingabefehler), oder er hatte nach dem Lernset noch keine in sich stabile Bewertungsmethode im Vergleich zu den anderen Untersuchern entwickelt. Es stellt auch eine Schwierigkeit dar, die klinische Erfahrung der Untersucher zu objektivieren. Somit könnte eventuell auch hierdurch die Abweichung von Rater 1 zu den anderen erklärt werden.

Die maximale Abweichung bei der Wheezing Rate lag bei sechs Skalenwerten, wohingegen sie bei der Klinischen Relevanz nur bei vier lag. Bei der näheren Betrachtung stellte sich für die Wheezing Rate heraus, dass die Abweichung um sechs Skalenwerte ursächlich auf zwei Untersucher (Rater 1 und 4) zurückzuführen war. Sie unterschieden sich in der Beurteilung einer Lungengeräuschaufzeichnung deutlich von denen der anderen. Zur Beurteilung, ob die Abweichung ein Versehen, oder aber eine wirklich andere Einschätzung bezüglich dieser Aufzeichnung darstellt, wurden beide Durchgänge näher betrachtet. Es wurde deutlich,

dass Rater 1 im zweiten Durchgang stark von seiner ersten Beurteilung abwich, wohingegen Rater 4 nur um einen Wert zu sich selbst differierte. Somit kann diskutiert werden, dass bei Rater 1 entweder im ersten Durchgang ein Versehen oder aber eine hohe Unsicherheit bei der Beurteilung dieser Geräuschaufzeichnung vorlag.

In anderen Arbeitsgruppen hatte sich gezeigt, dass Fehleinschätzungen bezüglich eines Nebengeräusches speziell dann festgestellt werden konnten, wenn die zugrunde liegenden Lungengeräusche eine hohe Intensität und die zu untersuchenden Nebengeräusche eine niedrige Amplitude aufwiesen [Kiyokawa H, 2001]. Auch dies kann zu einer Verunsicherung des Einzelnen und den punktuell aufgetretenen größeren Abweichungen in der Beurteilung geführt haben.

Abschließend lässt sich sagen, dass die außergewöhnlichen Abweichungen, mit Differenzen von vier und sechs Skalenwerten, bei beiden Beurteilungsaspekten nur vereinzelt vorkamen. Bezogen auf die Gesamtbeurteilung hatten sie keinen großen Einfluss auf die Reliabilität der Untersucher.

### **5.2.2 Re-Test Reliabilität der Untersucher**

Bereits in vorangegangenen Studien wurde diskutiert, ob eine hohe Zuverlässigkeit der Untersucher darauf hinweist, dass jeder Untersucher für sich ein individuelles Beurteilungsschema entwickelt hat, welches in sich konstant und wiederholbar ist [Brooks D, 1995; Stroud AE 2002]. Auch in der vorliegenden Studie, in der eine hohe Zuverlässigkeit der Untersucher gefunden wurde, kann dieses Phänomen eine Rolle gespielt haben. Weitere qualitative wissenschaftliche Untersuchungen sind jedoch notwendig, um diese individuellen Kriterien aufdecken zu können.

Das Eintreten eines Lerneffektes bei Wiederholung der Beurteilung muss ebenfalls diskutiert werden. Zwischen den beiden Beurteilungsdurchgängen lag bei jedem Untersucher mindestens ein Tag. Eine genauere Vorgabe bezüglich des Zeitpunktes der Wiederholung erfolgte darüber hinaus nicht. Von den Untersuchern wurden unterschiedliche Zeitabstände zwischen den beiden Durchgängen gewählt, so dass eine Aussage hinsichtlich eines Einflusses auf das Ergebnis aufgrund dessen nicht einheitlich getroffen werden kann. In Folgestudien könnte eine genauere Vorgabe erfolgen. Hierdurch wäre es möglich zu untersuchen, ob und inwieweit der zeitliche Abstand einen Einfluss auf die Re-Test Reliabilität der Untersucher hat.

Ein weiterer Aspekt, der in vorangegangenen Arbeiten im Hinblick auf die Zuverlässigkeit untersucht wurde, ist die klinische Erfahrung der einzelnen Untersucher. Diesbezüglich konnte jedoch in diversen Arbeiten kein signifikanter Unterschied zwischen erfahrenen und unerfahrenen Untersuchern festgestellt werden [Smyllie HC, 1965; Brooks D, 1993 und 1995; Fritz JM, 2000; Leher A, 2005].

Als deutlich wichtiger für die Zuverlässigkeit zeigte sich hingegen der Einfluss einer Schulungsintervention. Ein positiver Einfluss ist mittlerweile weitgehend belegt [Leher A, 2005]. Dina Brooks und Mitarbeiter untersuchten in einer Studie die Zuverlässigkeit bei der Auskultation von Lungengeräuschen. Sie konnten in einem zweiten Durchgang nach Schulung eine deutliche Verbesserung sowohl in der Übereinstimmung als auch der Zuverlässigkeit der Untersucher verzeichnen [Brooks D, 1995]. Ebenso zeigte sich in einer Studie von Melbye nach einer Schulung der Untersucher ein deutlicher Anstieg in der Zuverlässigkeit. Schwerpunkt der Untersuchung war hier die Einschätzung von FEV1-Werten in Zusammenhang mit dem Grad der Obstruktion [Melbye H, 1998]. In anderen Studien konnte dieser Effekt ebenfalls nachgewiesen werden [Homma Y 1985; Wigder HN, 1996; Yoshii C, 2002; Favrat B, 2004].

Auch in der vorliegenden Arbeit erfolgte vorab eine Schulungsintervention. Die Untersucher wurden durch die Präsentation von 50, dem Testset gleichwertigen, Lungengeräuschaufzeichnungen als Lernset auf die Beurteilung vorbereitet. Sie wurden dadurch mit der Beurteilung und den Kriterien der Einschätzung vertraut gemacht und über Aufbau und Entstehung der Frequenzbänder informiert.

Es erscheint sinnvoll, die Idee der Schulungsintervention weiter zu verfolgen. In einer Weiterentwicklung dessen könnte ein differenziertes Konzept erarbeitet und sein Einfluss in Folgestudien untersucht werden.

### 5.2.3 Zusammenhang der Beurteilungsaspekte

In der vorliegenden Studie lag der Schwerpunkt auf der Kategorisierung eines Symptoms und nicht der Beurteilung einer Erkrankung. Um jedoch einen klinischen Bezug herzustellen, wurde für jede Lungengeräuschaufzeichnung zudem um die Einschätzung der Klinischen Relevanz gebeten.

Bei der Auswertung wurden hohe Korrelationen zwischen den Beurteilungen der subjektiven Wheezing Rate und der Einschätzung der Klinischen Relevanz deutlich. Dieses Ergebnis spricht für einen von den Untersuchern subjektiv wahrgenommenen Zusammenhang zwischen den beiden Aspekten. Einer Geräuschaufzeichnung mit einer hohen Wheezing Rate wurde auch eine hohe Klinische Relevanz zugeordnet.

Hier finden sich Parallelen zum klinischen Alltag. Wheezing wird sehr häufig zur Einschätzung des Zustandsbildes eines Patienten mit obstruktiver Lungenerkrankung herangezogen und mit dem Schweregrad der Obstruktion oft nahezu gleichgesetzt [Pasterkamp H, 1997].

Ob ein realer Zusammenhang zwischen der subjektiven Wheezing Rate und der Klinischen Relevanz besteht, ist zumindest in dieser Studie, mit den verwendeten Kriterien und Vorgabe, nicht objektivierbar.

### 5.3 Ausblick

Das Nebengeräusch Wheezing gibt dem Arzt im klinischen Alltag einen wichtigen Hinweis auf eine bestehende Atemwegsobstruktion. In den Lungengeräuschaufzeichnungen ist es als Ansammlung bandförmiger dunkler Linien gut erkennbar. Daher bietet sich der Einsatz der computergestützten Geräuschanalyse in diesem Bereich an. Eine genaue Einschätzung des Schweregrades der Obstruktion anhand der Beurteilung von Wheezing (z.B. unter Verwendung der Wheezing Rate), kann somit objektiv und in einem längeren Betrachtungszeitraum erfolgen.

Es ist wichtig, die Entwicklung von Standards hinsichtlich Messung und Beurteilung von Lungengeräuschen voranzutreiben. Erst mit der Erarbeitung dieser Standards wird es möglich, eine Referenzdatenbank zu erstellen, auf die die Entwicklung und Erprobung eines Algorithmus zur automatischen Mustererkennung aufbaut.

In der vorliegenden Arbeit erfolgte die Bewertung von Wheezing bei den Lungengeräuschaufzeichnungen anhand einer Bewertungsskala von 1-7. Die Untersucher waren anhand dieser Skaleneinteilung in der Lage, Wheezing über die Wheezing Rate mit hohen Korrelationen im Vergleich zu einander, als auch zu sich selbst zu beurteilen. Des Weiteren war es ihnen möglich, eine klinische Relevanz abzuschätzen.

Die Ergebnisse belegen, dass die Verwendung der 7er-Skala reproduzierbare und verwertbare Daten liefert, was dafür spricht, diese Art der Skalierung in Zukunft weiter zu verwenden. Die Entwicklung einer genauen Einteilung, beispielsweise mit Übersetzen der einzelnen Skalenwerte in Prozentränge für die Wheezing Rate, ist denkbar und könnte in weiteren Studien erprobt werden.

## 6 Zusammenfassung

Die computergestützte Analyse von Lungengeräuschen hat in den letzten Jahrzehnten eine zunehmende Weiterentwicklung erfahren. Ziel dessen ist sowohl die weitere Erforschung ihrer Entstehungsmechanismen, als auch die Entwicklung einer automatischen Mustererkennung, die eine objektive Beurteilung ermöglicht. Trotz europäischer und internationaler Bemühungen um Vergleichbarkeit (CORSA/ILSA), existiert bis heute kein Standard für die Auswertung und Beurteilung von Lungengeräuschen. Somit bestand das Ziel dieser Studie darin, unter Verwendung einer standardisierten Analyse- und Beurteilungsmethode, einen Beitrag zu der Entwicklung eines Standards zu leisten.

In der vorliegenden Arbeit wurden die Lungengeräusche von insgesamt 98 Patienten und Probanden mit einem von der „Arbeitsgruppe Lung-Sound“ (Universität Marburg) entwickelten Prototyp zur Lungengeräuschanalyse unter standardisierter Atmung gemessen. Aus dem Untersuchungskollektiv wurden 24 Patienten mit Asthma bronchiale (FEV1% MW+STD 83±21), 18 mit COPD (FEV1% MW+STD 56±19) und 25 lungengesunde Probanden (FEV1% MW+STD 106±17) ausgewählt. Männer und Frauen in unterschiedlichen Altersgruppen (Mittelwert in Jahren: Asthma 41±18; COPD 64±13; Probanden 34±12) wurden in die Studie eingeschlossen.

Aus diesem Messkollektiv wurden 67 Lungengeräuschaufzeichnungen ausgewählt. Bei allen 42 Lungengeräuschaufzeichnungen aus der Patienten-Gruppe (Asthma und COPD) waren Wheezing-Ereignisse wahrnehmbar.

Die 67 Aufzeichnungen wurden fünf Ärzten der Abteilung Pneumologie des Universitätsklinikums Marburg mit vergleichbarer klinischer Erfahrung als „Testset“ zur audio-visuellen Beurteilung vorgelegt. Sie sollten das Nebengeräusch „Wheezing“ anhand der „Wheezing Rate“ und der „Klinischen Relevanz“ unter Verwendung einer Skaleneinteilung von 1-7 zur Einschätzung des Schweregrades wiederholt bewerten. Vorab war den fünf Untersuchern ein „Lernset“ mit vergleichbaren Daten präsentiert worden.

Hinsichtlich der Korrelationen (Reliabilität) der Untersucher (Rater) konnten für beide Beurteilungsaspekte signifikant hohe Werte gefunden werden ( $p < 0,01$ ). Die errechneten Korrelationskoeffizienten bei der Wheezing Rate reichten von  $r = 0,868$  bis  $r = 0,952$ , bei der Klinischen Relevanz von  $r = 0,882$  bis  $r = 0,924$ . Insgesamt kon-

zentrierten sich die Abweichungen der Untersucher bei der Wheezing Rate zu 95,5% und bei der Klinischen Relevanz zu 89,7% (Anteil an der Gesamtmenge  $n=67$ ) auf  $\leq 1$  Skalenwert.

Auch bei der Zuverlässigkeit, beurteilt mit dem Re-Test, zeigten die Untersucher für beide Beurteilungsaspekte signifikant hohe Werte ( $p < 0,01$ ). Die Korrelationen reichten bei der Wheezing Rate von  $r=0,857$  bis  $r=0,976$  und bei der Klinischen Relevanz von  $r=0,879$  bis  $r=0,962$ . Bei der Wheezing Rate lag im Re-Test der Anteil der Abweichungen von  $\leq 1$  Skalenwert bezogen auf die Gesamtmenge ( $n=67$ ) bei 96,7% und bei der Klinischen Relevanz bei 94,9%. Das Ergebnis spiegelt somit eine Beurteilungskonstanz der Untersucher bei Wiederholung unter gleich bleibenden Bedingungen wider.

Bei der Betrachtung des Zusammenhangs zwischen beiden Beurteilungsaspekten ergaben sich ebenfalls signifikant hohe Korrelationen ( $p < 0,05$ ) ( $r=0,8307$  bis  $r=0,9532$ ), die auf einen subjektiv wahrgenommenen Zusammenhang hinweisen.

Die 25 Lungengeräuschaufzeichnungen des lungengesunden Probandenkollektivs wurden zu 100% von allen Untersuchern in beiden Durchgängen korrekt beurteilt. Das Erkennen der An- und Abwesenheit von Wheezing erfolgte somit zuverlässig. Zusammenfassend konnte in dieser Arbeit gezeigt werden, dass klinisch erfahrene Untersucher in der Lage sind, das Nebengeräusch Wheezing anhand audio-visuell dargebotener Lungengeräuschaufzeichnungen übereinstimmend und zuverlässig zu beurteilen. Die verwendete Mess- und Analyseverfahren lieferte verwert- und differenzierbare Daten und die Reproduzierbarkeit der Methode zur standardisierten audio-visuellen Bewertung konnte aufgezeigt werden.

Die Verwendung einer Bewertungsskala von 1-7 zur Beurteilung von Wheezing kann daher auch für Folgestudien als Standard empfohlen werden.

## 6.1 Summary

Computerized respiratory sound analysis is a non-invasive and objective method used to detect abnormal respiratory noises. It has a great potential for monitoring of bronchial obstruction. Lately there has been a tendency towards the development of an automatic spectral classification system of breath sounds. Despite European and international efforts, there currently do not exist homogeneous standards used for analysis and assessment of respiratory noises. For this reason, the goal of this study, using standardized analysis and conclusive methods, was to take a step forward in the development of universal standards for research and analysis.

In this study 98 patients and healthy volunteers were recorded under standardized breathing (air-flow) using a prototype lung sound analyzer from the "Arbeitsgruppe Lung-Sound". This model only exists at the University of Marburg. From this collective, 24 patients with bronchial asthma (FEV1% MW+STD 83±21), 18 with COPD (FEV1% MW+STD 56±19) and 25 healthy volunteers (FEV1% MW+STD 106±17) were chosen. Men and women at different age (mean value: asthma 41±18; COPD 64±13; healthy volunteers 34±12) were included in the study.

From the above collective 67 lung sound recordings were used for the "Testset". All of the 42 recordings of the asthma and COPD patients had wheezing noise.

The 67 lung sound recordings were presented in two rounds to five physicians from the department of Pneumology at the Marburg University Clinic, who possessed comparable clinical experience. They had to rate these recordings from a simultaneous audio and visual presentation on the computer using a 1-7 rating-scale for evaluating the severity. The rating contained the assessment of the continuous adventitious sound of "wheezing" using the aspects "wheezing rate" and "clinical relevance". Prior to the ratings, the five physicians had been shown a "Lernset" (standardized training session), containing comparable data.

In the case of interrater reliability, there could be found a significant correlation ( $p < 0,01$ ) between the five raters for both aspects (wheezing rate:  $r = 0,868$  to  $r = 0,952$ ; clinical relevance:  $r = 0,882$  to  $r = 0,924$ ). The assessment of the raters did not deviate by more than one scale point in 95,5% of the cases for the aspect wheezing rate and 89,7% for the clinical relevance (share of the total number=67).

The validity, assessed by the re-test, also showed a significant correlation ( $p < 0,01$ ) for each rater and both aspects (wheezing rate:  $r = 0,857$  to  $r = 0,976$ ; clinical

relevance:  $r=0,879$  to  $r=0,962$ ). In 96,7% of the cases for the wheezing rate and 94,9% for the clinical relevance (share of the total number=67) did the assessment of the raters not deviate by more than one scale point. This result shows a continuous and valid rating ability for each rater during repeated presentation.

Evaluating the connection between the two aspects wheezing rate and clinical relevance, there could also be found a significant correlation ( $p<0,05$ ) ( $r=0,8307$  to  $r=0,9532$ ). This displays a subjective connection between the wheezing rate and the clinical relevance, perceived by the raters.

The 25 lung sound recordings of the healthy volunteers had been identified in total by each rater in both rounds. The presence of wheezing could be detected correctly using this method.

In conclusion, this study shows that clinical experienced physicians are able to rate the continuous adventitious sound of wheezing by using an audio-visual presented lung sound recording with high agreement and validity.

The lung sound analyser and method, used to achieve lung sound recordings, produced differentiated data as well as valid data, and was proven to be reproducible. Using a 1-7 rating-scale can be therefore recommended for further studies as a standardized technique.

## 7 Literaturverzeichnis

1. American Thoracic Society Committee on Pulmonary Nomenclature: Report of the ATS-ACCP ad hoc subcommittee on nomenclature. *Am Thorac Soc News* 1977; 3: 5-6.
2. Angelilli ML, Thomas R: Inter-rater evaluation of a clinical scoring system in children with asthma. *Ann Allergy Asthma Immunol* 2002; 88 (2): 209-14.
3. Baughman R, Loudon RG: Wheezing in asthma and its response to bronchodilators: Waveform analysis using fast fourier transform technique. *Am Rev Respir Dis* 1982; 125: 96.
4. Baughman RP, Loudon RG: Quantification of wheezing in acute asthma. *Chest* 1984; 86 (5): 718-722.
5. Baughman RP, Loudon RG: Lung sound analysis for continuous evaluation of airflow obstruction in asthma. *Chest* 1985; 88 (3): 364-368.
6. Bentur L, Beck R, Shinawi M, Naveh T, Gavriely N: Wheeze monitoring in children for assessment of nocturnal asthma and response to therapy. *Eur Respir J* 2003; 21: 621-626.
7. Brooks D, Wilson L, Kelsey CK: Accuracy and reliability of „specialized“ physical therapists in auscultating tape-recorded lung sounds. *Physiotherapy Canada* 1993; 45: 21-24.
8. Brooks D, Thomas J: Interrater reliability of auscultation of breath sounds among physical therapists. *Physical Therapy* 1995; 75 (12): 1082-1088.
9. Charbonneau G, Racineux JL, Sudraud M, Tuchais E: An accurate recording system and its use in breath sounds spectral analysis. *J Appl Physiol* 1983; 7: 1120-1127.
10. Charbonneau G, Sudraud M and Soufflet G: Method for the evaluation of flow rate from pulmonary sounds. *Bull Eur Physiopathol Respire* 1987; 33: 461-463.

11. Dalmay F, Antonini MT, Marquet P, Menier R: Acoustic properties of the normal chest. *Eur Respir J* 1995; 8: 1761-1769.
12. Dantzker DR, Pringleton S, Pierce JA, Niewoehner DE, Thurlbeck WM, Buist AS (American Thoracic Society): Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. The official statement of the Am Thor Soc 1986; 225-243.
13. Dittmar A: Frequenzanalyse vesikulärer Atemgeräusche gesunder Probanden, unter besonderer Berücksichtigung des Alters, des body-mass-index und der packyears. Dissertation 2002; Universitätsbibliothek Marburg.
14. Druzgalski CK, Donnerberg RL, Campbell RM: Techniques of recording respiratory sounds. *J of Clin Engineering* 1980; 5 (4): 321-330.
15. Earis JE, Cheetham BMG: Current methods used for computerized respiratory sound analysis. *Eur Respir Rev* 2000; 10 (77): 586-590.
16. Ehrhardt EE, Lowe JE: Observer variation in equine abdominal auscultation. *Equine Vet* 1990; 22 (3): 182-5.
17. Elphick HE, Lancaster GA, Solis A, Majumdar A, Gupta R, Smyth RL: Validity and reliability of acoustic analysis of respiratory sounds in infants. *Arch Dis Child* 2004; 89 (11): 1059-63.
18. Favrat B, Pecoud A, Jaussi A: Teaching cardiac auscultation to trainees in internal medicine and family practice: does it work? *BMC Med Educ* 2004; 4: 5.
19. Fenton FR, Pasterkamp H: Automated spectral characterisation of wheezing in asthmatic children. *IEEE Jan* 1985; BME-32, 1: 50-55.
20. Forgacs P, Lond MD, MRCP: Crackles and wheezes. *The Lancet* 1967; 2: 203-205.
21. Forgacs P: Lung sounds. *British Journal of Diseases of the Chest* 1969; 63 (1): 1-12.

22. Forgacs P: The Functional Basis of Pulmonary Sounds. *Chest* 1978a; 73 (3): 399.
23. Forgacs P: Lung sounds. Cassel and Coller. McMillan Publishers London Ltd. 1978b.
24. Fritz JM, Delitto A, Vignovic M, Busse RG: Interrater reliability of judgements of the centralization phenomenon and status change during movement testing in patients with low back pain. *Arch Phys Med Rehabil* 2000; 81 (1): 57-61.
25. Gavriely N, Palti Y, Alroy G: Spectral characteristics of normal breath sounds. *J Appl Physiol* 1981; 307-314.
26. Gavriely N, Shee TR, Cugell DW, Grotberg JB: Flutter in flow-limited collapsible tubes: a mechanism for generation of wheezes. *J Appl Physiol* 1989a; 66: 2251-61.
27. Gavriely N, Kelly KB, Grotberg JB, Loring SH: Critical pressures required for generation of forced expiratory wheezes. *J Appl Physiol* 1989b; 66: 1136-1142.
28. Gavriely N, Nissan M, Rubin A-H, Cugell DW: Spectral characteristics of chest wall breath sounds in normal subjects. *Thorax* 1995; 50: 1292-1300.
29. Gavriely N, Cugell DW: Airflow effects on amplitude and spectral content of normal breath sounds. *J Appl Physiol* 1996; 80: 5-13.
30. Glass GV, Stanley JC: Statistical methods in education and psychology. New Jersey: Prentice Hall, Inc. Englewood Cliffs 1970; 521.
31. Gross R, Schölmerich P, Gerok W: Die Innere Medizin. Schattauer-Verlag GmbH 1999; 9. Auflage.
32. Gross V, Dittmar A, Penzel T, Schuettler F, von Wichert P: The relationship between normal lung sounds, age and gender. *Am J Respire Crit Care Med* 2000; 162: 905-909.

33. Gross V, Koehler U, Penzel T, Reinke C, von Wichert P, Vogelmeier C: Effect of subcutaneous fatty tissue on normal lung sounds. *Biomed Tech (Berl)* 2003; 48: 182-185.
34. Grotberg JB, Davis SH: Fluid-dynamic flapping of a collapsible channel: sound generation and flow limitation. *J Biomech* 1980; 13: 219-230.
35. Grotberg, JB, Reiss, EL: Subsonic flapping flutter. *J Sound Vibr* 1984; 92: 349-361.
36. Hannon RR, Lyman RS: Studies on pulmonary acoustics. II. The transmission of tracheal sounds through freshly exenterated sheep`s lungs. *Am Rev Tuberc* 1929; 19: 360.
37. Homma Y, Matsuzaki M, Ogasawara H, Munakata M: Phonopneumograph possible for real-time tracing. *Comput Biomed Res* 1985; 18 (6): 502-9.
38. Kalantri S, Joshi R, Lokhande T, Singh A, Morgan M, Colford JM Jr, Pai M: Accuracy and reliability of physical signs in the diagnosis of pleural effusion. *Respir Med* 2007; 101 (3): 431-8. Epub 2006 Sep 11.
39. Kelemen SA, Cseri T, Marozsan I: Information obtained from tussigrams and the possibilities of their application in medical practice. *Bull Eur Physiopathol Respir* 1987; 13 (10): 51s-56s.
40. Kiyokawa H, Greenberg M, Shirota K, Pasterkamp H: Auditory detection of simulated crackles in breath sounds. *Chest* 2001; 119 (6): 1886-92.
41. Kompis M, Russi EW: Computer-based lung sound simulation. *Med Biol Eng Comput* 1997; 35 (3): 231-8.
42. Kraman SS: Effects of lung volume and airflow on the frequency spectrum of vesicular lung sounds. *Respir Physiol* 1986; 66: 1-9.
43. Laennec: De l`auscultation médiante ou traité du diagnostic de maladies des poumons et du cœur. Brosset et Chaudé, Paris 1819.

44. Leher A, Graf K, PhoDuc JM, Rammelsberg P: Is there a difference in the reliable measurement of temporomandibular disorder signs between experienced and inexperienced examiners? *J Orofac Pain* 2005; 19 (1): 58-64.
45. Loudon RG: Auscultation of the lung. *Clin Notes Respir Dis* 1982; 21 (2): 3-7.
46. Loudon RG, Murphy RL: Lung Sounds. State of the Art. *Am Rev Respir Dis* 1984; 130: 663-973.
47. Man SF, McAlister FA, Anthonisen MR, Sin DD: Contemporary management of chronic pulmonary disease: clinical applications. *Jama* 2003; 290 (17): 2313-6.
48. Marini JJ, Pierson DJ, Hudson LD, Lakshminarayan S: The significance of wheezing in chronic airflow obstruction. *Am Rev Respir Dis* 1979; 120: 1069-1072.
49. Masoli M, Fabian D, Holt S, Beasley R: The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy* 2004; 59: 469-478.
50. Meade TW, Gardner MJ, Cannon P, Richardson PC: Observer variability in recording the peripheral pulses. *Br Heart J* 1968; 30: 661-65.
51. Medbo A, Melbye H: Lung function testing in the elderly – can we still use FEV(1)/ FVC <70% as a criteria of COPD? *Respir Med* 2007; 101 (6): 1097-105.
52. Melbye H, Aaraas I, Hana J, Hensrud A: Improving pulmonary auscultation as a tool in the diagnosis of bronchial obstruction – results of an educational intervention. *Scand J Prim Health Care* 1998; 16.
53. Meslier N, Charbonneau G, Racineux JL: Wheezes. *Eur Respir J* 1995; 8: 1942-1948.
54. Mussell MJ: The need for standards in recording and analysing respiratory sounds. *Med Biol Eng Comput* 1992; 30 (2): 129-39.

55. Noviski N, Cohen L, Springer C, Bar-Yishaye E, Avital A, Godfrey S: Bronchial provocation determined by breath sounds compared with lung function. *Arch Dis Child* 1991; 66 (8): 952-5.
56. Pasterkamp H, Fenton R, Tal A, Chernick V: The effect of anti-cholinergic treatment on postexertional wheezing in asthma studied by phonopneumography and spirometry. *Am Rev Respir Dis* 1985; 132: 16-21.
57. Pasterkamp H, Wiebicke W, Fenton R: Subjective assessment vs computer analysis of wheezing in asthma. *Chest* 1987a; 91 (3): 376-381.
58. Pasterkamp H, Montgomery M, Wiebicke W: Nomenclature used by health care professionals to describe breath sounds in asthma. *Chest* 1987b; 92 (2): 346-352.
59. Pasterkamp H, Kraman SS, Wodicka GR: Respiratory Sounds. *Advances beyond the Stethoscope. Am J Respir Crit Care Med* 1997; 156: 974-987.
60. Ploysongsang Y, Baughman RP, Loudon RG, Rashkin MC: Factors influencing the production of wheezes during expiratory maneuvers in normal subjects. *Respiration* 1988; 54: 50-60.
61. Rietveld S, Dooijes EH, Rijssenbeck-Nouwens LH, Smit F, Prins PJ, Kolk AM, Everaerd WA: Characteristics of wheeze during histamine-induced airways obstruction in children with asthma. *Thorax* 1995; 50: 143-148.
62. Rietveld S, Oud M, Dooijes EH: Classification of asthmatic breath sounds: preliminary results of the classifying capacity of human examiners versus artificial neural networks. *Comput Biomed Res* 1999; 32: 440-448.
63. Sanchez I, Avital A, Wong I, Tal A, Pasterkamp H: Acoustic vs spirometric assessment of bronchial responsiveness to methacholine in children. *Pediatr Pulmonol* 1993; 15: 28-35.
64. Sankur B, Kahya PY, Güler EC, Engin T: Comparison of AR-based algorithms for respiratory sound classification. *Comput Biol Med* 1994; 24: 67-76.

65. Schreur HJW, Sterk PJ, Vanderschoot J, van Klink HC, van Vollenhoven E, Dijkman JH: Lung sound intensity in patients with emphysema and in normal subjects at standardised airflow. *Thorax* 1992; 47: 674-679.
66. Schüttler F, Penzel T, von Wichert P: Digital recording and computer-based analysis of lung sounds. *Proc of Annual Internat Conf of IEEE Eng in Med and Biol* 1996; 18: 327.
67. Shabtai-Musih Y, Grotberg JB, Gavriely N: Spectral content of forced expiratory wheezes during air, He, and SF6 breathing in normal humans. *J Appl Physiol* 1992; 72: 629-635.
68. Shim H, Williams H: Relationship of wheezing to the severity of obstruction in asthma: *Arch Intern Med* 1983; 143: 890-892.
69. Smyllie HC, Blendis LM, Amitage D: Observer disagreement in physical signs of the respiratory system. *Lancet* 1965; 2: 412-413.
70. Sovijärvi ARA, Vanderschoot J, Earis JE: European Respiratory Society Task Force Report: Computerized Respiratory Sounds Analysis (CORSA): recommended standards for terms and techniques. *Eur Respir Rev* 2000a; 10 (77): 585-649.
71. Sovijärvi ARA, Dalmaso F, Vanderschoot J, Malmberg LP, Righini G, Stoneman SAT: Definition of terms for applications of respiratory sounds. *Eur Respir Rev* 2000b; 10 (77): 597-610.
72. Stroud AE, Lawrie BW, Wiles CM: Inter- and intra-rater reliability of cervical auscultation to detect aspiration in patients with dysphagia. *Clin Rehabil* 2002; 16 (6): 640-5.
73. Sulzer J, Schüttler F, Penzel T, von Wichert P: A simple method for detecting bronchial breathing sounds. *Proc 22<sup>nd</sup> Internat Conf of Lung Sounds* 1997; Tokyo, Japan.

74. Sulzer J: Analyse von Atemgeräuschen: Entwicklung eines computergestützten Verfahrens zur Unterscheidung von Vesiculär- und Bronchialatmen mit Hilfe einer atemflußabhängigen Fast Fourier Transformation. Dissertation 2000, Universitätsbibliothek Marburg.
75. Wigder HN; Johnson DR, Cohan S, Felde R, Colella R: Assessment of lung auscultation by paramedics. *Ann Emerg Med* 1996; 28 (3): 309-12.
76. Workum P, Delbono EA, Holford SK, Murphy RLH: Observer agreement, chest auscultation, and crackles in asbestos-exposed workers. *Chest* 1986; 89: 27-9.
77. Yoshii C, Anzai T, Yatera K, Kawajiri T, Nakashima Y, Kido M: A new medical education using a lung sound auscultation simulator called "Mr. Lung". *J UOEH* 2002; 24 (3): 249-55.

## 8 Anhang

### 8.1 Abbildungsverzeichnis

<i>Abb. 1: Beispiel einer Lungengeräuschaufzeichnung mit Wheezing (tracheales Mikrophon) in den vier Ventilationsphasen nach Anwendung der Fast Fourier Transformation.</i>	3
<i>Abb. 2: Messung eines Patienten bei standardisierter Atmung unter Online-Kontrolle.</i>	14
<i>Abb. 3: Beispiel einer Lungengeräuschaufzeichnung eines lungengesunden Probanden (ohne Wheezing) inklusive Atemfluss-Kurve.</i>	20
<i>Abb.4: Beispiel einer Lungengeräuschaufzeichnung eines Patienten mit obstruktiver Lungenerkrankung (mit deutlichem in- und expiratorischen Wheezing).</i>	20
<i>Abb. 5: Graphische Darstellung der Differenzwerte Rater 3 vs. Rater 4/ Wheezing Rate bei insgesamt 67 Beurteilungen.</i>	23
<i>Abb. 6: Häufigkeitsverteilung der Differenzwerte Rater 1 zu den Ratern 2-5 in % für die Wheezing Rate.</i>	24
<i>Abb. 7: Graphische Darstellung der Differenzwerte Rater 3 vs. Rater 4/ Klinische Relevanz bei insgesamt 67 Beurteilungen.</i>	25
<i>Abb. 8: Häufigkeitsverteilung der Differenzwerte Rater 1 zu den Ratern 2-5 in % für die Klinische Relevanz.</i>	27
<i>Abb. 9: Graphische Darstellung der Differenzwerte Rater 3 vs. Rater 3 (2. Dg)/ Wheezing Rate bei insgesamt 67 Beurteilungen.</i>	28
<i>Abb. 10: Graphische Darstellung der Differenzwerte Rater 2 vs. Rater 2 (2. Dg)/ Klinische Relevanz bei insgesamt 67 Beurteilungen.</i>	30

## 8.2 Tabellenverzeichnis

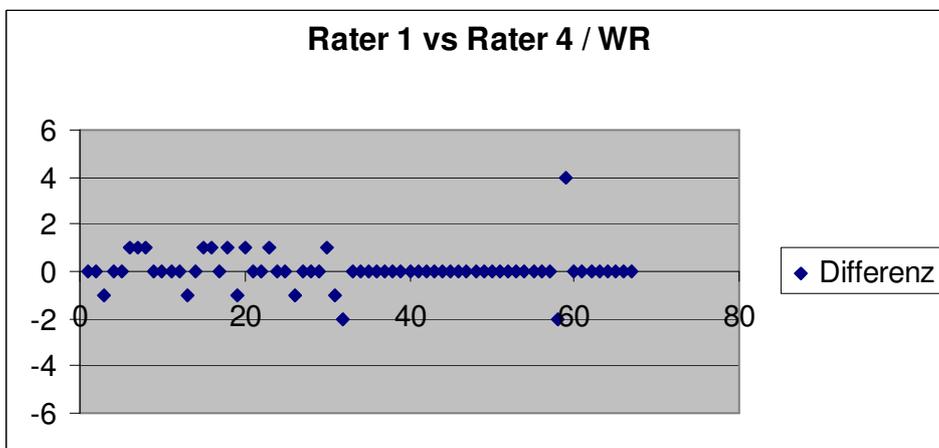
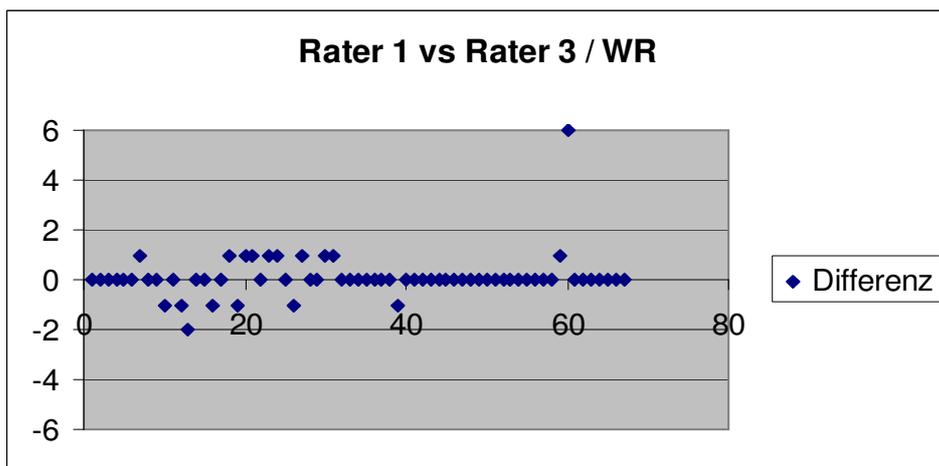
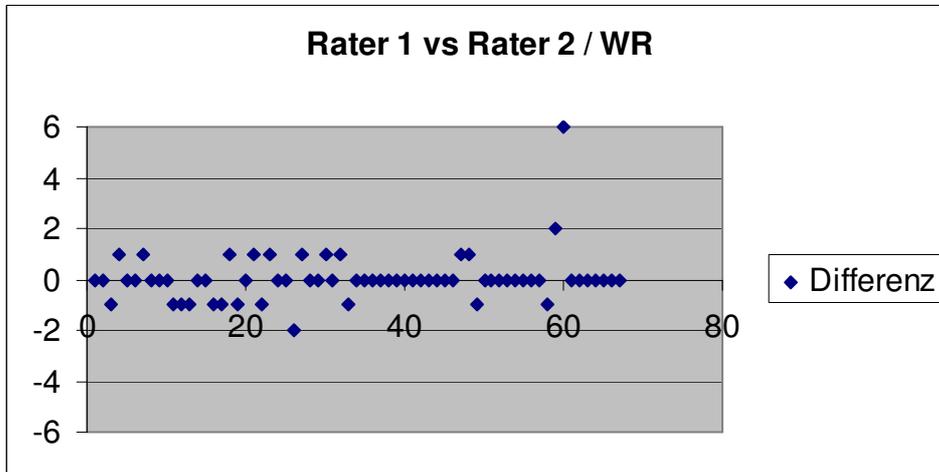
<i>Tabelle 1: Übersicht über die Einteilung der Lungengeräusche nach der aktuellen Nomenklatur der „International Lung Sound Association“.</i>	2
<i>Tabelle 2: Einige Lungenerkrankungen und ihre assoziierten Geräusche [Sankur B, 1994].</i>	4
<i>Tabelle 3: Klinische Parameter des Untersuchungskollektivs im Testset.</i>	12
<i>Tabelle 4: Parameter der Geräuschaufzeichnungen im Testset.</i>	19
<i>Tabelle 5: Korrelation aller 5 Untersucher nach Spearman-Rho für die Wheezing Rate.</i>	22
<i>Tabelle 6: Tabellarische Darstellung der Abweichungen mit Mittelwert und Standardabweichung für die Wheezing Rate.</i>	24
<i>Tabelle 7: Korrelation aller 5 Untersucher nach Spearman-Rho für die Klinische Relevanz.</i>	25
<i>Tabelle 8: Tabellarische Darstellung der Abweichungen mit Mittelwert und Standardabweichung für die Klinische Relevanz.</i>	26
<i>Tabelle 9: Re-Test Reliabilität aller 5 Untersucher durch den Vergleich 1. und 2. Durchgang für die Wheezing Rate.</i>	28
<i>Tabelle 10: Tabellarische Darstellung der Abweichungen für die Wheezing Rate im Re-Test.</i>	29
<i>Tabelle 11: Re-Test Reliabilität aller 5 Untersucher durch den Vergleich 1. und 2. Durchgang für die Klinische Relevanz.</i>	30
<i>Tabelle 12: Tabellarische Darstellung der Abweichungen für die Klinische Relevanz im Re-Test.</i>	31
<i>Tabelle 13: Korrelation der Wheezing Rate zur Klinischen Relevanz.</i>	32

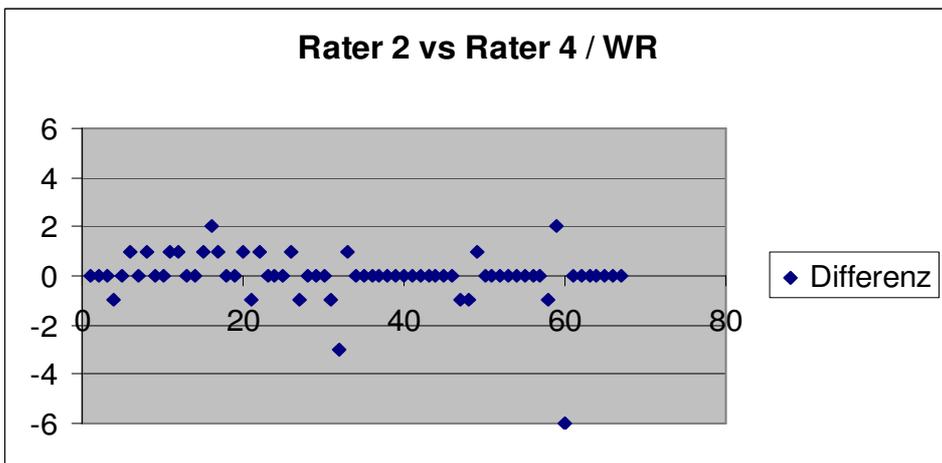
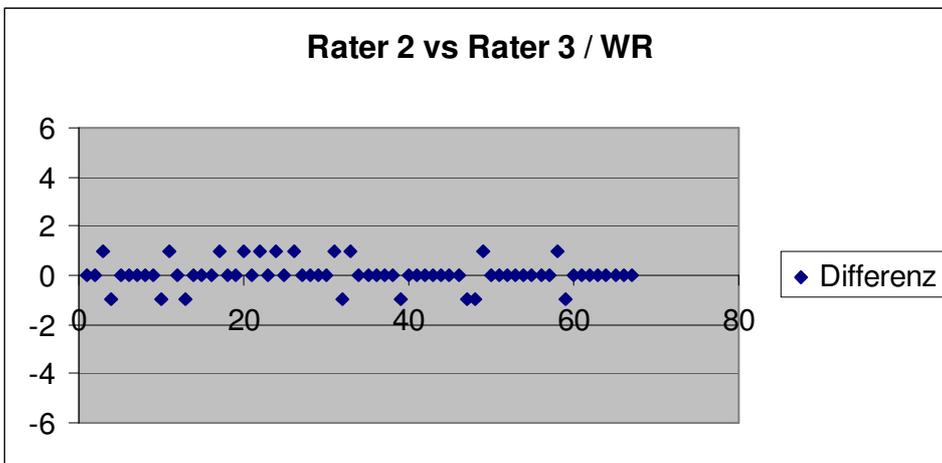
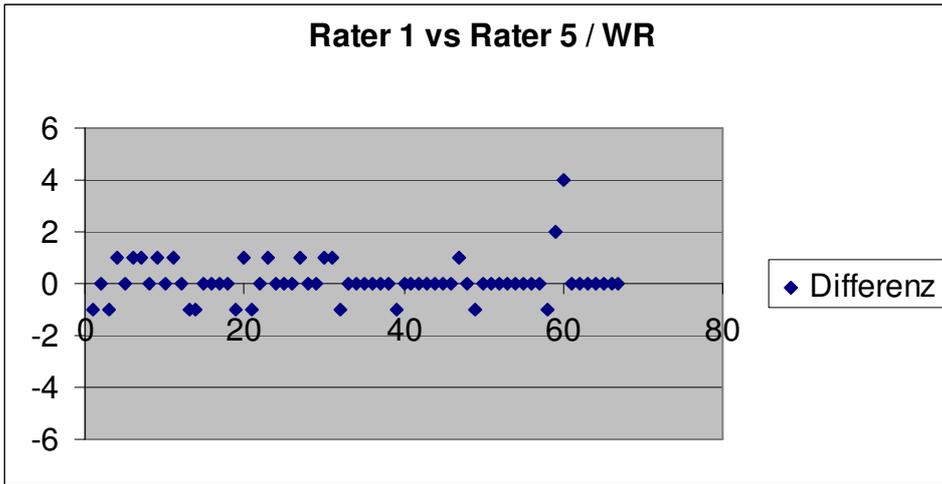
### 8.3 Abkürzungsverzeichnis

Abb.	Abbildung
Abw.	Abweichung
A/D	Analog/Digital
AG	Atemgeräusch
ATS	American Thoracic Society
COPD	Chronisch obstruktive Lungenerkrankung
CORSA	Computerized Respiratory Sound Analysis
dB	Dezibel
Dg	Durchgang
FEV1	Forciertes expiratorisches Volumen in 1 Sekunde
FFT	Fast Fourier Transformation
FVC	Forcierte Vitalkapazität
Hz	Hertz
ID	Identifikation
ICR	Intercostalraum
ILSA	International Lung Sounds Association
KR	Klinische Relevanz
l	Liter
l/s	Liter pro Sekunde
MARS	Marburg Respiratory Sounds (Multimedia Database)
MATLAB	Programmroutine
MEF50	Maximum Expiratory Flow bei 50% der FVC
mm	Millimeter
mm <sup>2</sup>	Quadratmillimeter
ms	Millisekunde
MW	Mittelwert
n	Gesamtmenge
PC	Personalcomputer
PEF	Peak Expiratory Flow
STD	Standardabweichung
WHO	World Health Organization
WR	Wheezing Rate

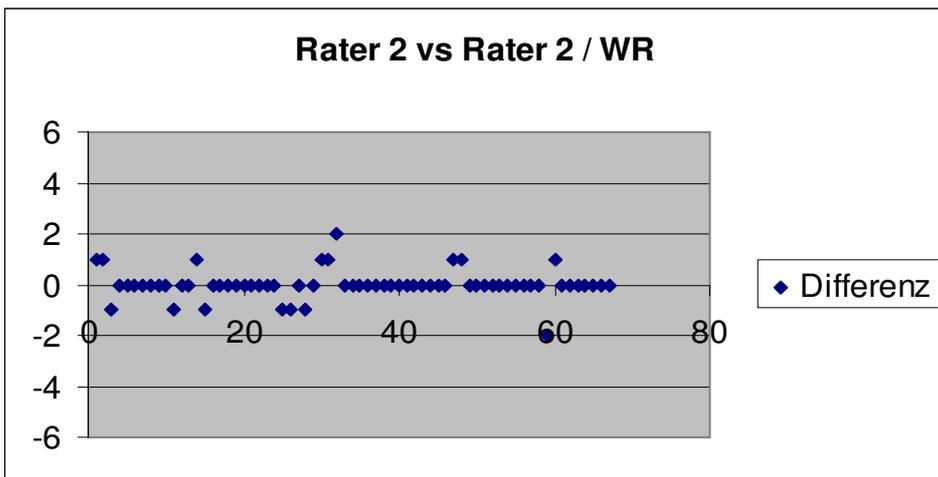
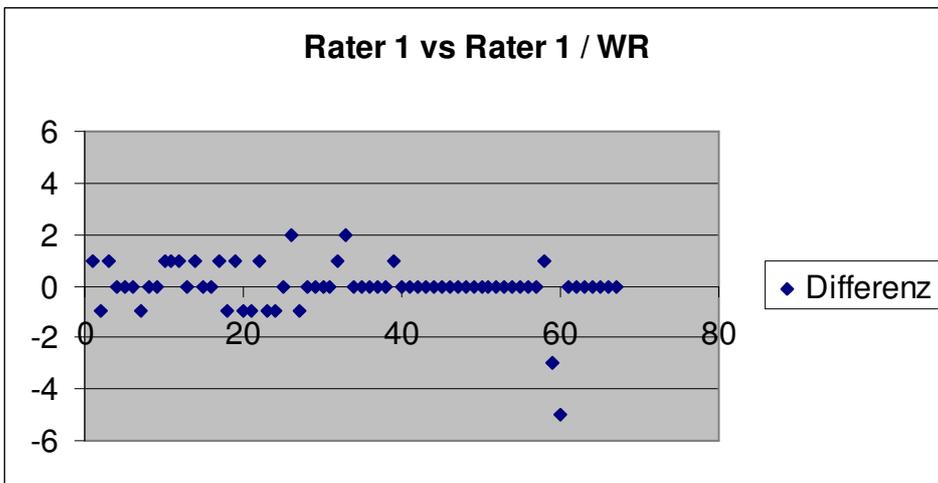
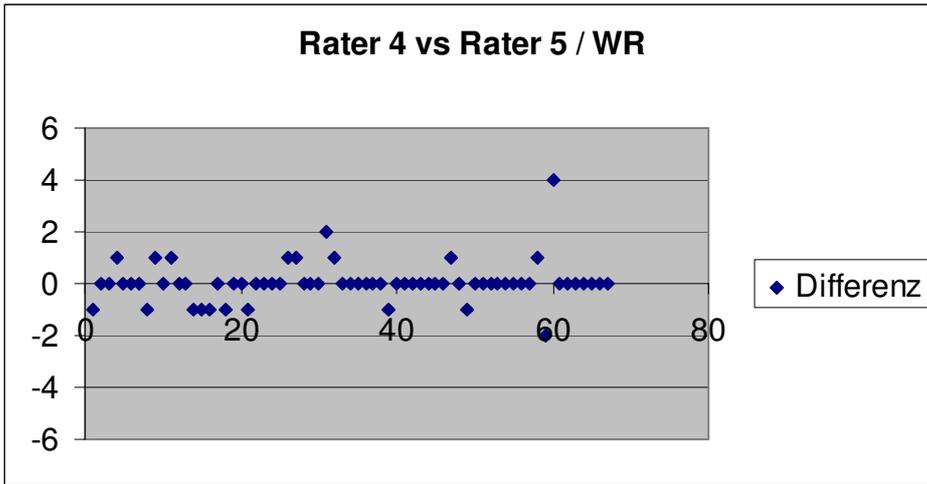
## 8.4 Graphische Darstellungen

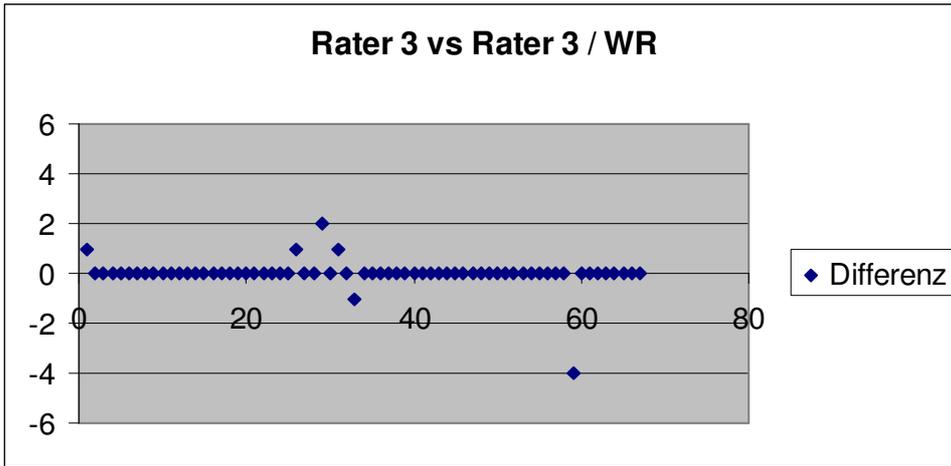
### Differenzwerte Wheezing Rate



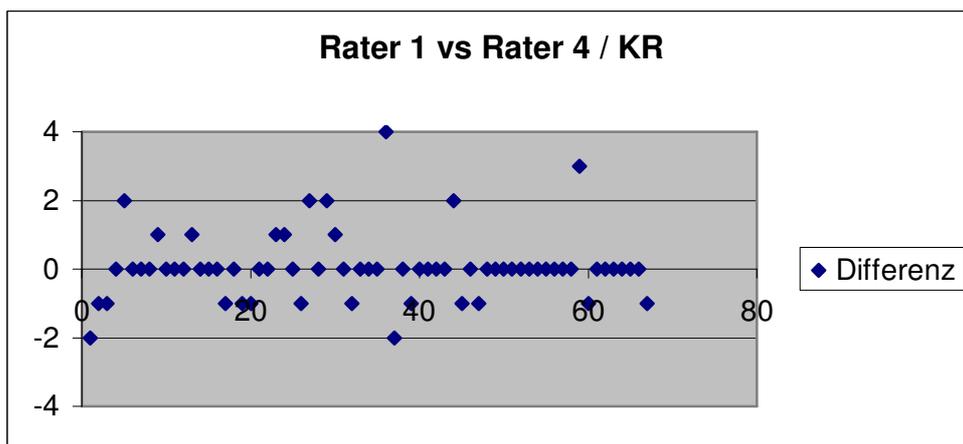
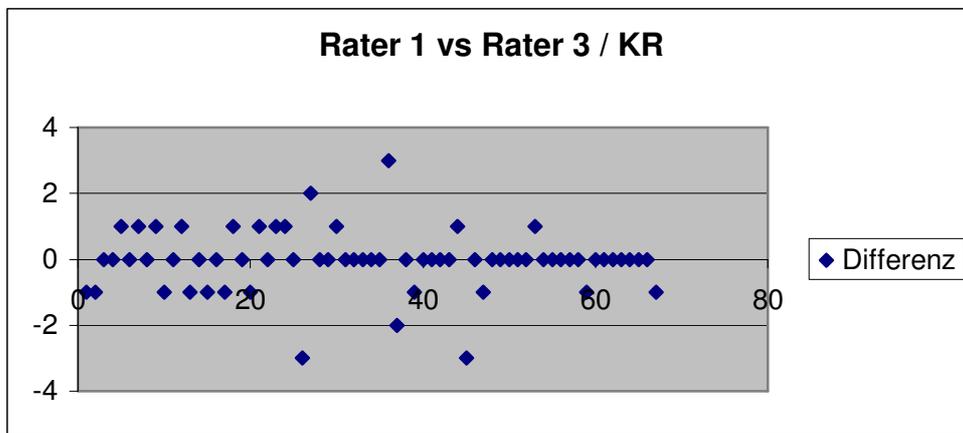
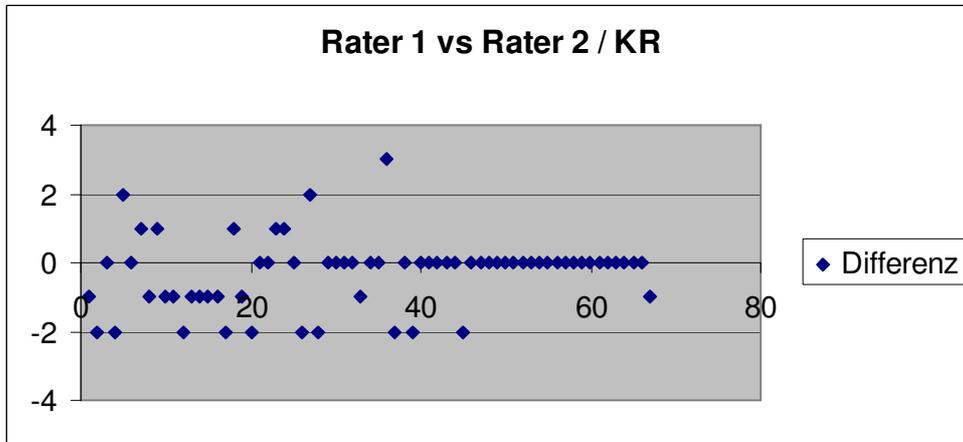


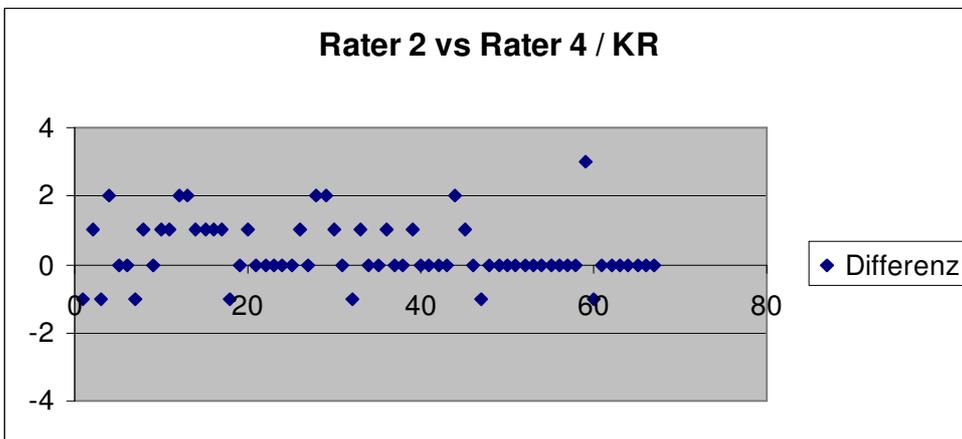
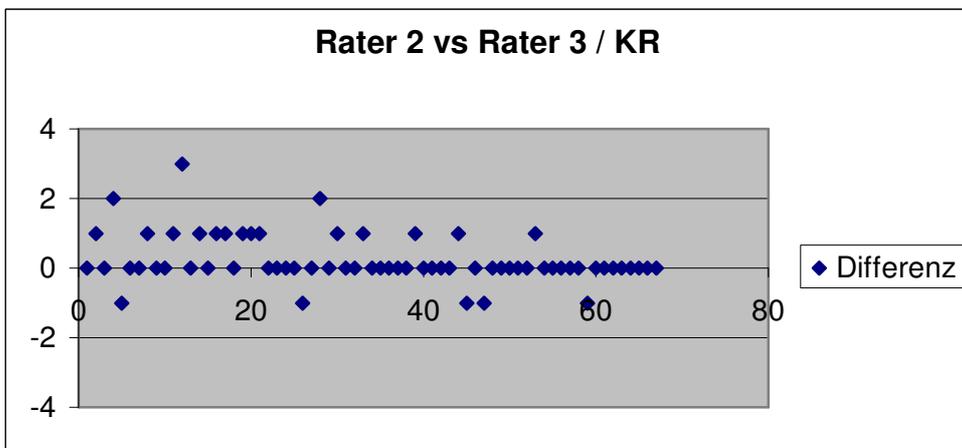
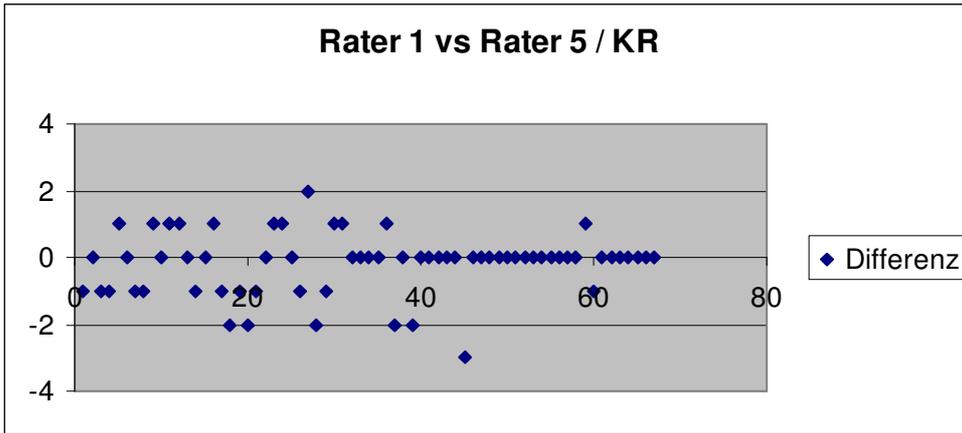


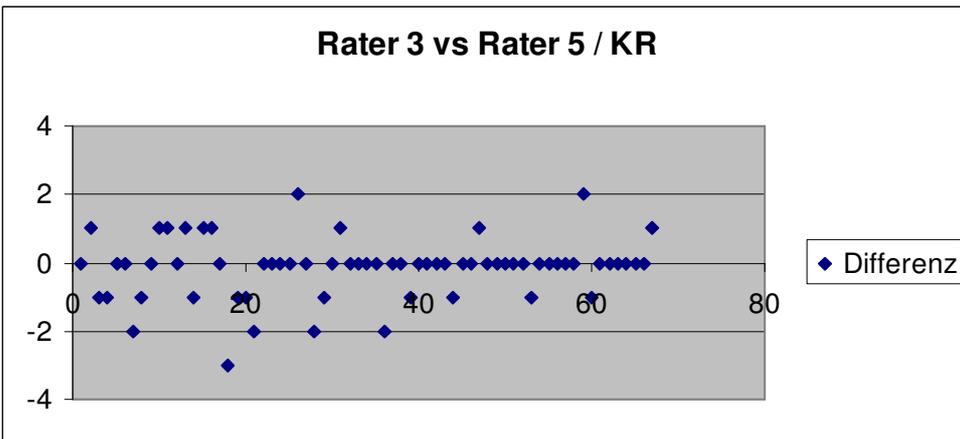
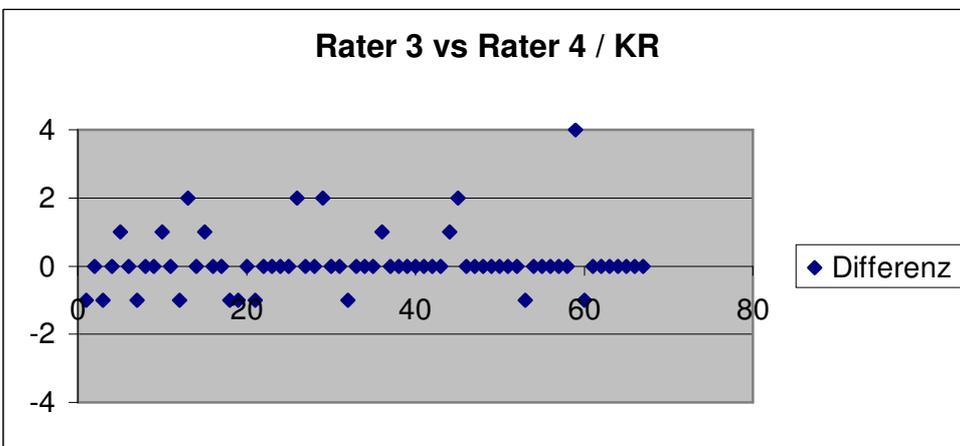
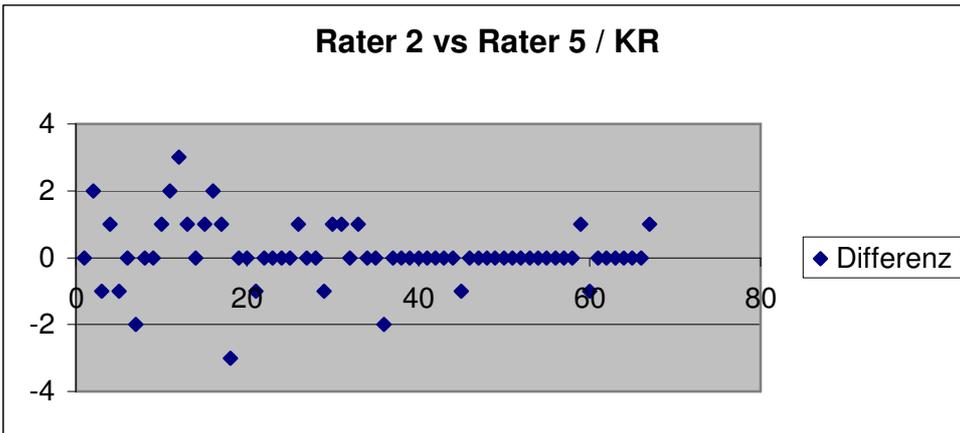


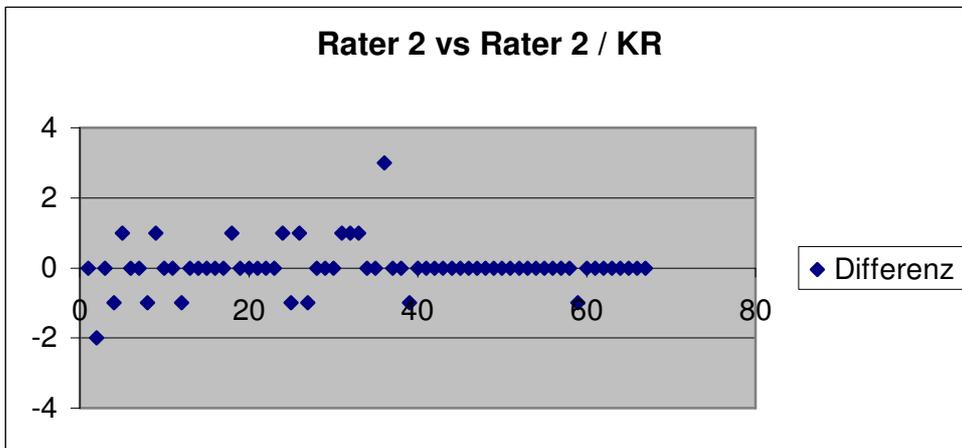
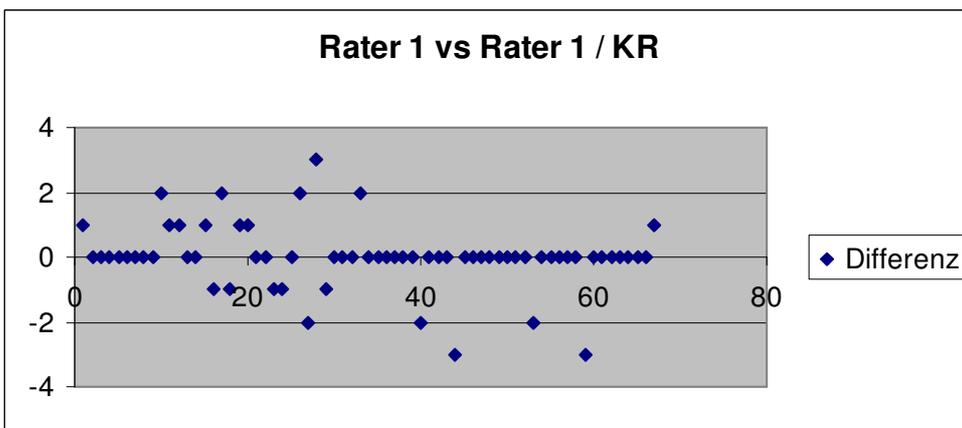
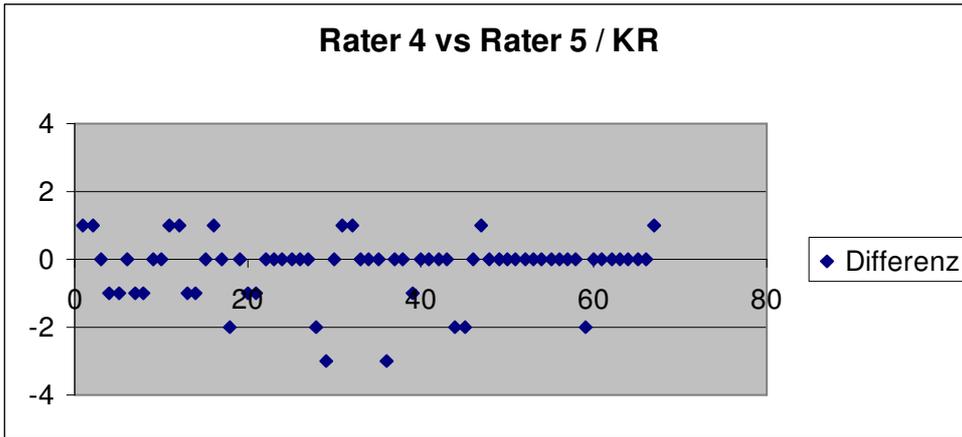


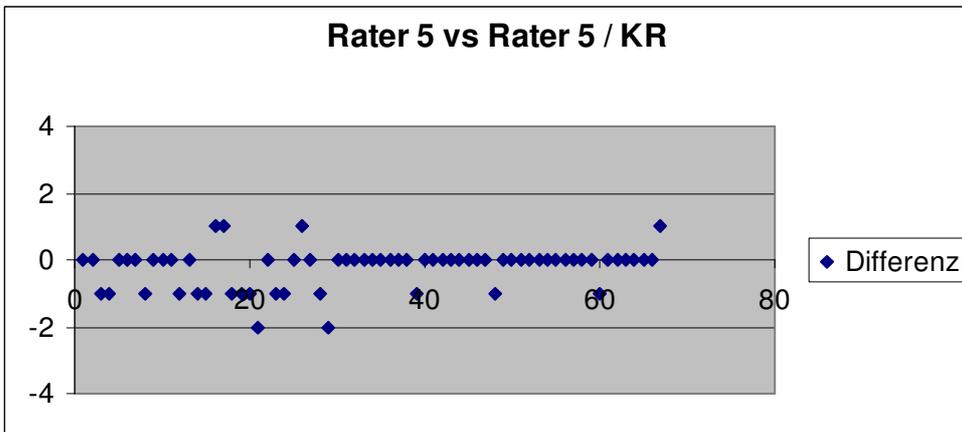
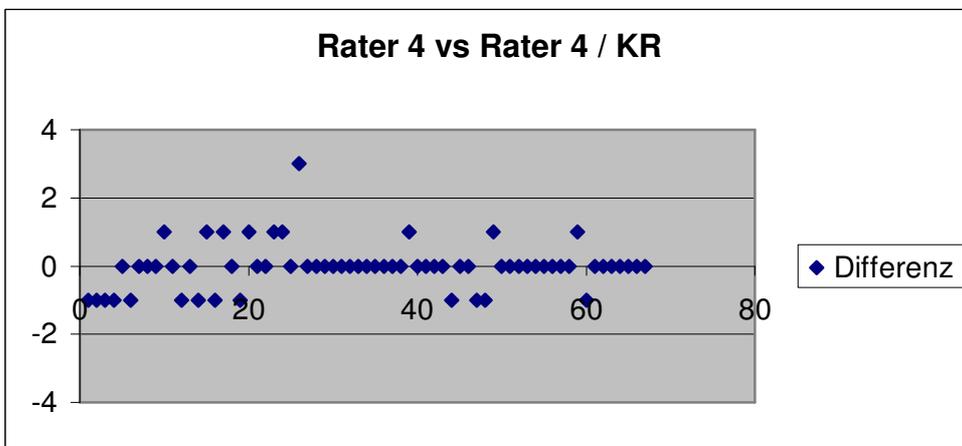
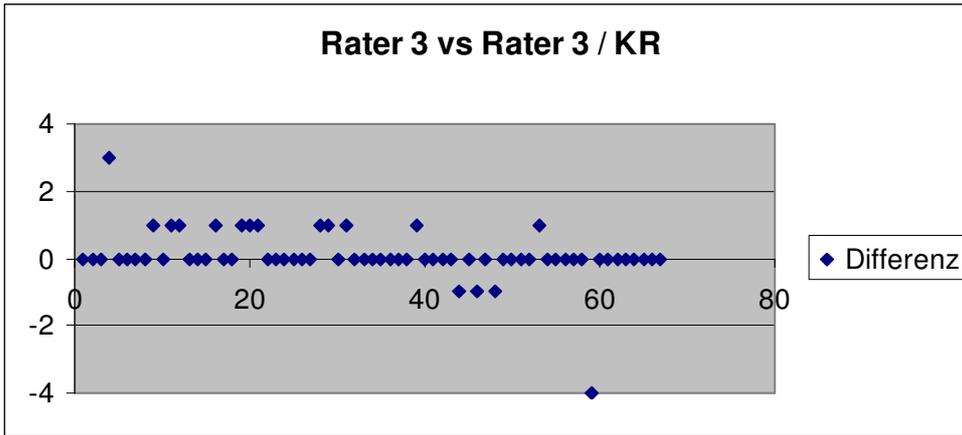
Differenzwerte Klinische Relevanz











## 8.5 Verzeichnis der akademischen Lehrer

Meine akademischen Lehrer in Halle-Wittenberg, Maastricht und Marburg waren die Damen und Herren:

Arnold<sup>3</sup>, Aumüller<sup>3</sup>, Barth<sup>3</sup>, Basler<sup>3</sup>, Baum<sup>3</sup>, Behr<sup>3</sup>, Beitz<sup>1</sup>, Bergmann<sup>1</sup>, Bertalanffy<sup>3</sup>, Beyer<sup>3</sup>, Bien<sup>3</sup>, Blume<sup>1</sup>, Bolm<sup>3</sup>, Bortfeld<sup>1</sup>, Braun<sup>1</sup>, Brilla<sup>3</sup>, Brodde<sup>1</sup>, Brummer<sup>2</sup>, Cetin<sup>3</sup>, Christ<sup>1</sup>, Christiansen<sup>3</sup>, Cobet<sup>1</sup>, Czubayko<sup>3</sup>, Daemen<sup>2</sup>, Daut<sup>3</sup>, Donner-Banzhoff<sup>3</sup>, Effendy<sup>3</sup>, Engel<sup>3</sup>, Fehmann<sup>3</sup>, Feuser<sup>3</sup>, Fischer<sup>1</sup>, Fittkau<sup>1</sup>, Franke<sup>1</sup>, Fricke<sup>1</sup>, Fruhstorfer<sup>3</sup>, Fuhrmann<sup>3</sup>, Gallitelli<sup>1</sup>, Gemsa<sup>3</sup>, Geus<sup>3</sup>, Glanz<sup>1</sup>, Gläser<sup>1</sup>, Göke<sup>3</sup>, Görg<sup>3</sup>, Gotzen<sup>3</sup>, Grau<sup>3</sup>, Gressner<sup>3</sup>, Griez<sup>2</sup>, Griss<sup>3</sup>, Groß<sup>3</sup>, Grzeschik<sup>3</sup>, Haerting<sup>1</sup>, Hamulyak<sup>2</sup>, Happle<sup>3</sup>, Hartung<sup>3</sup>, Hasilik<sup>3</sup>, Hebebrand<sup>3</sup>, Hellinger<sup>3</sup>, Helm<sup>1</sup>, Herzum<sup>3</sup>, Hesse<sup>3</sup>, Heß<sup>3</sup>, Höhne<sup>1</sup>, Höltermann<sup>3</sup>, Hoffmann<sup>3</sup>, Hofmann<sup>3</sup>, Holzhausen<sup>1</sup>, Isenberg<sup>1</sup>, Iwig<sup>1</sup>, Jahreis<sup>1</sup>, Janke<sup>3</sup>, Jaques<sup>3</sup>, Jenderka<sup>1</sup>, Joseph<sup>3</sup>, Junge<sup>3</sup>, Kälble<sup>3</sup>, Kalbfleisch<sup>3</sup>, Kekule<sup>1</sup>, Kern<sup>3</sup>, Kleine<sup>3</sup>, Kleinert<sup>1</sup>, Klemenz<sup>1</sup>, Klenk<sup>3</sup>, Klonisch<sup>1</sup>, Klose<sup>3</sup>, Koehler<sup>3</sup>, Koelsch<sup>1</sup>, Konermann<sup>3</sup>, Koolman<sup>3</sup>, Krause<sup>3</sup>, Kretschmer<sup>3</sup>, Krieg<sup>3</sup>, Kroll<sup>3</sup>, Lammel<sup>3</sup>, Lang<sup>3</sup>, Lange<sup>3</sup>, Lasch<sup>1</sup>, Lautenschläger<sup>1</sup>, Lemberg<sup>3</sup>, Lennartz<sup>3</sup>, Lippert<sup>3</sup>, Lorenz<sup>1</sup>, Lorenz<sup>3</sup>, Ludewig<sup>1</sup>, Lütcke<sup>3</sup>, Machulla<sup>1</sup>, Maisch<sup>3</sup>, Markwardt<sup>1</sup>, Marneros<sup>1</sup>, Marzotko<sup>1</sup>, McGregor<sup>3</sup>, Mischereit<sup>1</sup>, Mogck<sup>3</sup>, Moll<sup>1</sup>, Moll<sup>3</sup>, Moosdorf<sup>3</sup>, Mueller<sup>3</sup>, Müller<sup>1</sup>, Müller<sup>3</sup>, Mutters<sup>3</sup>, Nies<sup>3</sup>, Neubauer<sup>3</sup>, Oertel<sup>3</sup>, Pankow<sup>1</sup>, Penzel<sup>3</sup>, Peschke<sup>1</sup>, Peter<sup>3</sup>, Petermann<sup>3</sup>, Pfab<sup>3</sup>, Pohl<sup>1</sup>, Radke<sup>1</sup>, Radsak<sup>3</sup>, Rawert<sup>3</sup>, Remschmidt<sup>3</sup>, Renz<sup>3</sup>, Reichert<sup>1</sup>, Richter<sup>3</sup>, Röhm<sup>3</sup>, Rosenow<sup>3</sup>, Rothe<sup>1</sup>, Rothmund<sup>3</sup>, Rudolf<sup>1</sup>, Schachtschabel<sup>3</sup>, Schäfer<sup>3</sup>, Schlegel<sup>3</sup>, Scheufler<sup>1</sup>, Schmidt<sup>3</sup>, Schnabel<sup>3</sup>, Schneider<sup>3</sup>, Schneyer<sup>3</sup>, Schönhofer<sup>3</sup>, Schruers<sup>2</sup>, Schüffel<sup>3</sup>, Schuermann<sup>3</sup>, Schultka<sup>1</sup>, Schulz<sup>3</sup>, Schwarz<sup>3</sup>, Seifart<sup>3</sup>, Seitz<sup>3</sup>, Seyberth<sup>3</sup>, Sitka<sup>1</sup>, Slenczka<sup>3</sup>, Slesina<sup>1</sup>, Sommer<sup>3</sup>, Sorger<sup>1</sup>, Sorkau<sup>1</sup>, Steiniger<sup>3</sup>, Stiletto<sup>3</sup>, Stempel<sup>3</sup>, Sturm<sup>3</sup>, Sziegoleit<sup>1</sup>, Thiele<sup>1</sup>, Vogelmeier<sup>3</sup>, Vohland<sup>3</sup>, Voigt<sup>3</sup>, Wagner<sup>3</sup>, Walther<sup>3</sup>, Wasserloos<sup>1</sup>, Weihe<sup>3</sup>, Werner<sup>3</sup>, Westermann<sup>3</sup>, v. Wichert<sup>3</sup>, Wiegandt<sup>3</sup>, Wilhelm<sup>3</sup>, Willhard<sup>1</sup>, Wulf<sup>3</sup>, Wußling<sup>1</sup>, Zernahle<sup>1</sup>, Zielke<sup>3</sup>.

<sup>1</sup> in Halle

<sup>2</sup> in Maastricht

<sup>3</sup> in Marburg

## 8.6 Danksagung

Ich danke Herrn Prof. Dr. med. U. Koehler für die freundliche Überlassung des Themas sowie seine langjährige Unterstützung.

Ein besonderer Dank gilt Herrn Dr. Volker Gross für die professionelle wie freundschaftliche Betreuung und seine unendliche Geduld und Hilfsbereitschaft.

Für die Hilfen und Anregungen im Bereich der statistischen Berechnungen danke ich Herrn Diplom-Psychologen Thomas Ploch.

Des Weiteren danke ich der „AG Lung-Sound“ für ihre tatkräftige Unterstützung sowie den Mitarbeitern der Lungenfunktion der Abteilung Pneumologie (Ambulanz/ Poliklinik) des Universitätsklinikums Marburg für die Hilfe bei der Rekrutierung der Patienten.

Ich danke meinen Freunden und meinem Partner für die stetige Motivation, Rücksichtnahme und das Verständnis.

Besonders möchte ich meinen Eltern danken, die mir das Medizinstudium und die Promotion durch ihre immerwährende Unterstützung erst ermöglicht haben.

Zu guter Letzt danke ich noch allen Korrekturlesern!