

# Interreadervarianz bei der HRCT- und CXR-Befundung in einer Längsschnittstudie bei ehemals asbeststaubexponierten Personen

Von der Medizinischen Fakultät  
der Rheinisch-Westfälischen Technischen Hochschule Aachen  
zur Erlangung des akademischen Grades  
einer Doktorin der Medizin  
genehmigte Dissertation

vorgelegt von

Tanja Maren Illa Carl

aus

Bonn-Bad Godesberg

Berichter: Herr Universitätsprofessor  
Dr. med. Thomas Kraus

Herr Universitätsprofessor  
Dr. med. Rolf Günther

Tag der mündlichen Prüfung: 10. Oktober 2005

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

# Danksagung

Großen Dank möchte ich Herrn Prof. Dr. med. Thomas Kraus für seine ausgezeichnete Betreuung aussprechen und dafür, dass er immer ein offenes Ohr für Fragen hatte, stets gut erreichbar war und für seine unglaublich freundliche, zuversichtliche Art.

Ich danke Frau Nicole Heussen für die statistische Beratung.

Harald Tlatlik gilt großer Dank für die ausgiebigen inhaltlichen Diskussionen, für das Korrekturlesen und für die Hilfen bei der Erlernung von LaTeX und bei der Erstellung von Graphiken.

Desweiteren danke ich Stefan Vogel für die zahlreichen Unterstützungen im Umgang mit LaTeX und Julia Beißner für die Unterstützung in statistischen Fragen.

Zuletzt möchte ich den Mitarbeitern des Institutes für Arbeitsmedizin für die Organisation von Doktorandenseminaren danken und allen, die mit interessierten Fragen und Kritik zu dieser Arbeit beigetragen haben.

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>5</b>
<b>2. Derzeitiger wissenschaftlicher Kenntnisstand</b>	<b>7</b>
2.1. Asbest: Geschichte, Mineralogie, Vorkommen und Verwendung . . . . .	7
2.2. Asbestassoziierte Erkrankungen . . . . .	8
2.2.1. Asbestose . . . . .	10
2.2.2. Pleurafibrose . . . . .	10
2.2.3. Bronchialkarzinom . . . . .	11
2.2.4. Malignes Mesotheliom . . . . .	12
2.2.5. Larynxkarzinom . . . . .	13
2.3. Asbestassoziierte Berufskrankheiten . . . . .	13
2.4. Vorsorge-Untersuchung . . . . .	14
2.4.1. ILO-Klassifikation (für Röntgenbilder) . . . . .	15
2.4.2. CT-Untersuchung . . . . .	19
2.4.3. CT-Klassifikation . . . . .	23
2.5. Interreadervarianz . . . . .	23
2.5.1. Kurze Einführung in die $\kappa$ -Statistik . . . . .	25
2.5.2. Bisherige Studien zur Interreadervarianz . . . . .	29
<b>3. Hypothesen</b>	<b>39</b>
<b>4. Methoden</b>	<b>40</b>
4.1. Kollektiv und Untersuchungsmethoden . . . . .	40
4.1.1. Untersuchungskollektiv . . . . .	40
4.1.2. Demographische Daten . . . . .	40
4.1.3. Untersuchungsaufbau und Untersuchungsmethoden . . . . .	41
4.1.4. Zweitbeurteilung . . . . .	43
4.1.5. Röntgen versus HRCT . . . . .	44
4.2. Kategorisierung . . . . .	46
4.2.1. Röntgen-Kodierungsbogen . . . . .	47
4.2.2. CT-Kodierungsbogen . . . . .	48
4.3. Datenerfassung und -auswertung . . . . .	49
<b>5. Ergebnisse, Diskussion und Ausblick</b>	<b>51</b>
5.1. $\kappa$ -Werte bezogen auf Paar und Fragestellung . . . . .	51
5.2. Ergebnisse: Betrachtung der einzelnen Hypothesen . . . . .	58

---

5.2.1.	Hypothese 1: Interreadervarianz beim Röntgen größer als bei CT?	58
5.2.2.	Hypothese 2: Rö: Interreadervarianz bei der Beurteilung des Parenchyms größer als bei der Pleurabeurteilung? . . . . .	60
5.2.3.	Hypothese 3: CT: Interreadervarianz bei der Beurteilung des Parenchyms größer als bei der Pleurabeurteilung? . . . . .	62
5.2.4.	Hypothese 4: Rö: sinkende Interreadervarianz im zeitlichen Verlauf?	63
5.2.5.	Hypothese 5: CT: sinkende Interreadervarianz im zeitlichen Verlauf?	66
5.3.	Diskussion und Schlussfolgerungen . . . . .	70
5.3.1.	Diskussion des Studiendesigns . . . . .	71
5.3.2.	Allgemeine Diskussion durchschnittlicher $\kappa$ -Werte . . . . .	72
5.3.3.	Diskussion Hypothese 1 . . . . .	73
5.3.4.	Diskussion Hypothese 2 . . . . .	75
5.3.5.	Diskussion Hypothese 3 . . . . .	76
5.3.6.	Diskussion Hypothese 4 und 5 . . . . .	77
5.4.	Zusammenfassung . . . . .	78
5.5.	Ausblick . . . . .	79
<b>A.</b>	<b>Röntgen-Bogen</b>	<b>82</b>
<b>B.</b>	<b>CT-Bogen</b>	<b>83</b>
	<b>Abkürzungsverzeichnis</b>	<b>84</b>
	<b>Tabellenverzeichnis</b>	<b>85</b>
	<b>Abbildungsverzeichnis</b>	<b>87</b>
	<b>Literaturverzeichnis</b>	<b>88</b>

# 1. Einleitung

In der BRD wurde die Verwendung von Asbest 1993 verboten. Durch die oft lange Latenzzeit bis zur Entwicklung asbestbedingter Erkrankungen wird ein Erkrankungsgipfel erst etwa 2015 erwartet. In den letzten Jahren gab es etwa 1.000 asbestbedingte Todesfälle pro Jahr in der BRD, womit Asbest bei den tödlich verlaufenden Berufskrankheiten die weitaus häufigste Todesursache darstellt (50,2 % der Tode der fünf häufigsten zum Tode führenden Berufskrankheiten). Gleichzeitig gehören die Asbesterkrankungen zu den Krebserkrankungen mit den kürzesten Zeitspannen zwischen Erkrankungsbeginn und Tod (Drechsel-Schlund 2003). Vor diesem Hintergrund ist es wichtig, die heutige Vorsorgeuntersuchungsstrategie zu überdenken, um zukünftig eine frühestmögliche Diagnosestellung asbestinduzierter Erkrankungen zu ermöglichen.

Die vorliegende Arbeit basiert auf Daten einer Längsschnittstudie und bearbeitet die Übereinstimmungsraten zweier Beurteiler bei der Befundung von pneumokoniotisch veränderten Röntgenbildern. Sie wurde im Zeitraum von 1993-1997 in sechs Zentren durchgeführt. Es wurde ein Hochrisikokollektiv zusammengestellt, aus dem 636 Personen bereit waren, jährlich einen Vorsorgetermin mit Röntgen- und CT-Diagnostik wahrzunehmen. Die Untersuchungen geschahen unter standardisierten Bedingungen, die Befunde wurden in standardisierte Bögen eingetragen (Röntgen: ILO-1980-Bogen, CT: selbstentworfenen Bogen, in Anlehnung an den ILO-1980-Bogen). Die Bilder wurden jeweils von zwei Beurteilern befundet, wobei der Zweitbeurteiler nicht über die Befunde des Erstbeurteilers informiert war. Insgesamt gab es sieben Beurteiler, die zehn Beurteilerpaare bildeten. Sie hatten sich durch einen großen Erfahrungsreichtum in der Befundung von Pneumokoniosebildern ausgezeichnet. Die Variablen der Bögen wurden zu Kategorien zusammengefasst, anschließend wurden zur Beurteilung der Übereinstimmung  $\kappa$ -Werte errechnet.

Ziel dieser Arbeit war es, folgende fünf Hypothesen zu prüfen:

- 1) Bei der Interpretation von Röntgenbildern wird die Interreadervarianz größer sein als bei der Interpretation von CT-Aufnahmen. CT-Aufnahmen sind leichter zu lesen und führen deshalb zu übereinstimmenderen Aussagen, obwohl bei der Interpretation von Röntgenbildern bezüglich Pneumokoniosen viel mehr Übung besteht.
- 2) Bei Röntgen-Aufnahmen wird die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen größer sein als bei pleuralen Veränderungen. Allgemein lassen sich pneumokoniotische Veränderungen besser im Pleurabereich erkennen als im Parenchym.

- 3) Bei CT-Aufnahmen wird die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen größer sein als bei pleuralen Veränderungen mit der gleichen Begründung wie 2).
- 4) Im zeitlichen Verlauf wird die Interreadervarianz bei der Beurteilung von Röntgenbildern sinken, da ein Lernprozess stattfinden wird.
- 5) Im zeitlichen Verlauf wird die Interreadervarianz bei der Beurteilung von CT-Bildern sinken, da ein Lernprozess stattfinden wird.

Nachdem dem Leser der derzeitige wissenschaftliche Kenntnisstand nähergebracht wird, beschreibt das vierte Kapitel die auf die Hypothesen angewandten Methoden. Das fünfte Kapitel stellt die Ergebnisse dar und enthält die Diskussion der Ergebnisse. Die Arbeit endet mit einer Zusammenfassung und einem Ausblick.

## 2. Derzeitiger wissenschaftlicher Kenntnisstand

### 2.1. Asbest: Geschichte, Mineralogie, Vorkommen und Verwendung

Asbeste sind natürliche, mineralische faserförmige Stoffe. Der Terminus „Asbest“ ist aus dem Griechischen abgeleitet und bedeutet „unauslöschlich“ (Konietzko 1998). Durch seine außergewöhnlichen Eigenschaften wie Hitzebeständigkeit, mechanische Festigkeit, Elastizität, Biugsamkeit, Säurebeständigkeit und außergewöhnliche isolierende Eigenschaften fand Asbest bereits vor 4.000 Jahren Verwendung: die Chinesen und Ägypter stellten Tücher und Matten aus Asbest her, Herodots Berichten zufolge wurden im Altertum Leichengewänder für höhergestellte Persönlichkeiten aus Asbest verwendet. Im Mittelalter geriet dieses Material in Vergessenheit. Erst von Karl dem V. wird berichtet, dass er flammenwiderstehende Gewebe nutzte, die scheinbar mit Asbest durchwebt waren (Konietzko 1998). Seit der Mitte des 19. Jahrhunderts wurde Asbest weltweit in großen Mengen verarbeitet: seit 1860 wurden über 100 Mio. Tonnen Asbest gefördert (Förster 1998) und fanden in über 3.000 Bereichen Anwendung (Kraus et al. 1998; Levin et al. 2000). Asbestenthaltendes Gestein wurde und wird hauptsächlich in Kanada, Südafrika, China, den USA und der Sowjetunion (Konietzko 1998) abgebaut und in Fabriken stufenweise zu Fasern verarbeitet.

Asbest ist der kommerzielle (nicht mineralische) Sammelbegriff für eine Familie von sechs natürlich vorkommenden faserförmigen silikatischen Mineralien sehr unterschiedlicher Natur (Konietzko 1998). Nach ihrer Morphologie werden die Fasern in zwei Gruppen eingeteilt: Amphibole (gerade, stäbchenförmig) und Serpentine (gewunden). Die Struktur der Serpentine ist nicht so stabil und widerstandsfähig wie die der Amphibole, was für den menschlichen Körper insofern von Bedeutung ist, als Serpentine abgebaut und mit sehr langer Latenzzeit eliminiert werden können, während Amphibole dauerhaft pathogen wirkend im Körper verbleiben (Müller et al. 1996; Levin et al. 2000). Allgemein sind jedoch alle Asbestfasertypen in der Lage, alle asbestassoziierten Erkrankungen hervorzurufen (Levin et al. 2000). In Deutschland wurden hauptsächlich Weißasbest (Chrysotil) und Blauasbest (Krokydolith) verarbeitet. Das serpentinische Asbest Chrysotil hatte einen Anteil von 90-95 % am deutschen Einsatz, während das amphibole Krokydolith nur einen Anteil von 5-10 % hatte (Müller et al. 1996; Kraus et al. 1998).

Asbestfasern finden sich überall in der Luft. Einmal in die Luft gelangte feine Asbestfasern bleiben über viele Stunden hinweg in der Luft, selbst unter windstillen Bedingungen. Schon leichte Luftbewegungen wirbeln Asbestfasern, die sich auf Oberflächen

abgelagert haben, wieder auf (Levin et al. 2000). In Reinluftgebieten (Harz) lassen sich 200 Fasern pro Kubikmeter Luft nachweisen. An Großstadtkreuzungen ist der Gehalt bereits fünfmal so hoch. Man kann von einer 100- bis 10.000fach erhöhten Exposition an gefährdeten Arbeitsplätzen gegenüber der allgemeinen Umweltbelastung ausgehen (Müller et al. 1996). Auch im Trinkwasser befinden sich Asbestfasern. In Deutschland findet man Konzentrationen von 1-10 Mio. Fasern pro Liter (in den USA sind es stellenweise 300 Mio.). Es konnte bislang jedoch keine erhöhte Inzidenz von Krebserkrankungen des Verdauungstraktes nachgewiesen werden. Diskutiert wird eine mögliche Gefahr durch Luftbefeuchter, die asbestfaserhaltiges Trinkwasser vernebeln und dadurch wiederum die Luftbelastung erhöhen (Müller et al. 1996).

In Deutschland wurde der Verarbeitungsgipfel von asbestfaserhaltigen Produkten Ende der 70er Jahre mit annähernd 200.000 t Asbest pro Jahr erreicht (Woitowitz et al. 1981). In den 80er Jahren ließ sich eine starke Senkung infolge von Arbeitsschutzmaßnahmen und vermehrtem Einsatz von Ersatzstoffen verzeichnen. In der DDR lag das Maximum 1980 bei 74.000 t, danach nahm der Import nur leicht ab. 1984 wurde dort der Import von Krokydolith verboten. Etwa 70 % des Asbestes wurden zu Asbestzementprodukten verarbeitet, der Rest verteilte sich auf etwa 3.000 weitere Anwendungsgebiete (Levin et al. 2000), darunter die Herstellung von Brems- und Kupplungsbelägen, Textilien, Fußbodenbelägen, Hochdruckdichtungen, Spritzmassen, bauchemischen Produkten und Isolationen (Kraus et al. 1998). Letztlich waren fast alle industriellen und handwerklichen Berufe asbeststaubbelastet (Kraus et al. 1998) und man geht davon aus, dass mehr als 500.000 Personen im Laufe ihres Berufslebens Asbeststaub ausgesetzt waren und darum einer Risikogruppe zugerechnet werden können, zumal der TRK-Wert an den Arbeitsplätzen noch Anfang der 80er Jahre oft überschritten wurde (Kraus et al. 1998).

Seit 1993 besteht laut Gefahrstoffverordnung ein generelles Asbestverbot in Deutschland: sowohl Verarbeitung als auch Inverkehrbringen sind nicht mehr gestattet (Kraus et al. 1998). Im Baubereich wurde Asbest vollständig durch geeignete Substitute verdrängt, allerdings besteht weiterhin eine Gefährdung bei der Sanierung asbesthaltiger älterer Gebäude. Mehrere Wochen nach Asbestsanierungen wurden Spitzenkonzentrationen von bis zu 900.000 Fasern pro Kubikmeter Luft gemessen (Müller et al. 1996). Der Grenzwert für Asbestfaserblastungen der Luft bei Abbruch-, Sanierungs- und Instandhaltungsarbeiten liegt seit 1995 bei 15.000 Fasern pro Kubikmeter Luft (Kraus et al. 1998).

## 2.2. Asbestassoziierte Erkrankungen

Durch Asbest bedingte Erkrankungen wurden in der modernen Medizin erst sehr spät bekannt. Es wird angenommen, dass Montago Merry 1899 den ersten Asbestose-Fall beschrieb (Konietzko 1998). Bereits 1918 weigerten sich einige Krankenkassen in den USA und Kanada, asbeststaubexponierte Personen zu versichern, weil unter ihnen die Erkrankungszahl deutlich erhöht war (Levin et al. 2000). 1936 wurde die „schwere Asbeststaublungenkrankheit (Asbestose)“ in Deutschland in die Liste der entschädigungspflichtigen Berufskrankheiten aufgenommen (Drechsel-Schlund 2003).



Das derzeitige Wissen über die gesundheitlichen Folgen von Asbeststaubexposition basiert auf epidemiologischen und arbeitsmedizinisch-klinischen Erkenntnissen. Fasern mit einer Länge von 0,5 bis 5  $\mu\text{m}$  gelangen bis tief in die Lungenbläschen (Alveolen) und provozieren die dort vorhandenen Makrophagen und Fibroblasten. Die Makrophagen setzen nach der Phagozytose freier Asbestfasern Entzündungsmediatoren frei, die andere Zellen des Immunsystems anlocken, sind aber nicht in der Lage, die Fasern abzubauen. Sowohl die Entzündungsmediatoren, als auch die freien Asbestfasern schädigen die Alveolarepithelzellen. Es entwickelt sich eine chronisch-entzündliche Situation, die auf Dauer in einer Zunahme des Makrophagenvolumens, polymorphkerniger Leukozyten, Fibroblasten und extrazellulärer Matrix resultiert. Dieser fibrotische Prozess schreitet fort und führt nach und nach zu einer versteiften, verkleinerten Lunge mit vermindertem Gasaustauschvermögen (Levin et al. 2000). Einige Asbestfasern migrieren durch das Lungengewebe zur Pleura, wahrscheinlich durch lymphatische Spalten. Auch dort reichern sie sich an und schädigen auf Dauer die Pleura. Ein Teil der Asbestfasern gelangen anschließend mit der Lymphe in andere Körperregionen (Levin et al. 2000). Auf diese Weise können Asbestfasern zu einer Fibrose des Lungenparenchyms („Asbestose“) oder der Pleura (Plaques und Verkalkungen) und/oder zur Entstehung von Karzinomen führen. Das kanzerogene Potenzial von Krokydolith wird aufgrund seiner höheren Biobeständigkeit und seiner besonderen Fasergeometrie höher eingestuft als das des Chrysotils (Kraus et al. 1998).

Folgende Karzinome werden als Folge von Asbestexposition anerkannt: das Bronchialkarzinom (Lungenkrebs), das maligne Mesotheliom der Pleura, des Peritoneums und des Perikards und das Larynx-Karzinom (Kehlkopfkrebs) (Kraus et al. 1998; Berry et al. 2000). Umstritten sind die Fragen nach einem Zusammenhang zwischen Asbestexposition und Non-Hodgkin-Lymphom/Plasmozytom bzw. Colon-Karzinom.

Die Asbest-Exposition erhöht das relative Risiko, an einer asbestassoziierten Krankheit zu erkranken um ein Vielfaches, doch nicht jeder Exponierte erkrankt. Die Erkrankungswahrscheinlichkeit wird durch die Intensität und Dauer der Exposition, die Faserart, die individuelle Disposition und das Vorhandensein von Kokarzinogenen beeinflusst (Müller et al. 1996; Albin et al. 1996; Kraus et al. 1998). Man geht davon aus, dass jede asbestassoziierte Krankheit unabhängig von anderen asbestassoziierten Krankheiten auftreten kann, z.B. ein Bronchialkarzinom unabhängig von einer Asbestose (Levin et al. 2000). Auch nach einem Stopp des äußeren Asbest-Einflusses kann eine asbest-assoziierte Krankheit fortschreiten, da die bereits inhalierten Fasern in der Lunge persistieren und weiterhin permanent Entzündungsreize hervorrufen (Kiesel et al. 1995; Levin et al. 2000).

Aufgrund sehr langer Latenzzeiten zwischen Expositionsbeginn und dem Ausbruch der Krankheiten wird der Erkrankungsgipfel frühestens ab 2015 erwartet (Kjærgaard et al. 2000). Französische Forscher gehen von einem Gipfel zwischen 2030 und 2040 aus. In Frankreich erwartet man 1140-1300 Mesotheliom-Tote pro Jahr in diesem Zeitraum (Banaei et al. 2000) .

### 2.2.1. Asbestose

Asbestfasern werden eingeatmet und reichern sich in der Lunge an. Dort lösen sie Entzündungsreaktionen aus und führen letztlich zu einer verstärkten Bindegewebsbildung der Lunge (Fibrose). Es handelt sich hierbei um eine symmetrische, disseminierte, diffuse interstitielle Fibrose des Lungenparenchyms, die sich bevorzugt basal subpleural findet (Kraus et al. 1998; Levin et al. 2000). Voraussetzung für eine starke Reaktion ist eine langjährige und intensive Asbestexposition (Koskinen et al. 1998), aber auch kurze, besonders intensive Expositionen sollen im Einzelfall eine Asbestose auslösen können (Levin et al. 2000). Auch nach Expositionsende kann die Krankheit weiter fortschreiten (s. Abs. 2.2). Die Latenzzeit bis zur klinisch manifesten Fibrose beträgt im Durchschnitt 15-20 Jahre (Levin et al. 2000). Klinisch kündigt sich die Asbestose meist mit chronischem Reizhusten an, der von Belastungsdyspnoe und später sogar Ruhedyspnoe gefolgt werden kann. Manchmal treten Brustschmerz und auskultatorisch ein Knisterrasseln auf, bei der Lungenfunktionstestung kann eine restriktive Ventilationsstörung mit Gasaustauschstörungen nachweisbar sein. In späten Stadien können ein chronisches unspezifisches respiratorisches Syndrom mit obstruktiver Ventilationsstörung und eine Rechtsherzbelastung auftreten (Kraus et al. 1998; Levin et al. 2000). Röntgen-Thoraxaufnahmen in posterior-anterior (p.a.)-Orientierung sind ein wichtiges diagnostisches Mittel. Sie werden gemäß der ILO-Staublungenklassifikation von 1980 beurteilt und kategorisiert (s. Abs. 2.4.1). Abbildung 2.2 zeigt röntgenologische Veränderungen, die sich bei asbestbedingten Pneumokoniosen finden (Lange 1986). Im Zeitraum von 1980 bis 2000 wurden in Deutschland 18.250 Asbestosen anerkannt, davon seit 1995 etwa 2000 jährlich (Drechsel-Schlund 2003).

### 2.2.2. Pleurafibrose

Die Pleurafibrose in ihren unterschiedlichen Ausprägungen ist die häufigste Konsequenz einer Berufsbelastung durch Asbest. Mit einer durchschnittlichen Latenzzeit von 30-40 Jahren treten folgende Reaktionsmuster auf (Müller et al. 1996):

1. Bei der **Asbestpleuritis** finden sich rezidivierende, unspezifische, fibrinreiche Pleuraergüsse nahezu ohne Entzündungszellen, gegebenenfalls mit charakteristischen Rundherdatelektasen. Diese Pleuraergüsse sind oft die ersten asbeststaubassoziierten Veränderungen, noch vor der Entwicklung einer Asbestose, und treten häufig bereits in den ersten zehn Jahren nach Expositionsbeginn auf (Levin et al. 2000).
2. Die **Hyalinosis complicata** der Pleura parietalis mit einer Dicke von 0,2 bis 1 cm tritt nach einer Pleuritis auf und geht manchmal mit einer diffusen, oft beidseitigen (meist Unter- oder Mittellappen-) Fibrose der Pleura pulmonalis einher, sowie Pleuraplaques im Brustwand- oder Zwerchfellbereich.
3. **Hyaline** (oder im fortgeschrittenen Stadium auch verkalkte) **Pleuraplaques** können als Brückenbefund angesehen werden; ca. 70 % dieser charakteristischen, meist beidseitigen Plaques sind Folgen einer Asbeststaubexposition.

Im Gegensatz zur Asbestose, wo ein klarer Zusammenhang zwischen Erkrankung und Expositionsintensität besteht, spielt bei der Entstehung der Pleuraplaques die Asbestmenge keine ausschlaggebende Rolle (Koskinen et al. 1998). In der Allgemeinbevölkerung wird eine Prävalenz hyaliner Plaques von weniger als 1,2 % angenommen, bei beruflich asbeststaub-exponierten Kollektiven fanden sich Werte bis zu 42 % (Kraus et al. 1998). Daher werden Pleuraplaques als Marker einer früheren Asbestexposition angesehen (Levin et al. 2000). Mögliche andere ätiologische Faktoren wie pleurale Residuen nach Rippenfrakturen oder Tuberkulose sollten bei der Diagnosestellung berücksichtigt werden (Kraus et al. 1998). Unter den Pleuraverdickungen können sich kleine Atelektaseherde oder Pseudotumoren entwickeln, die meist kleiner als 2 cm im Durchmesser sind. Obwohl sie oft eine charakteristische Form aufweisen, besteht die Gefahr einer Fehleinstufung, da sich dahinter auch ein beginnender Lungentumor verbergen kann (Levin et al. 2000). Einige Forscherteams haben dargestellt, dass bereits kleine Pleuraläsionen die Lungenfunktion einschränken (Levin et al. 2000), andere gehen davon aus, dass nur bei ausgedehnter Verschwartung eine restriktive Ventilationsstörung erfolgt (Kraus et al. 1998). Pleuraveränderungen lassen sich in p.a.-Aufnahmen nachweisen und müssen nach der ILO-Staublungen-Klassifikation kategorisiert werden (s. Abs. 2.4.1).

### 2.2.3. Bronchialkarzinom

Das Bronchialkarzinom (Lungenkrebs) stellt die verbreitetste asbestverursachte Neoplasie und die Haupttodesursache durch Asbest in den entwickelten Ländern dar (Levin et al. 2000). Je nach Quelle wird mit Latenzzeiten von 20 bis 39 Jahren gerechnet (Baur et al. 1994; Levin et al. 2000; Puntoni et al. 2001). Das asbestassoziierte Bronchialkarzinom unterscheidet sich in keiner Hinsicht (morphologisch, topologisch oder histologisch) von Bronchialkarzinomen anderer Genese (Brodkin et al. 1997; Levin et al. 2000).

Das Lungenkrebsrisiko ist bei bestehender Asbestose deutlich erhöht (Kraus et al. 1998). Das Zusammentreffen von diffuser interstitieller Lungenfibrose und Lungenkrebs ist häufiger (53 % eines Untersuchungskollektiv) als das Auftreten von Lungenkrebs ohne diffuse interstitielle Lungenfibrose (15 % des selben Kollektivs) (Katabami et al. 2000). Dennoch kommen neuere Studien zu dem Ergebnis, dass das Lungenkrebsrisiko auch ohne radiologisch oder histologisch nachweisbare Asbestose erhöht ist (Müller et al. 1996; de Klerk et al. 1996; Finkelstein et al. 1997; Brodkin et al. 1997; Levin et al. 2000). Ätiologisch spielt inhalatives Rauchen als ein äußerst potentes Kokarzinogen eine wichtige Rolle (Popp 1996; Müller et al. 1996; de Klerk et al. 1996; Brodkin et al. 1997; Levin et al. 2000). Während das relative Risiko eines Asbeststaubexponierten an einem Bronchialkarzinom zu erkranken gegenüber der Allgemeinbevölkerung um das ca. fünffache erhöht ist, ist es bei gleichzeitigem Zigarettenrauch um das ca. 50fache erhöht (Hausmann et al. 2000). Da das Aufgeben des Rauchens unter asbestexponierten Arbeitern mit einem gesenkten Krebsrisiko verbunden war, wenn auch nicht das Level von Nie-Rauchern erreicht wurde, sollte die Aufklärungsaufgabe des Arztes nicht unterschätzt werden (Levin et al. 2000). Vermutlich sind auch Pleuraplaques mit einem erhöhten Lungenkrebsrisiko verbunden (Kraus et al. 1998).

Die Behandlung des asbestassoziierten Bronchialkarzinoms entspricht der Behandlung

von Bronchialkarzinomen anderer Genese, da sie sich nicht voneinander unterscheiden. Die 5-Jahres-Überlebensrate liegt derzeit noch bei weniger als 10 %, weil meist erst Spätstadien diagnostiziert werden. (Levin et al. 2000). Im Zeitraum von 1980 bis 2000 wurden in Deutschland 6.797 Bronchial- und Larynx-Karzinome anerkannt, davon 693 alleine im Jahr 2000 (Drechsel-Schlund 2003).

#### 2.2.4. Malignes Mesotheliom

Das Mesothel ist ein einschichtiges Plattenepithel, das Bauch- und Brusthöhle auskleidet. Dementsprechend gibt es tumoröse Veränderungen (Mesotheliome) in der Pleura (Brustfell), im Perikard (Herzbeutel) oder im Peritoneum (Bauchfell). Durch Asbest ausgelöste Mesotheliome sind zu über 90 % Pleuramesotheliome (Kraus et al. 1998). Das diffuse Pleuramesotheliom wird als seltener Tumor beschrieben (Levin et al. 2000): Anfang der 80er Jahre fand man zwei Pleuramesotheliome bei 1.000 Obduktionen (Kraus et al. 1998). Nach einer beruflichen Asbestexposition ist das relative Risiko, ein Pleuramesotheliom zu entwickeln mehrere hundertfach erhöht (Hausmann et al. 2000). Man geht davon aus, dass heutzutage 70-90 % der Pleuramesotheliome auf eine Asbeststaubexposition zurückzuführen sind (Müller et al. 1996; Kraus et al. 1998; Havemann et al. 1997). Darum wird das Mesotheliom auch als ein Signaltumor der Asbeststaubexposition angesehen (Levin et al. 2000; Hausmann et al. 2000). Es manifestiert sich mit einer durchschnittlichen Latenzzeit von 20 bis 40 Jahren (Albin et al. 1996; Levin et al. 2000). Auch deutlich kürzere Latenzzeiten von unter zehn Jahren konnten in Einzelfällen nachgewiesen werden (Puntoni et al. 2001). Besonders bei Krokydolith scheint eine kurze Expositionsdauer zu reichen, um die Entwicklung eines Mesothelioms zu verursachen (Kraus et al. 1998), in Einzelfällen nur wenige Wochen (Müller et al. 1996). Weder Zigarettenrauchen noch eine starke Pleuraverdickung erhöhen das Risiko, ein Mesotheliom zu entwickeln (Hillerdal 1994; Rudd 1996; Levin et al. 2000). Im Frühstadium finden sich viele kleine Knötchen in der Pleura, die später zu bis zu mehrere Zentimeter dicken Tumorplatten konfluieren können (Müller et al. 1996). Die Diagnostik früher Stadien ist derzeit noch unbefriedigend (Kraus et al. 1998). Häufigstes Leitsymptom ist ein Pleuraerguss (Kraus et al. 1998), oft liegen Brustschmerzen und Dyspnoe vor (Levin et al. 2000).

Die Prognose eines Mesothelioms ist besonders schlecht. Die durchschnittliche Überlebenszeit beträgt für Pleuramesotheliome 8-12 Monate (Ribak et al. 1992; Levin et al. 2000) und für Peritonealmesotheliome 4-7 Monate (Ribak et al. 1992). In den letzten Jahren scheint sich die Dreifach-Therapie: Operation, Chemotherapie und Bestrahlung gegenüber anderen Therapien bewährt zu haben (Levin et al. 2000). Patz et al. fordern ein besseres Staging-System, um die Therapieerfolge verschiedener Therapieansätze besser vergleichen zu können (Patz et al. 1996). Im Zeitraum von 1980 bis 2000 wurden in Deutschland 6.663 maligne Mesotheliome anerkannt, davon 660 alleine im Jahr 2000 (Drechsel-Schlund 2003).

### 2.2.5. Larynxkarzinom

Über den Kausalzusammenhang zwischen Larynxkarzinom (Kehlkopfkrebs) und Asbestexposition gibt es starke Kontroversen. Das Larynxkarzinom ist ein relativ seltener Tumor mit einer Inzidenz in der Allgemeinbevölkerung von 4-7/100.000 Einwohner. Die in Studien nachgewiesenen leicht erhöhten relativen Risiken asbeststaubexponierter Personen können verzerrt sein. Zum Beispiel wurden Rauch- und Trinkgewohnheiten meistens nicht hinreichend berücksichtigt (Kraus et al. 1998). Die Morphologie fraglich asbeststaubassoziiertes Larynxkarzinome weist keine Besonderheiten auf: Symptomatik, Morphologie und Histologie sind wie bei Larynxkarzinomen anderer Genese (Kraus et al. 1998). Dementsprechend erfolgt auch die Therapie wie bei Larynxkarzinomen anderer Genese (Levin et al. 2000). Bereits in der DDR wurden 76 Larynxkarzinomfälle als asbestassoziierte Berufskrankheit anerkannt, auch in der BRD geschah dies in Einzelfällen (Kraus et al. 1998). Seit der Neufassung der Berufskrankheitenverordnung vom 31. Oktober 1997 werden Larynxkarzinomerkrankungen unter den gleichen Voraussetzungen wie Bronchialkarzinomerkrankungen als Berufskrankheiten anerkannt (Kraus et al. 1998; Wiethage et al. 1999). Bis zum Jahr 2000 wurden in Deutschland bereits 229 Fälle anerkannt (Drechsel-Schlund 2003).

## 2.3. Asbestassoziierte Berufskrankheiten

Eine Berufskrankheit ist eine durch besondere Einwirkungen bei beruflichen Tätigkeiten verursachte, meist chronische Erkrankung. Sie gilt als Versicherungsfall im Sinne der gesetzlichen Unfallversicherung und verpflichtet den Unfallversicherungsträger zu medizinischer Behandlung, Rehabilitationsmaßnahmen und Entschädigungsleistungen. Bei Verdacht auf das Vorliegen einer Berufskrankheit besteht Anzeigepflicht beim Unfallversicherungsträger oder Gewerbearzt. Eine Entschädigung setzt eine Minderung der Erwerbsfähigkeit um mind. 20 % voraus (Pschyrembel 1998). Im Jahr 2002 ist die neue Berufskrankheiten-Verordnung § 9 Abs. 6 des Sozialgesetzbuches VII in Kraft getreten (s. Abb. 2.1)

*Tabelle 2.1.: Asbestbedingte Berufskrankheiten in der Liste der Berufskrankheiten 2002*

<b>BK-Nr. 4103</b>	Asbeststaublungerkrankung (Asbestose) oder durch Asbeststaub verursachte Erkrankung der Pleura
<b>BK-Nr. 4104</b>	Lungenkrebs oder Kehlkopfkrebs <ul style="list-style-type: none"> <li>• in Verbindung mit Asbeststaublungerkrankung (Asbestose)</li> <li>• in Verbindung mit durch Asbeststaub verursachter Erkrankung der Pleura oder</li> <li>• bei Nachweis der Einwirkung einer kumulativen Asbestfaserstaub-Dosis am Arbeitsplatz von mindestens 25 Faserjahren (<math>25 \cdot 10^6 [\frac{\text{Fasern}}{\text{m}^3} \cdot \text{Jahre}]</math>)</li> </ul>
<b>BK-Nr. 4105</b>	Durch Asbest verursachtes Mesotheliom des Rippenfells, des Bauchfells und des Perikards.

Asbestbedingte Berufskrankheiten nahmen im Jahr 2000, hinter der Lärmschwerhörigkeit mit 6.872 anerkannten Fällen, Platz zwei der anerkannten Berufskrankheiten mit 3.259 Fällen ein. Gefolgt wurden sie von Hautkrankheiten mit 1.680 Fällen (BK-DOK '99 1999). Der seit 1982 zu verzeichnende deutliche Aufwärtstrend des Anteils von Krebserkrankungen unter allen anerkannten Berufskrankheiten ist vor allem auf den steilen Anstieg der Mesotheliome und der asbestassoziierten bösartigen Lungentumoren zurückzuführen (Wiethege et al. 1999). Unter den anerkannten berufsbedingten Krebserkrankungen haben die asbestbedingten von 1978 bis 2000 mit 14.079 einen Raum von 69,8 % eingenommen. Im Jahr 2000 gab es insgesamt 1.785 Krebsfälle, von denen 1.353 asbestbedingt waren, was 75,8 % entspricht (Butz 2002; Drechsel-Schlund 2003). Durch die lange Latenzzeit wird die Zahl jährlich weiterhin steigen. Alleine 2000 wurden 3.259 Erstentschädigungen durch die gesetzliche Unfallversicherung vorgenommen, 2.059 mehr als vier Jahre zuvor (BK-DOK '99 1999; Kraus et al. 1998). Wenn man davon ausgeht, dass für jede erkrankte Person Kosten von etwa 250.000 € anfallen, wird auch der sozio-ökonomische Faktor für die Träger der Unfallversicherungen deutlich (Kraus et al. 1998).

## 2.4. Vorsorge-Untersuchung

Im Jahr 2003 waren in Deutschland ca. 500.000 früher asbeststaubexponierte Personen registriert (ZAs 2003). Die Relevanz effizienter Vorsorgeuntersuchungen wird deutlich, wenn man sich vor Augen hält, dass der Malignomgipfel erst frühestens ab 2015 erwartet wird (s. Abs. 2.2). Abgesehen vom persönlichen Leid der betroffenen Personen entstehen durch asbestverursachte Malignome auch hohe Kosten für den Unfallversicherungsträger (s. Abs. 2.3). Effiziente Vorsorgeuntersuchungen haben verschiedene Vorzüge. Zum einen kann die Krankheit bereits zu einem Zeitpunkt entdeckt werden, wo ein heilendes Eingreifen noch möglich ist, zum anderen können die betroffenen Personen über Jahre hinweg begleitet und über die neuesten Erkenntnisse aufgeklärt werden (z.B. über die Relevanz des Einstellens der Rauchgewohnheiten). Desweiteren stellt die Vorsorgeuntersuchung eine größere Sicherheit dar, dass die Arbeiter medizinisch überwacht werden und gegebenenfalls Kompensationen erhalten (Levin et al. 2000).

Sowohl aktuell durch Asbest gefährdete Arbeitnehmer als auch ehemals exponierte Personen müssen bei den Berufsgenossenschaften (BG) gemeldet werden. Bereits 1972 haben die Berufsgenossenschaften vereinbart, nachgehende Untersuchungen sicherzustellen, und dafür in Augsburg die weltweit einzigartige Einrichtung einer Zentrale Erfassungsstelle (ZAs) bei der Textil-Berufsgenossenschaft eingerichtet. Sie überwacht die Durchführung der Untersuchungen nach dem Berufsgenossenschaftlichen Grundsatz G 1.2 „Asbestfaserhaltiger Staub“ seit Anfang der 70er Jahre. 2003 waren dort 490.912 Personen gemeldet, die aus 49.698 Betrieben stammten. Davon sind 61.103 Personen aktuell durch Instandhaltungs- und Abriss-Arbeiten gefährdet (ZAs 2003). Dem Versicherten werden freiwillige Vorsorgeuntersuchungen in Abständen von höchstens 5 Jahren angeboten, die er aber nicht wahrnehmen muss. Da Frühveränderungen in Röntgenbildern oft schwer zu beurteilen sind, gibt die Zentrale Erfassungsstelle gegebenenfalls eine

Zweitbeurteilung bei besonders qualifizierten Ärzten in Auftrag. Eine Auswahl dieser Zweitbeurteiler wirkte bei der hier vorliegenden Studie mit. Die Vorsorgeuntersuchung besteht aus Anamnese, körperlicher Inspektion, p.a.-Thorax-Röntgenaufnahme, die gemäß der ILO-1980 Staublungenklassifikation beurteilt und eingestuft wird und einem Lungenfunktionstest. In Abb. 2.2 werden die röntgenologischen Veränderungen bei asbestbedingten Pneumokoniosen dargestellt. Ende der 90er Jahre fanden etwa 40.000 Vorsorgeuntersuchungen pro Jahr statt, die Zahl stieg und steigt seit der Wiedervereinigung stetig (Kraus et al. 1998).

**Tabelle 2.2.:** Röntgenologische Veränderungen bei asbestbedingten Pneumokoniosen (Lange S 1986)

- kleine Winkelergüsse
- kleine irreguläre Fleckschatten (bei Silikose sind sie rund), die basal beginnen
- selten massive Fibrosen (bei Silikose häufig)
- Kugelatektase, „Asbestpseudotumor“
- tafelfergartige Pleuraplaques, die ab einer Größe von ca. 3 mm sichtbar sind
- flächige, diffuse Pleuraverdickungen, wenn sie tangential vom Röntgenstrahl getroffen werden
- verwaschene Herz- und Pleurakonturen durch pleuroparenchymatöse Fibrose
- verdickte Septen
- kalkdichte Linien (auf Pleuraverdickungen) parallel zur Zwerchfellkontur oder Thoraxwand. Die Plaques sind typisch, kommen aber auch bei anderen Erkrankungen vor: Talkuminalation, Hämatothorax, Emphyem oder therapeutischen Pneumothorax wegen Tuberkulose.

Auf den bisherigen Erfahrungen basierend existieren Kriterien, nach denen der Verdacht auf eine asbeststaubassoziierte Erkrankung geäußert und an den Unfallversicherungsträger gemeldet werden muss (Kraus et al. 1998). Diese sind in Tab. 2.3 und Tab. 2.4 dargestellt. Das bisherige Konzept der Vorsorgeuntersuchungen hat sich bezüglich der Sekundärprävention bereits klinisch relevanter Asbeststaublungen-Erkrankungen bewährt, ist aber nicht aussagekräftig genug bei der Frühdiagnostik von Lungenkrebs-erkrankungen und Mesotheliomen. Da der Malignom-Gipfel erst ab 2015 erwartet wird, wurde der Einsatz anderer Verfahren mit höherer Sensitivität wie die Computertomographie (CT) geprüft, um Frühstadien zu einem Zeitpunkt zu erkennen, wo ein heilender Eingriff noch möglich sein könnte. Auch erscheint der bisherige Untersuchungsabstand von häufig fünf Jahren als möglicherweise zu groß (Kraus et al. 1998).

### 2.4.1. ILO-Klassifikation (für Röntgenbilder)

1950 entwarf die International Labor Organization (ILO) erstmals einen Fragebogen, der der Erfassung radiologischer Erscheinungen der Staublung (Pneumokoniose) diene. 1980 wurde er zum vierten Mal überarbeitet und ist in dieser Form noch heute weltweit



**Tabelle 2.3.:** BK-Anzeige: Rö-Parenchym; Kriterien für die Erstattung einer ärztlichen BK-Anzeige gemäß BK-Nr.4103 -basierend auf Röntgenbefunden und körperlicher Untersuchung; s, t, u: unregelmäßige kleine Schatten; Dichte 1/0, 1/1: s. Abs 2.4.1

Röntgenbefund der Lungen nach ILO-Klassifikation 1980		Auskultations- und Lungenfunktionsbefund
Dichte der Schatten	Form	
a) 1/0	s,t bzw. u	Knisterrasseln und/oder VKI $\leq 90$ % von VKS nach EGKS Mindestsollwert unter BTPS-Bedingungen
b) $\geq 1/1$	s,t bzw. u	auch wenn klinisch keine Auffälligkeiten und keine Einschränkungen der VKI messbar sind

**Tabelle 2.4.:** BK-Anzeige: Rö-Pleura; Kriterien für die Erstattung einer ärztlichen BK-Anzeige gemäß BK-Nr.4103 -basierend auf Röntgenbefunden und körperlicher Untersuchung

Pleuraveränderung	Dicke/Ausdehnung	Anamnese
hyaline Pleuraplaques	Dicke $\geq 3$ mm und/oder Ausdehnung $\geq 2$ cm	anamnestische Hinweise auf Asbeststaubexposition
verkalkte Pleuraplaques	unabhängig	
Hyalinosis complicata bzw. Pleuraerguss	unabhängig	
„diffuse“ Pleura-Verdickung	Dicke $\geq 2 - 3$ mm im Bereich der Mittel-/Unterfelder	

akzeptiert zur Aufzeichnung radiologischer Veränderungen bei asbeststaubverursachten Lungen- und Pleuraerkrankungen (Welch et al. 1998). In den letzten Jahren wurde die ILO-2000 in Kontinuität zur ILO-1980 entwickelt. Die definitive Umstellung wird für das Jahr 2004 erwartet. Die Klassifikation umfasst einen Satz von Standardfilmen, den geschriebenen Text und eine Anzahl von Anmerkungen. Gegebenenfalls kann der Standardfilm vorrangig gegenüber dem definierenden Text sein (Internationales Arbeitsamt Genf 1981). Im Anhang A wird ein ILO-1980-Bogen abgebildet. Auf ihm werden die Befunde der deutschen Vorsorgeuntersuchungen eingetragen. Im folgenden werden die Grundzüge des ILO-1980-Bogens dargelegt. Für verschiedene Lungen- und Pleurabefunde wird einzeln bewertet, welche Form, Größe, Streuung und Verbreitung vorliegen.

### Lungenbefunde:

- kleine Schatten

Streuung: Die Kategorie wird durch Abschätzung des Streuungsgrades im Vergleich mit den Standardfilmen erhalten. Kategorie 0 bedeutet, dass keine kleinen Schatten vorhanden sind oder sie weniger gestreut sind als es der unteren Grenze von Kategorie 1 entspricht. Die Kategorien 1 bis 3 bezeugen eine zunehmende Streuungsdichte. Für diese vier großen Kategorien sind Standardfilme erhältlich. Eine 12-Stufenskala (0/- bis 3/+) ermöglicht eine differenziertere Einschätzung der Streuung. Die Aufnahme wird mit den Standardfilmen verglichen und in eine der vier Kategorien eingeteilt. Wenn dabei eine benachbarte Kategorie ernstlich in Betracht kommt, wird dies angegeben. Zum Beispiel steht 2/1 dafür, dass eine Streuung vorliegt, die eher Kategorie 2, gegebenenfalls aber auch Kategorie 1 zuzurechnen ist. Man erhält eine Skala von 0/- bis 3/+. Dabei bedeutet 0/-, dass offenkundig keine Schatten vorhanden sind, während 3/+ bedeutet, dass eine



wesentlich höhere Streuung vorliegt als 3/3.

Verbreitung: Die Lungen sind in 6 Felder: Ober-, Mittel- und Unterfeld, jeweils rechts und links, eingeteilt. Die Felder, in denen kleine Schatten zu sehen sind, werden markiert.

Die Streuungskategorie wird durch den Gesamteindruck aller betroffenen Felder und Vergleich mit den Standardfilmen bestimmt. Wo eine starke Streuungsdifferenz (mind. drei kleine Kategorien) in verschiedenen Feldern vorliegt, werden die Felder mit der wesentlich kleineren Streuung nicht beachtet. Die Streuung wird nicht durch Bildung eines arithmetischen Mittels errechnet, sondern die ILO fordert „einen Integrationsprozess im Kopf des Beurteilers“.

Form und Größe der kleinen Schatten müssen angegeben werden. Es werden zwei Formen unterschieden: rundlich und unregelmäßig. Für die rundlichen Schatten kodieren die Kürzel p, q, r die Größe (p = Durchmesser < 1,5 mm, q = 1,5–3 mm, r = 3–10 mm). Für die unregelmäßigen Schatten übernehmen diese Aufgabe die Kürzel s, t, u (s = Kaliber < 1,5 mm, t = 1,5–3 mm, u = 3–10 mm).

- große Schatten

Große Schatten sind größer als 10 mm im Durchmesser (steigende Staffelung von A nach C).

#### **Pleurabefunde:**

- Pleuraverdickung (Lokalisation, Dicke und Ausdehnung)

##### Brustwand:

Typ: umschrieben (Plaque) und/oder diffus

Lokalisation: linke und rechte Seite werden getrennt angegeben

Dicke: a = größte Breite < 5 mm, b = 5–10 mm, c  $\geq$  10 mm

Ausdehnung:

1 = Gesamtlängen-Äquivalent <  $\frac{1}{4}$  der seitlichen Brustwand,

2 = <  $\frac{1}{2}$  der seitlichen Brustwand,

3 = >  $\frac{1}{2}$  der seitlichen Brustwand

Zwerchfell: Plaques der Pleura diaphragmatica werden seitengetrennt angegeben.

Obliteration kostophrenischer Winkel: seitengetrennte Angaben

- Pleuraverkalkung

Lokalisation und Ausdehnung für beide Thoraxhälften getrennt

Lokalisation: Brustwand, Zwerchfell und andere Lokalisationen.

Ausdehnung:

- 1 = eine Zone mit einem  $\varnothing$  von 20 mm oder Zonen-Summen- $\varnothing$  von 20 mm,
- 2 = eine Zone  $> 20$  mm, aber  $< 100$  mm oder Zonen-Summen- $\varnothing$  mit den selben Werten,
- 3 = eine Zone von  $> 100$  mm oder Zonen-Summen- $\varnothing$  von  $> 100$  mm

**Symbole für Begleitbefunde:**

Begleitbefunde können vermerkt werden, sie sind in Tab. 2.5 wiedergegeben.

*Tabelle 2.5.: Erklärung der Symbole auf dem ILO-Bogen*

keine	keine Begleitbefunde
ax	Koaleszenz kleiner pneumokoniotischer Schatten
bu	bullöses Emphysem
ca	Karzinome der Lunge oder der Pleura
cn	Verkalkung innerhalb eines Pneumokonioseknötchens
co	abnorme Herzform und/oder -größe
cp	Cor pulmonale
cv	Kaverne, Einschmelzungen
di	Distorsion intrathorakaler Organe
ef	Pleuraerguss
em	Emphysem
es	Eierschalenverkalkung hilärer oder mediastinaler Lymphknoten
fr	Rippenfrakturen
hi	Vergrößerung hilärer oder mediastinaler Lymphknoten
ho	Honigwabellunge
idd	Zwerchfellunschärfe
idh	unscharfe Herzkontur
kl	Kerley-Linien
od	andere Befunde von Bedeutung („other disease“), Erläuterung unter „ergänzende Befunde“
pi	Pleuraverdickung in den Interlobärspalten oder Mediastinum
px	Pneumothorax
rp	rheumatoide Pneumokoniose
tb	Tuberkulose

**Bemerkungen**

Desweiteren besteht Raum für Bemerkungen. Diese sollten besonders dann festgehalten werden, wenn Fremddursachen für einen Befund angenommen werden, der von anderer Seite auch als Staublung aufgefasset werden könnte (Internationales Arbeitsamt Genf 1981). Die Tabellen 2.7 und 2.8 stellen die Bedeutungen der Abkürzungen des ILO-Bogens kurzgefasst dar.

Der Bogen wurde für epidemiologische Zwecke entwickelt. Er ist allerdings zu unpräzise, um ihn als einzige Determinante zur Krankheitserkennung und -bestimmung zu verwenden (Welch et al. 1998).

Im folgenden werden die Fehlerquellen diskutiert. Die ILO fordert die Dokumentation aller Erscheinungen, solange sie nicht definitiv keine Pneumokoniose sind. Es wird ver-

Tabelle 2.7.: ILO-Klassifikation des Parenchyms

<b>Streuung</b>			
0/- 0/0 0/1	Kategorie 0: kleine Schatten fehlen oder sind weniger gestreut, als der unteren Grenze der Kategorie 1 entspricht.		
1/0 1/1 1/2	Kategorien 1, 2 und 3 repräsentieren eine zunehmende Streuungsdichte, wie durch die entsprechenden Standardfilme belegt.		
2/1 2/2 2/3			
3/2 3/3 3/+			
<b>Verbreitung (Lokalisation)</b>			
RO	rechtes Oberfeld	LO	linkes Oberfeld
RM	rechtes Mittelfeld	LM	linkes Mittelfeld
RÜ	rechtes Unterfeld	LU	linkes Unterfeld
<b>Form</b>			
kleine runde Fleckschatten		kleine unregelmäßige Fleckschatten	
Bezeichnung	Durchmesser ( $\varnothing$ )	Bezeichnung	$\varnothing$
p	<1,5 mm	s	<1,5 mm
q	1,5-3 mm	t	1,5-3 mm
r	3-10 mm	u	3-10 mm
<b>große Schatten</b>			
A	$\varnothing$ einzeln oder in Summation 1-5 cm		
B	ein oder mehr Schatten, größer oder zahlreicher als A, Schattensumme nicht größer als das Äquivalent von RO		
C	ein oder mehr Schatten, Schattensumme größer als das Äquivalent von RO		

mutet, dass dadurch falsch-positive Ergebnisse entstehen, die sonst vielleicht nicht als Fibrose eingestuft würden. Dieses fehlerbelastete Vorgehen scheint die derzeitige unpräzise Diagnostik-Praxis zu reflektieren (Frumkin et al. 1990).

Für die Einordnung in die passenden Kategorien können zwar bei der ILO in Genf Standardfilme angefordert werden, dennoch unterliegt die Einordnung einer großen Subjektivität. Grenzwertige Bilder (0/1 = gesund und 1/0 = krankhaft) werden einmal als pneumokoniosefrei, ein anderes Mal als krankhaft bezeichnet. Dies geschieht sowohl bei unterschiedlichen Radiologen, die das gleiche Bild auswerten, als auch bei einem einzigen Untersucher, der das Bild zu unterschiedlichen Zeiten betrachtet. Darum empfiehlt die ILO, dass mindestens zwei, besser drei, unabhängige Klassifikationen eines Röntgenbildes durchgeführt werden sollten (Internationales Arbeitsamt Genf 1981).

In den USA können Ärzte an ILO-Staublungenklassifikations-Fortbildungen teilnehmen. Als A-Reader wird bezeichnet, wer an einem Kurs teilgenommen hat, als B-Reader, wer zudem eine Prüfung abgelegt hat (Welch et al. 1998). In Deutschland gibt es keine derartige Prüfung. Es gibt derzeit 16 Beurteiler, die sich durch einen großen Erfahrungsreichtum in der Beurteilung von Pneumokoniose-Röntgenbildern auszeichnen.

## 2.4.2. CT-Untersuchung

Die Computertomographie hat sich aus der Röntgendiagnostik entwickelt und ist ihr in vielen Bereichen überlegen. Eine Röntgenstrahlenquelle rotiert senkrecht zur Körperachse um die bewegungslos in einer Röhre liegende Person. Auf der der Strahlenquelle gegenüberliegenden Seite empfangen Detektoren die durch den Körperdurchtritt abgeschwächte Strahlung. Nach einer kompletten Runde um die Körperachse auf einer fixen Höhe können die Werte im Computer verrechnet werden und es entsteht ein Körperquerschnittsbild. Der Körper wird auf diese Weise in Scheiben zerlegt, so dass eine überlagerungsfreie Darstellung der einzelnen Organe möglich ist. Die Aufnahme erscheint in

Tabelle 2.8.: ILO-Klassifikation der Pleura

Pleuraverdickung	
<b>Typ</b>	umschrieben diffus
<b>Lokalisation</b>	linke Seite rechte Seite
<b>Breite(Dicke)</b>	a größte Breite <5 mm b größte Breite 5–10 mm c größte Breite >10 mm
<b>Ausdehnung</b>	1 Gesamtlängen-Äquivalent < $\frac{1}{4}$ der seidl. Brustwand, 2 < $\frac{1}{2}$ der seitlichen Brustwand, 3 > $\frac{1}{2}$ der seitlichen Brustwand.
Obliteration des kostophrenischen Winkels	
seitengetrennte Angabe	
Pleuraverkalkung	
seitengetrennte Angabe	
<b>Lokalisation</b>	Brustwand Zwerchfell andere Lokalisationen
<b>Ausdehnung</b>	1 einzelner Zonen- oder Summen- $\varnothing$ 20 mm 2 einzelner Zonen- oder Summen- $\varnothing$ 21–99 mm 3 einzelner Zonen- oder Summen- $\varnothing$ > 100 mm

verschiedenen Grautonschattierungen, wobei jede Schattierung für eine spezielle Dichte steht (Hounsfield-Einteilung). So wird eine sichere Abgrenzung verschiedener nebeneinander liegender Gewebe ermöglicht. Im Vergleich zu den konventionellen Röntgenbildern weisen CT-Bilder eine höhere Kontrast- und eine geringere Struktur- und Formaflösung auf.

Die Computertomographie ist erheblich teurer als eine konventionelle Röntgenaufnahme und führt zu einer größeren Strahlenbelastung des Körpers. Dennoch lassen sich manche asbestinduzierten Veränderungen nur hier erkennen. Deshalb wird gefordert, in speziell abgewägten Fällen auch die Computertomographie in der Vorsorgeuntersuchung von asbestexponierten Personen einzusetzen (Kraus et al. 1998). Am Rande sei erwähnt, dass kernspintomographische (MRT-)Aufnahmen in diesem Gebiet nutzlos sind. Um aussagekräftige MRT-Bilder zu erhalten, muss ein relativ dichtes Gewebe vorliegen, was bei der luftgefüllten Lunge nicht der Fall ist.

In Tab. 2.9 sind die Lungenveränderungen auf CT-Aufnahmen dargestellt, die typisch für asbestverursachte Pneumokoniosen sind. Diese Veränderungen sind jedoch nicht absolut spezifisch für eine beginnende Asbestose, deshalb muss man bei der Diagnosestellung vorsichtig sein und andere Indikatoren mit einbeziehen, besonders wenn die oben genannten Veränderungen einzeln auftreten (Bergin et al. 1994).

Mittels CT werden Lungenabschnitte erreicht, die man auf konventionellen Röntgenaufnahmen aufgrund von Überlagerungen kaum sehen kann (z.B. hinter dem Herzen, neben der Wirbelsäule und einige Pleura-Abschnitte) (Kraus et al. 1998; Levin et al. 2000). Im Gegensatz zu konventionellen Röntgenaufnahmen kann die gesamte Pleurazirkumferenz dargestellt werden. Durch Hounsfield-Einheiten-Dichtemessungen werden Abgrenzungen von epipleuralen Fettauflagerungen möglich. Auf konventionellen Röntgenaufnahmen lässt sich subpleurales Fett nicht als solches identifizieren und führt oft zu Fehleinschätzungen und Verdachtsanzeigen, da lediglich erkannt wird, dass sich et-

**Tabelle 2.9.:** Veränderungen bei asbestverursachter Pneumokoniose auf CT-Aufnahmen (Kraus et al. 1998)

Parenchym: <ul style="list-style-type: none"> <li>• interlobuläre, septal verdickte Linien</li> <li>• intralobuläre, non-septale, verdickte Linien und zentrilobuläre isolierte, mikronoduläre Strukturen</li> <li>• subpleurale, kurvilineare Linien</li> <li>• noduläre, irregulär begrenzte Schatten</li> <li>• milchglasähnliche Strukturen</li> <li>• kleine, subpleural gelegene zystische Veränderungen</li> <li>• parenchymale Bänder</li> <li>• Honigwabengunge</li> </ul>
Pleura: <ul style="list-style-type: none"> <li>• Plaques, mit oder ohne Verkalkungen</li> <li>• generalisierte (gleichförmige) Pleuraverdickung mit oder ohne subpleurale Lungenfibrose</li> <li>• Verschwartung (Hyalinosis complicata) bei rezidivierenden asbest-assoziierten Ergüssen nach Ausschluss anderer Ursachen.</li> </ul>

was Raumforderndes unterhalb der Pleura befindet (Kraus et al. 1998). Bezüglich nicht kalzifizierter Knoten zeigt sich bei CT-Aufnahmen eine dreimal, bezüglich Bronchial-Ca eine sechsmal höhere Sensitivität als bei konventionellen Röntgenaufnahmen (Levin et al. 2000).

Neben Untersuchungen mit normalen Computertomographen wurden seit den 80er Jahren auch solche mit hochauflösenden Computertomographen (HRCT) durchgeführt. Hierbei wurde festgestellt, dass HRCT-Aufnahmen wiederum sensitiver und verlässlicher sind als einfache CT-Aufnahmen (Aberle et al. 1988). Aberle und ihre Mitarbeiter untersuchten die radiologischen Veränderungen eines Kollektivs mit klinisch manifester Asbestose. Im HRCT wurden 96 %, im CT 83 % der Parenchymveränderungen dargestellt. Bei den Pleuraverdickungen war das Verhältnis HRCT 100 % zu CT 93 % (Aberle et al. 1988)). Friedman et al. kamen mit ihrer Studie zu ähnlichen Ergebnissen (Friedman et al. 1988, s. Tab. 2.10). Auch die Korrelation zwischen pathologisch-histologischen Asbestosebefunden und HRCT-Befunden war in vielen Studien sehr hoch (Aberle et al. 1988; Akira et al. 1990; Akira et al. 1991; Bessis et al. 1992).

**Tabelle 2.10.:** Positiver prädiktiver Wert von Rö, CT und HRCT in einer Studie von Friedman et al. 1988

	Röntgen[ %]	CT [ %]	HRCT [ %]
Parenchymveränderungen	51	83	100
Pleuraveränderungen	56	79	100

Zusammenfassend kann festgehalten werden, dass mittels HRCT eine frühe und exakte Erfassung reaktiver asbeststaubinduzierter Befunde sowohl des Lungengewebes als auch der Pleura möglich ist. Hier können Befunde erhoben werden, die zu diesem Zeitpunkt noch auf keinem anderen nicht-invasiven Weg erhoben werden können. Friedman et al. heben die Vorzüge des HRCT noch in einem weiteren Bereich hervor: mittels HRCT werden falsch-positive Resultate, wie sie bei der Fehlinterpretation von Röntgenbildern aufgrund von subpleuralem Fett, Plaques oder Emphysemen, die das Bild verschleiern, nicht selten entstehen, umgangen. Die Interpretation von Thoraxaufnahmen ehemals asbestexponierter Personen ist oft sehr schwierig und subjektiv, so dass gefordert wird, dass positive Ergebnisse mit HRCT-Aufnahmen bestätigt werden sollten (Friedman et al. 1988).

Mit einer Spiralaufnahmetechnik mittels Low-Dose Spiral-CT konnte die Strahlendosis unter 1 mSv gesenkt werden. Die jährliche natürliche Strahlenexposition liegt durchschnittlich bei 3 bis 3,5 mSv. Demzufolge sind heutzutage keine Einschränkungen der Anwendung der Computertomographie aufgrund zu hoher Strahlenbelastungen mehr nötig (Kraus et al. 1998).

Die Computertomographie gewinnt in der Arbeitsmedizin zunehmende Bedeutung. Schon Ende der 70er Jahre wurde sie zur Begutachtung fraglich berufsbedingter Lungen- und Pleuraveränderungen angewandt. In der gutachterlichen Befundung, besonders bei Brückenbefunden gemäß BK-Nr. 4104, hat die HRCT inzwischen oft einen entscheidungsrelevanten Stellenwert erlangt. Dies ist insbesondere deshalb so, weil die, zur Anerkennung einer Berufskrankheit häufig nötigen, Faserjahre (s. Abs 2.3) oft nicht präzise berechenbar sind, da gerade für den Zeitraum vor 1976 oft keine objektiven Messdaten zur Verfügung stehen (Kraus et al. 1998).

Wahrscheinlich wird die HRCT in Zukunft auch in der arbeitsmedizinischen Vorsorge eine zunehmend größere Rolle spielen. Zwischenzeitlich wird zum Beispiel akzeptiert, dass eine richtungweisende Einschätzung pleuraler Befunde und die Anerkennung/Ablehnung einer Berufskrankheit ohne CT/HRCT oft nicht möglich ist (Hering et al. 2003).

Während für die Röntgenaufnahmen die international anerkannte ILO-Klassifikation angewandt wird, bestand zu Beginn der vorliegenden Studie noch keine einheitliche, entsprechende Klassifikationsmethode für CT-Aufnahmen. Zwar waren schon verschiedentlich Vorschläge dazu gemacht worden, doch wurden sie nicht anerkannt und angenommen. Sowohl um Fehlbeurteilungen aufgrund ungeeigneter Untersuchungsparameter und unzureichender Befunderfassung zu umgehen, als auch um Verlaufsbeurteilungen unter vergleichbaren Bedingungen durchführen zu können, ist ein standardisiertes Verfahren auch für die Computertomographie von großer Bedeutung. Der bei dieser Studie verwendete Fragebogen wurde inzwischen von einer internationalen Arbeitsgruppe weiter entwickelt, um die Befundung zu standardisieren und international zu harmonisieren (Hering et al. 2004).

### 2.4.3. CT-Klassifikation

Vor Durchführung der vorliegenden Studie wurde von einem interdisziplinären Team ein standardisierter Fragebogen entworfen, der nach den ersten praktischen Erfahrungen weiterentwickelt wurde. Er ist in Anlehnung an den ILO-Bogen entwickelt worden und im Anhang B abgebildet. Die eingeführten Ziffern des Bogens (601-1068) sind fortlaufend zur ILO-Kodierung gewählt worden, so dass eine integrative Datenerfassung ermöglicht wurde. Er sollte als Diskussionsgrundlage für eine (möglichst internationale) Vereinheitlichung von Untersuchungsstrategie und Befunderfassung dienen. Diese Diskussion fand in den letzten Jahren rege statt und im Jahr 2000 wurde ein standardisierter Klassifikationsbogen entwickelt, der nun international Anwendung finden soll.

Eine analoge Klassifikation der CT-Befunde zur ILO-Klassifikation ist nicht möglich, da es sich um zwei sehr unterschiedliche Aufnahmeverfahren mit daraus resultierender unterschiedlicher Informationsvermittlung handelt. Die Informationen auf Röntgenaufnahmen entstehen durch Summationseffekte der Schatten, während die Computertomographie überlagerungsfreie Aufnahmen ermöglicht. Dennoch erschien eine Beibehaltung der Symbole für die Formen aus didaktischen Gründen sinnvoll. Daher wurden für den CT-Bogen die Bezeichnungen  $p'$ ,  $q'$ ,  $r'$  für kleine rundliche Fleckschatten und  $s'$ ,  $t'$ ,  $u'$  für kleine irreguläre Fleckschatten gewählt. Zudem wurden noch die Bezeichnungen  $v'$ ,  $w'$ ,  $x'$  für lineare Schatten eingeführt. Zu genaueren Definitionen siehe Tabelle 2.11. Auch bezüglich der Befunderfassung der Pleuraveränderungen richtet sich die CT-Dokumentation nach der ILO-Klassifikation. Es werden Lokalisation, Ausdehnung und Dicke der umschriebenen und gleichförmigen (Rö: diffusen) Pleuraverdickungen angegeben. Aufgrund der Schnittbildtechnik ist jedoch eine viel präzisere Aussage möglich. Zusätzlich zu den Symbolklassifikationen des ILO-Bogens wurden noch die Symbole „be“ (Bronchiektasen), „br“ (Bronchitis, Nachbarschaftszeichen), „cl“ (verkalkte Lymphknoten), „gg“ (ground glass pattern), „ra“ (Rundatelektasen) und „fp“ (extrapleurale Fettauflagerungen) eingeführt. Nach kurzer Einarbeitungszeit erwies sich der Bogen als gut praktikabel. Die Bearbeitungszeit lag zwischen 8 und 20 Minuten. In Tab. 2.11 ist die CT-Klassifikation des Parenchyms, in Tab. 2.13 die der Pleura dargestellt.

## 2.5. Interreadervarianz

Die ILO empfiehlt aufgrund großer Differenzen in der Beurteilung von Pneumokoniose-Röntgenbildern, dass mindestens zwei Fachleute ein Bild mit Pneumokoniose-Verdacht beurteilen (s. Abs. 2.4.1). In der Literatur haben sich verschiedene Arbeitsgruppen mit der Fragestellung auseinandergesetzt, wie hoch die Interreadervarianz ist. Die Interreadervarianz ist ein Maß für die Übereinstimmung zweier Personen bei der Beurteilung eines Befundes. Das zur Zeit anerkannteste Mittel zu ihrer Berechnung ist der sogenannte  $\kappa$ -Koeffizient.

Tabelle 2.11.: CT-Klassifikation des Parenchyms

Verbreitung (Lokalisation)			
RO	rechtes Oberfeld	LO	linkes Oberfeld
RM	rechtes Mittelfeld	LM	linkes Mittelfeld
RU	rechtes Unterfeld	LU	linkes Unterfeld
Streuung			
0	keine		
1	lockere Streuung, Lungenstrukturen eindeutig identifizierbar		
2	mitteldichte Streuung, normale Lungenstrukturen teilweise überlagert		
3	dichte Streuung, normale Lungenstrukturen kaum zu identifizieren, weitgehend oder vollständig überlagert		
Form			
kleine rundliche Fleckschatten			
Bezeichnung	Durchmesser ( $\varnothing$ )		
p'	< 1,5 mm (auch mikronodulär)		
q'	1,5-3 mm		
r'	3-10 mm		
kleine unreguläre Fleckschatten			
Bezeichnung	$\varnothing$		
s	< 1,5 mm, fein, linear, Länge < 1 cm		
t	1,5-3 mm, mittelgrob, noch linear, Länge < 1 cm		
u	3-10 mm, grob, vorwiegend klecksig, Länge < 1 cm		
vorwiegend lineare Schatten			
v'	Dicke/Kaliber > 1 mm, Länge 1-2 cm		
w'	Dicke/Kaliber > 1 mm, Länge 2-4 cm		
x'	Dicke/Kaliber > 1 mm, Länge > 4 cm, Parenchymbänder (Craw's feet)		
große Schatten			
0	keine Schatten		
1	$\varnothing$ einzeln oder in der Summation 1-5 cm		
2	$\varnothing$ einzeln oder in der Summation > 5 cm - $\frac{1}{2}$ der Schicht		
3	$\varnothing$ einzeln oder in der Summation > $\frac{1}{2}$ der Schicht		
honeycombing			
0	keine		
1	einzelne honigwabenartige Strukturen im Feld in < als ein $\frac{1}{3}$ der Schicht		
2	honigwabenartige Strukturen in $\frac{1}{3}$ - $\frac{1}{2}$ der Schicht. Lungenstrukturen nur noch andeutungsweise zu identifizieren		
3	honigwabenartige Strukturen in > $\frac{1}{2}$ der Schicht. Lungenstrukturen nicht mehr zu identifizieren		
Emphysem			
0	kein Emphysem		
1	Emphysem und/oder Bullae mit $\varnothing > 1$ cm, Überblähung bis zu $\frac{1}{3}$ der Schicht		
2	Emphysem oder mehrere Bullae oder Überblähung in $\frac{1}{3}$ - $\frac{1}{2}$ der Schicht		
3	Emphysem oder mehrere Bullae oder Überblähung in > $\frac{1}{2}$ der Schicht		



**Tabelle 2.13.: CT-Klassifikation der Pleura**

Pleuraverdickung	
<b>Typ</b>	umschrieben: tafelbergartig (erhaben) oder flach bis hügelartig (nicht-erhaben) aus dem Pleuraniveau diffus: konfluierende Pleuraverdickungen im Niveau der Pleura
<b>Lokalisation</b>	s. Parenchym +RD/LD: rechtes/linkes Diaphragma
<b>Breite(Dicke)</b>	a < 5 mm b 5–10 mm c > 10 mm
<b>Ausdehnung</b>	1 < $\frac{1}{4}$ der Zirkumferenz einer Seite 2 $\frac{1}{4}$ – $\frac{1}{2}$ der Zirkumferenz einer Seite 3 > $\frac{1}{2}$ der Zirkumferenz einer Seite
Pleuraverkalkung	
<b>Lokalisation</b>	RO, RM, RU, LO, LM, LU, RD, LD

### 2.5.1. Kurze Einführung in die $\kappa$ -Statistik

Möchte man die Übereinstimmung der Befunde zweier Ärzte bezüglich eines Röntgenbildes erfahren, kann man einen sogenannten  $\kappa$ -Koeffizienten errechnen. Bei der Berechnung von  $\kappa$  fließt im Gegensatz zu anderen Methoden der Übereinstimmungsberechnung (z.B. Prozent-Übereinstimmung) auch die Übereinstimmung ein, die durch Zufall erzielt würde. In einer Population mit einer hohen Erkrankungshäufigkeit wird die Übereinstimmung positiver Diagnosen schon allein durch Zufall größer sein als bei einer Population mit einer mittlerer Erkrankungshäufigkeit.  $\kappa$  bleibt von der Prävalenz einer Krankheit im Gegensatz zu Prozent-Übereinstimmungen unbeeinflusst.

Dies machen folgende Beispiele deutlich: Bei zwei verschiedenen Populationen beurteilen zwei Ärzte bei jeweils 100 Patienten, ob sie eine Lungenentzündung haben oder gesund sind. Im ersten Beispiel ist die Prävalenz sehr niedrig (ca. 1-5 %), während sie im zweiten Beispiel recht hoch ist mit ca. 40-45 % (s. Tab 2.15 und 2.16). Die Beispiele sind in Vierfeldertafeln aufgetragen. Als Prototyp ist Tab. 2.14 dargestellt. Zu genaueren Erklärungen muss auf Lehrbücher der Statistik verwiesen werden.

**Tabelle 2.14.: Prototyp Vierfeldertafel**

		Arzt Nr.1	
		Ja	Nein
Arzt Nr.2	Ja	Ja/Ja	Nein/Ja
	Nein	Ja/Nein	Nein/Nein

**Tabelle 2.15.:  $\kappa$ -Beispiel 1**

		Arzt Nr.1	
		Ja	Nein
Arzt Nr.2	Ja	1	3
	Nein	2	94

In ihren Befunden stimmen die Ärzte beide Male zu 95 % überein. Dennoch ist die Übereinstimmung unterschiedlich zu werten, wenn man beachtet, dass eine zufällige

**Tabelle 2.16.:**  $\kappa$ -Beispiel 2

		Arzt Nr.1	
		Ja	Nein
Arzt Nr.2	Ja	43	3
	Nein	2	52

Übereinstimmung im ersten Beispiel sehr viel wahrscheinlicher wäre als im zweiten.  $\kappa$  macht diesen Unterschied deutlich. Zu seiner Berechnung benötigt man:

- die **beobachtete Übereinstimmung** (In den o.g. Beispielen die Zellen, in denen die Beobachter übereinstimmen geteilt durch die Gesamtsumme der Zellen. Sie ist in beiden Tabellen 95 % -Bsp. 1:  $\frac{1+94}{100} = 0,95$  und Bsp. 2:  $\frac{43+52}{100} = 0,95$ ) und
- die **erwartete Übereinstimmung N** (zufällige Übereinstimmung -auf den Randsummen (s. Tab. 2.17) basierend).

**Tabelle 2.17.:** Randsummen  $\kappa$ -Beispiel 1

		Arzt Nr.1		Randsumme
		Ja	Nein	
Arzt Nr.2	Ja	1	3	4
	Nein	2	94	96
Randsumme		3	97	100

Um die durch Zufall zu erwartende Zahl N in jeder Zelle zu errechnen, wird davon ausgegangen, dass jede Zelle der Tabelle die Randsummen-Verteilung reflektieren sollte. Das Ja/Nein-Verhältnis sollte das Gleiche innerhalb der Vierkammer-Tabelle sein wie in der Randsumme. N gibt an, wie oft die untersuchte Konstellation (Ja/Ja, Ja/Nein, Nein/Ja, Nein/Nein) eintritt. Im Beispiel sind 96 % der Randsumme in der „Nein-Kategorie“. Dementsprechend müssten 96 % der Nein-Entscheidungen des Arzt Nr.1 im Nein/Nein-Feld sein. Das wäre hier  $0,96 \times 0,97 = 0,9312$  (s. Tab. 2.18). Durch Subtraktion lassen sich alle anderen Kästchen ausfüllen (s.Tab. 2.5.1).

**Tabelle 2.18.:** Berechnung der erwarteten Übereinstimmung für  $\kappa$ -Beispiel 1, Teil 1

		Arzt Nr.1		Randsumme
		Ja	Nein	
Arzt Nr.2	Ja			4
	Nein		93,12	96
Randsumme		3	97	100

Addiert man nun die Inhalte der Kästchen „erwartete Übereinstimmung Ja/Ja“ und „erwartete Übereinstimmung Nein/Nein“ ( $0,12 + 93,12$ ), so erhält man die erwartete Übereinstimmung. Nun kann man sehen, dass alleine durch Zufall 93,24 der 100 Diagnosen der zwei Beurteiler übereinstimmen.

Vergleich der beobachteten Übereinstimmung mit der erwarteten Übereinstimmung:

- die erwartete Übereinstimmung ist 6,76 % entfernt von der perfekten Übereinstimmung ( $100-93,24$ ).

**Tabelle 2.19.:** Berechnung der erwarteten Übereinstimmung für  $\kappa$ -Beispiel 1, Teil 2

		Arzt Nr.1		Rand- summe
		Ja	Nein	
Arzt Nr.2	Ja	0,12	3,88	4
	Nein	2,88	93,12	96
Randsumme		3	97	100

- die beobachtete Übereinstimmung ist 5 % entfernt von der perfekten Übereinstimmung (100-95).
- Dementsprechend waren die beiden Beobachter 1,76 % besser als der Zufall erwarten ließ. Sie hätten aber 6,76 % besser sein können.

Folgendes ist die Formel zur Berechnung von  $\kappa$  nach Cohen für zwei Beurteiler:

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \quad (2.1)$$

wobei  $p_0$  die **beobachtete Übereinstimmung**

$$p_0 = \frac{\sum_{i=1}^k f_i^i}{N} \quad (2.2)$$

und  $p_e$  die **erwartete Übereinstimmung** ist:

$$p_e = \frac{\sum_{i=1}^k e_i^i}{N} = \frac{\sum_{i=1}^k f_i \cdot f_i}{N^2} \quad (2.3)$$

$f_i$  = Randsumme für Objekt  $i$

$i = 1, \dots, k$

$k$  = Anzahl der Merkmalskategorien

$N$  = Anzahl der beurteilten Objekte

$e_i^i$  = Zeilensumme (i) x Spaltensumme (i)/N; gibt an, wie viele Urteilsübereinstimmungen zu erwarten wären, wenn die beiden Beurteiler rein zufällig urteilen würden, so dass die Urteile stochastisch unabhängig wären. (Bortz et al. 1998)

Einfacher formuliert heißt das:

$$\kappa = \frac{\text{beobachtete Übereinstimmung} - \text{erwartete Übereinstimmung}}{1 - \text{erwartete Übereinstimmung}}$$

in diesem Fall also:

$$\kappa = \frac{0,9500 - 0,9324}{1,0000 - 0,9324} = \frac{0,0176}{0,0676} = 0,26$$

Im zweiten Beispiel ist die beobachtete Übereinstimmung ebenfalls 95 %, doch die Randsummenverteilung ist anders (s. Tab. 2.16). Bei gleichem Vorgehen wie im ersten

**Tabelle 2.20.:** Erwartete Übereinstimmung Beispiel 2

		Arzt Nr.1		Rand- summe
		Ja	Nein	
Arzt Nr.2	Ja	20,7	25,3	46
	Nein	24,3	29,7	54
Randsumme		45	55	100

Beispiel wird Tabelle 2.20 zur Berechnung der erwarteten Übereinstimmung erhalten. Es lässt sich eine erwartete Übereinstimmung von 50,4 % berechnen (20,7 % + 29,7 %).

In die  $\kappa$ -Formel eingesetzt ergibt sich folgende Rechnung:

$$\kappa = \frac{0,950 - 0,504}{1,000 - 0,504} = \frac{0,446}{0,496} \approx 0,9$$

Im zweiten Beispiel ist die beobachtete Übereinstimmung von 95 % viel weiter von der durch Zufall zu erwartenden Übereinstimmung von 50,4 % entfernt. Dies macht der  $\kappa$ -Wert (0,9), der deutlich höher ist als der  $\kappa$ -Wert des ersten Beispiels (0,26), deutlich.

In den oben genannten Beispielen haben zwei Ärzte eine Vielzahl von Patienten nach einem Ja/Nein-Schema beurteilt (krank/gesund). Gibt es allerdings eine größere Anzahl an Diagnosemöglichkeiten und möchte man nicht alle Unterschiede als gleichwertig bewerten, so muss man einen etwas weiterführenden Weg einschlagen. Zwischen verschiedenen Diagnosemöglichkeiten muss eine Wichtung vorgenommen werden. Es muss beachtet werden, dass eine falsche Diagnose mit ähnlichen Folgen für die Behandlung nicht so gravierend ist wie eine falsche Diagnose mit resultierender gänzlich anderer Behandlung. Im besten Fall kann mit einer Ordinalskala gearbeitet werden: je weiter weg vom richtigen Wert, desto schlechter.

Hierfür hat Cohen eine Lösung in Form des **gewichteten („weighted“) kappa** ( $\kappa_w$ ) gefunden:

$$\kappa_w = 1 - \frac{\sum_{i=1}^k \sum_{j=1}^k v_{ij} \cdot f_{ij}}{\sum_{i=1}^k \sum_{j=1}^k v_{ij} \cdot e_{ij}} \quad (2.4)$$

i = Objekt 1, ....., N

j = Kategorie 1, ....., k

k = Anzahl der Merkmalskategorien

N = Anzahl der beurteilten Objekte

$v_{ij}$  = Diskrepanzgewicht; übereinstimmende Urteile, also in der Diagonale der k x k Tafel eingetragenen Häufigkeiten, erhalten das Gewicht 0. Die diskrepanten Urteile (nichtdiagonale Häufigkeiten) werden um so höher gewichtet, je weiter die entsprechenden Kategorien voneinander entfernt sind. Die maximal mögliche Diskrepanz wird mit 1 gewichtet.

$f_{ij}$  = beobachtete Häufigkeiten

$e_{ij}$  = erwartete Häufigkeiten (Bortz et al. 1998)

Ein  $\kappa$  von +1 bedeutet, dass die Beurteiler in ihren Diagnosen genau übereinstimmen, 0 entspricht genau der Übereinstimmung, die der Zufall erwarten ließe, und -1 bedeutet, dass die Beurteiler genau gegenteilige Diagnosen stellen. In Tabelle 2.21 ist das Bewertungsschema für  $\kappa$ -Werte nach Landis und Koch dargestellt (Landis 1977).

**Tabelle 2.21.:**  $\kappa$ -Bewertung nach Landis und Koch

$\kappa$	> 0,8 gilt als beinahe perfekt
	0,61 – 0,8 gilt als beachtlich
	0,41 – 0,6 gilt als gut
	0,21 – 0,4 gilt als mittelmäßig
	0,0 – 0,2 gilt als schlecht

## 2.5.2. Bisherige Studien zur Interreadervarianz

Verschiedene Studien haben sich mit der Interreadervarianz beschäftigt. Ein Vergleich der Studien ist schwierig, da es keine standardisierten Bedingungen gab, unter denen die Studien durchgeführt wurden. Es wurden unterschiedliche Untersuchungsbögen verwendet und verschiedene Formen der Kategorisierung vorgenommen. Zudem ist die Beschreibung des genauen Kategorisierungsvorgangs in der Literatur häufig sehr unpräzise, so dass in manchen Fällen ein genaues Nachvollziehen der statistischen Methoden nicht möglich ist. In den Tabellen 2.22, 2.24, 2.25 und 2.26 sind die verschiedenen Studien zusammenfassend dargestellt. Nur einige Studien erwähnen explizit, dass bei ihnen der Mittelwert der  $\kappa$ -Werte angegeben wurde. Bei den restlichen Studien wird ein gleiches Vorgehen angenommen.

In den meisten Studien lagen den Beurteilern keine Informationen über die Patienten vor (Ausnahme: DeRaeve 2001: hier wussten die Beurteiler, dass alle Patienten asbestexponiert waren, Tiitola 2002: hier erfuhren die Beurteiler Namen und Alter). Die meisten Studien haben sich mit der Untersuchung der Interreadervarianz bei der Interpretation von Röntgenbildern befasst. Hier lag fast immer der ILO-1980-Bogen zugrunde, auch wenn er teilweise stark vereinfacht wurde. Die Studien, die sich mit der Interreadervarianz bei der Interpretation von CT-Aufnahmen befassten, standen vor dem Problem, dass sie eine Methode verwandten, die auf diesem Gebiet noch kein Standard war. Dementsprechend gab es keine Konsens-Klassifikationsbögen. Die Gruppen entwickelten ihre eigenen Klassifikationsbögen, weshalb die  $\kappa$ -Werte der verschiedenen Gruppen nicht vergleichbar sind.

Lediglich in zwei Studien wurden sowohl  $\kappa$ -Werte für die Interpretation von Röntgenbildern als auch von CT-Aufnahmen ermittelt. Bei Copley et al. ging es um die Beurteilung der Pleura (Copley et al. 2001), bei Begin et al. um die Beurteilung des Parenchyms (Begin et al. 1993)(s. unten). Copley et al. nahmen eine gänzlich andere Kategorisierung als die Autorin dieser Studie vor, während die Kategorisierung von Begin et al. relativ ähnlich war.

**Tabelle 2.22.:** Studien zum Thema Interreadervarianz bei der Beurteilung des Parenchyms auf Röntgenbildern (Streuung kleiner Fleckschatten) (KZ: Kategorienanzahl; 4HK: vier Hauptkategorien: 0 = 0/0, 0/1; 1 = 1/0, 1/1, 1/2; 2 = 2/1, 2/2, 2/3; 3 = 3/2, 3/3; d = dichotom ( $\geq 1/0$  vs  $\leq 0/1$ ))

Literatur	Kollektiv	Charakteristika	Beurteiler	Bogen	Vergleichsmodus	Kategorisierung	KZ	$\kappa$ -Werte
Musch et al. 1985	743	Männer aus Kohlebergwerksdorf (davon 392 Bergmänner)	3 B-Reader	ILO- Ü/C 1972	3 Paare	4HK	4	0,61
Muir et al. 1992	205	Aufnahmen von 48 Bergleuten	5 Beurteiler	ILO-1980	generalisiertes $\kappa$	12-Punkte-Skala -für 6 Zonen einzeln	12	0,15
						-für ges. Lunge =Integrationsprozess	12	0,08
						4HK	4	0,16
Begin et al. 1993	61 (5)	Asbestexponierte mit ILO- Kategorie 0 oder 1 (Kontr. gesund)	5 Beurteiler	ILO-1980	unbekannt	4HK	4	0,20
						d	2	0,44
Welch et al. 1998	119	Metallarbeiter	6 Beurteiler	ILO-1980	Median der fünf anderen Beurteiler	d	2	0,51
						4HK	4	0,39
Impivaara et al. 1998	7.095	Repräsentativer Querschnitt der finnischen Bevölkerung	2 Radiologen	ILO-1980 in leichter Abwandlung	1 Paar	4 Kategorien: 0/0; 0/1; 1/0; 1/1-2/2	4	0,48 $\kappa_w = 0,72$
	597	angereicherte Befunde: $\geq 1/0$	1.Durchgang					0,20 $\kappa_w = 0,55$
			2.Durchgang					0,35 $\kappa_w = 0,54$

**Tabelle 2.24.:** Studien zum Thema Interreadervarianz bei der Beurteilung der Pleura auf Röntgenbildern (KZ: Kategorienanzahl; \*: Verweis auf Zitting 1995 bei Impivaara 1998; °: Obliteration des kostophrenischen Winkels oder Verdickung des horizontalen Interlobularspalts)

Literatur	Kollektiv	Charakteristika	Beurteiler	Bogen	Vergleichsmodus	Kategorisierung	KZ	$\kappa$ -Werte
Frumkin et al. 1990	434	geringe Asbestexposition	6 B-Reader	ILO-1980 (stark vereinfacht)	15 Paare	dichotom (pos: bds. gr. Läsionen (mind. B oder 2) (Dicke(0-C), Ausdehnung(0-3)) )	2	0,24
Impivaara et al. 1998	7.095	Repräsentativer Querschnitt der finnischen Bevölkerung	2 Radiologen	ILO-1980*	1 Paar	umschriebene Pleuraverdickung ja/nein	2	0,73
						Pleurakalzifizierung, ja/nein	2	0,73
						Viszeralpleuraveränderungen ja/nein <sup>°</sup>	2	0,74
						Viszeralpleuraveränderungen außer fissuralen Verdickungen ja/nein	2	0,83
Welch et al. 1998	119	Metallarbeiter	6 Beurteiler	ILO-1980	Median	dichotom ( pos: Pneumokoniose (3A) und spezifische Befunde in Sektion 3B-D)	2	0,50
Copley et al. 2001	50	benigne Pleuraläsionen	2 Beurteiler	ILO-1980	Paar	jede Lungenhälfte separat,		
						Obliteration kostophrenischer Winkel ja/nein	2	$k_w=0,78$
						Dicke -nichtkalzifizierter diffuser Pleuraverdickung	4	$k_w=0,84$
						-umschriebener Pleuraplaques	4	$k_w=0,44$
						Ausdehnung der diffusen Verdickung	4	$k_w=0,87$
						Ausdehnung der Pleurakalzifizierung	4	$k_w=0,45$
Pleurakalz.durchmesser	3	$k_w=0,53$						

**Tabelle 2.25.:** Studien zum Thema Interreadervarianz bei der Beurteilung des Parenchyms auf CT-Bildern (KZ: Kategorienanzahl; (x): Anzahl gesunder Kontrollpersonen; 4HK: vier Hauptkategorien: 0 = 0/0, 0/1; 1 = 1/0, 1/1, 1/2; 2 = 2/1, 2/2, 2/3; 3 = 3/2, 3/3)

Literatur	Kollektiv	Charakteristika	Beurteiler	Bogen	Vergleichsmodus	Kategorisierung	KZ	$\kappa$ -Werte
Akira et al. 1991	23 (17)	asbestexponierte Personen mit keinen/minimalen Läsionen auf dem Röntgenbild	2 Beurteiler	eigener Bogen	Paar	4 Kategorien (jew. für 2 Lungenzonen: innen, außen): 0 %, $\leq 25$ %, $\leq 50$ % und $\geq 50$ % des Parenchyms betroffen; für 6 Aspekte (s.Text)	4	$>0,90$ ; $0,85^a$ (HRCT)
Begin et al. 1993	61(5)	Asbestexponierte mit ILO- Kategorie 0 oder 1	5 Beurteiler	ILO-1980 orientiert	unbekannt	4HK	4	0,28 (CT+HRCT)
						dichotom (pos: $\geq$ Kategorie 1)	2	0,63 (CT+HRCT)
	14	als pos. beurteilte Bilder				dichotom (pos: $\geq$ Kategorie 1)	2	0,73 (HRCT) 0,84 (CT)

<sup>a</sup>intra- und interlobular-Linien



**Tabelle 2.26.:** Studien zum Thema Interreadervarianz bei der Beurteilung der Pleura auf CT-Bildern (KZ: Kategorienanzahl; ?: KZ mit Hilfe des Quellcodes nicht nachvollziehbar;  $\kappa_w$ : gewichtetes  $\kappa$ ; (x): Anzahl gesunder Kontrollpersonen; \*: zu genaueren Erläuterungen siehe Text; °: 4 willkürliche Kategorien für beide Hemithoraces: (0=no, 1=mild, 2=moderate, 3=severe))

Literatur	Kollektiv	Charakteristika	Beurteiler	Bogen	Vergleichsmodus	Kategorisierung	KZ	$\kappa$ -Werte
Copley et al. 2001	50	benigne Pleuraläsionen	2 Beurteiler	eigener Bogen	Paar	4 Quadranten auf 5 Höhen Pleuraplaque, Anzahl der Quadranten Atelektasen Dicke - Umfang -diffuser Pleuraverdickungen	4 ? ? 4	$\kappa_w=0,87$ $\kappa_w=0,74$ $\kappa_w=0,59$ $\kappa_w=0,82$
				nach Aberle und Jarad	Paar	re.+li. Hemithorax auf 3 Höhen: diffuse Pleuraverdickung*  Anzahl der Pleuraplaques* Atelektasen* Kalzifikation* fissurale Pleuraverdickung*	4 4 3 2 3	$\kappa_w=0,78$ $\kappa_w=0,74$ $\kappa_w=0,74$ $\kappa_w=0,63$ $\kappa_w=0,57$
De Raeve et al. 2001	100	≥ 10 J. Asbestexposition im Arbeitsstättengebäude	3 Radiologen	eigener Bogen	3 Paare	5 Kategorien (0=normal bis 4=anormal mit hoher Wahrscheinlichkeit für Plaques)	5	$\kappa_w=0,38$
Tiitola et al. 2002	602 (49)	Bauarbeiter mit asbestbedingten Veränderungen	3 Radiologen	eigener Bogen	unbekannt	durchschnittl. Dicke der Pleuraverdickungen,- größte Dicke- 3 ILO-Kategorien Kalzifikation Ausdehnung konvergierende Adhärenzen° diaphragmatische Adhärenzen° andere Adhärenzen° Interlobarpleuraplaques° Interlobarpleuradicke°	3 3 ? ? 4 4 4 4 4	$\kappa_w=0,60$ $\kappa_w=0,60$ $\kappa_w=0,81$ $\kappa_w=0,75$ $\kappa_w=0,38$ $\kappa_w=0,31$ $\kappa_w=0,23$ $\kappa_w=0,49$ $\kappa_w=0,24$

## Parenchym auf Röntgenbildern

Für die Beurteilung des Parenchyms (immer Streuung kleiner Fleckschatten) auf Röntgenbildern ist eine Einteilung in die vier Hauptkategorien der ILO-Klassifikation verbreitet:

**0** = 0/–, 0/0 und 0/1;

**1** = 1/0, 1/1 und 1/2;

**2** = 2/1, 2/2 und 2/3;

**3** = 3/2, 3/3 und 3/+).

Allen Studien, die sich mit der Beurteilung des Parenchyms auseinandergesetzt haben, ist gemein, dass alle Beurteiler alle Bilder beurteilt haben.

Musch et al. erreichten bei der Beurteilung der Bilder von 743 Männern aus einem Kohlebergwerksdorf ein durchschnittliches  $\kappa$  von 0,61 (drei Beurteilerpaare: 0,52, 0,57, 0,74) (Musch et al. 1985). Sie stellten fest, dass die meisten Unterschiede zwischen den Beurteilern bei der Einteilung in die Kategorien 0 und 1 erschienen. Hier wurde der Einfluss gesammelter Erfahrungen, Trainingseffekte und Haltung gegenüber minimaler Veränderungen und der Frage, was der Beurteiler als normal/anormal empfindet, diskutiert. Abgesehen von diesen subjektiven Faktoren gäbe es Faktoren wie Leseplatzbedingungen und Müdigkeit, die auf die Entscheidungsfindung Einfluss nähmen. Die Übereinstimmung bei Bildern schlechter Qualität war signifikant schlechter als bei Bildern guter Qualität.

Muir et al. erreichten bei der Beurteilung der Bilder von 48 Bergleuten ein durchschnittliches  $\kappa$  von 0,16 (Muir et al. 1992). Sie konnten zeigen, dass die Übereinstimmung besser war, wenn verschiedene Lungenabschnitte einzeln beurteilt wurden, ohne dass ein mentaler Integrationsprozess stattfand und wenn die Anzahl der Kategorien niedriger war.

Begin et al. erreichten bei der Beurteilung der Bilder 61 asbestexponierter Personen ein durchschnittliches  $\kappa$  von 0,20 (Begin et al. 1993). Auch hier konnte gezeigt werden, dass eine geringere Kategorienzahl zu einer besseren Übereinstimmung führt. Es konnte ein signifikant besseres Ergebnis bei der Beurteilung von CT-Bildern gegenüber Röntgenbildern und von CT-Bildern gegenüber HRCT-Bildern gezeigt werden.

Welch et al. erreichten bei der Beurteilung der Bilder von 119 Metallarbeitern ein durchschnittliches  $\kappa$  von 0,39 (Welch et al. 1998). Sie konnten keinen Effekt der Filmqualität auf die Interreadervarianz nachweisen. Sie betonten, dass die ILO-Klassifikation zu unpräzise zur alleinigen Beurteilung einer Krankheit sei. Die Studie zeige, dass gut qualifizierte Beurteiler sich in ihrer Beurteilung stark unterschieden und somit die Diagnose einer asbestverursachten Krankheit auf dem kompletten klinischen Erscheinungsbild fußen müsse.

Impivaara et al. bildeten auf eine andere Art vier Kategorien: 0/0; 0/1; 1/0 und 1/1 – 2/2. Sie begründeten die Zusammenfassung von 1/1 bis 2/2 dadurch, dass in ihrer

Population sehr wenige Personen eine Streuung  $\geq 1/1$  aufwiesen (Impivaara et al. 1998). Sie erreichten bei der Beurteilung der Bilder von 7095 Personen im repräsentativen Querschnitt der finnischen Bevölkerung ein durchschnittliches  $\kappa$  von 0,48 und durchschnittliches gewichtetes  $\kappa$  von 0,72. Sie weisen daraufhin, dass die beiden Beurteiler schon viele Jahre zusammen gearbeitet hätten. Dementsprechend seien ihre Ergebnisse nicht zu verallgemeinern, zeigten jedoch, dass unter günstigen Umständen eine relativ hohe Übereinstimmung erzielt werden könne. Dennoch seien Unterschiede festzustellen, weshalb die ILO-Empfehlung, dass mindestens zwei Personen ein Bild beurteilen sollten, sehr begründet sei. Sie konnten bestätigen, dass die Varianz bei der Beurteilung grenzwertiger Befunde (0/1 vs 1/0) besonders groß ist. Sie zeigten, dass meist nur eine Abweichung um eine Kategorie vorlag. Dafür wurde sowohl ungewichtetes als auch gewichtetes  $\kappa$  berechnet. Das gewichtete  $\kappa$  war um den Faktor 1,5 bis 2,6 mal höher. Die Kategorien waren sehr klein (0/0; 0/1; 1/0; 1/1-2/2), so dass die Erkenntnis für eine klare Ähnlichkeit der Interpretationen spricht. Nur bei 0,1 % der Befunde waren Abweichungen um mehr als eine Kategorie zu finden. Insgesamt war die Übereinstimmung zwischen den beiden Beurteilern etwa so gut wie zwischen zwei Beurteilungen derselben Person.

### **Pleura auf Röntgenbildern**

Bei der Beurteilung der Pleura auf Röntgenbildern gibt es keine Kategorisierungstendenz.

In drei Studien gab es eine dichotome Kategorisierung, wobei Positivität unterschiedlich definiert wurde. Bei Frumkin et al. galt ein Befund dann als positiv, wenn beidseits große Läsionen zu finden waren (Frumkin et al. 1990). Große Läsionen waren als mind. B oder 2 definiert, wobei die Dicke in 0-c und die Ausdehnung in 0-3 angegeben wurde. Bei der Interpretation der Bilder von 434 gering asbestexponierten Personen erreichten sie ein  $\kappa$  von 0,24. Hierbei wurden nicht alle Bilder von allen Beurteilern gelesen. Sie kamen zu dem Ergebnis, dass die Genauigkeit bei der Diagnose von Pleuraverdickungen auf Röntgenbildern limitiert sei, selbst wenn die Bilder von B-Readern beurteilt würden. Sie weisen darauf hin, dass laut ILO alle fraglichen Befunde als Pneumokoniose festgehalten werden müssen, solange sie nicht definitiv keine Pneumokoniose sind. Dies könne die Übereinstimmung verschlechtern, würde aber die derzeitige Praxis widerspiegeln. Sie zeigten, dass die Interpretation weder durch schlechte Filmqualität noch durch konkurrierende Krankheiten beeinflusst wurde.

Die Kategorisierung von Impivaara et al. war durchweg dichotom (Impivaara et al. 1998). Sie erreichten bei der Beurteilung der Bilder von 7095 Personen im repräsentativen Querschnitt der finnischen Bevölkerung ein durchschnittliches  $\kappa$  von 0,73-0,83 (ILO-umschriebene Pleuraverdickungen vorhanden/nicht vorhanden: 0,73, Pleurakalzifizierung vorhanden/nicht vorhanden: 0,73, Veränderungen der Viszeralpleura (= Obliteration kostophrenischer Winkel oder Verdickungen der Interlobularlinien) vorhanden/nicht vorhanden: 0,74, Veränderungen der Viszeralpleura ohne Verdickungen der Interlobularlinien vorhanden/nicht vorhanden: 0,83). Damit zeigten sie, dass die Übereinstimmungen bei der Beurteilung der Pleura besser war als bei der Beurteilung paren-

chymaler Veränderungen, wobei die Dichotomie der Pleurakategorisierung auch zu den besseren Werten beigetragen haben wird.

Bei Welch et al. 1998 galt ein Befund dann als „positiv, wenn eine Pneumokoniose (3A) und spezifische Befunde in den Sektionen 3B-D zu finden waren“ (Welch et al. 1998). Sie erreichten bei der Interpretation der Bilder von 119 Metallarbeitern ein durchschnittliches  $\kappa$  von 0,50, was in etwa dem  $\kappa$ -Wert der Parenchymfragestellung entspricht (0,51).

Copley et al. kategorisierten sechs Aspekte unterschiedlich: 1) Obliteration des kostophrenischen Winkels: vorhanden/nicht vorhanden ( $\kappa=0,78$ ), 2+3) Dicke nichtkalzifizierter diffuser Pleuraverdickung (0,84) sowie umschriebener Pleuraplaques (0,44) (4 Kategorien: 0, < 5 mm, 5–10 mm, >10 mm), 4+5) Ausdehnung der diffusen Verdickung (0,87) sowie der Pleurakalzifizierung (0,45) (4 Kategorien: 0 %, < 25 %, 25–50 %, > 50 %), 6) Pleurakalzifizierungs-Durchmesser (0,53) (3 Kategorien: größter Durchmesser < 20 mm oder mehrere Bereiche, deren Durchmesser Summe < 20 mm, 20–100 mm, > 100 mm) (Copley et al. 2001). Sie stellten fest, dass die Beurteilung diffuser Pleuraverdickungen bei beiden Verfahren (Röntgen und CT) gut war, dennoch war die Übereinstimmung beim Röntgen schlechter (besonders bezüglich der Ausdehnung von Pleuraplaques). Sie hielten fest, dass bei Patienten mit isolierter Pleuraerkrankung eine Röntgenaufnahme ausreichend sein könnte, wobei ein CT trotzdem nötig sein könnte, um andere Krankheitsgeschehen zu identifizieren, wo Details des Parenchyms wichtig seien (wie z.B. Asbestose).

### Parenchym auf CT-Bildern

Zwei Studien befassten sich mit der Interreadervarianz bei der Beurteilung des Parenchyms auf CT-Aufnahmen.

Akira et al. entwickelten einen ILO-unabhängigen Bogen (Akira et al., 1991). Bei 23 Personen mit keinen oder minimalen Läsionen auf dem Röntgenbild und 17 Kontrollen wurden sechs Aspekte untersucht: 1) verdickte Intra- und Interlobularlinien, 2) subpleurale kurvilineare Linien, 3) pleural basierte noduläre Unregelmäßigkeiten, 4) verschwommene Bereiche verstärkter Verschattung, 5) kleine zystische Bereiche und 6) kleine Bereiche geringer Verschattung. Für eine innere und eine äußere Zone musste beurteilt werden, wie viel Prozent des Parenchyms betroffen waren: 0 %,  $\leq 25$  %,  $\leq 50$  % und  $\geq 50$  %. Bei der Beurteilung der Intra- und Interlobularlinien wurde ein Wert von 0,85, bei den anderen Aspekten wurden Werte von  $> 0,90$  erlangt. Die Untersucher wiesen auf die kleine Probandengruppe hin und empfahlen eine Kontrolluntersuchung mit einem größeren Kollektiv.

Begin et al. entwickelten einen ILO-1980-orientierten Kategorisierungsbogen, und wählten auch hier die vier ILO-Hauptkategorien bzw. eine dichotome Auswertung, bei der als positiv galt, was  $\geq$  Kategorie 1 war (Begin et al. 1993). Mit den vier Hauptkategorien erlangten sie ein  $\kappa$  von 0,28, mit den dichotomen Kategorien ein  $\kappa$  von 0,63. Mit einer reduzierten Anzahl an Bildern, denen gemeinsam war, dass sie im ersten Durchgang als positiv beurteilt worden waren, wurde in einer zweiten dichotomen Beurteilung zur Unterscheidung zwischen HRCT und umfassendem CT mittels HRCT ein  $\kappa$  von 0,73

und mittels umfassendem CT eines von 0,84 erlangt. Es konnte gezeigt werden, dass bei der gleichen Kategorisierung gut aufeinander abgestimmter Röntgen- und CT-Bögen die Übereinstimmung bei der Interpretation von CT-Bildern deutlich besser war, obwohl es für diese keine Referenzfilme gab. Das HRCT habe in bis zu 30 % der Fälle zu mehr Präzision und Klarheit der Befunde beigetragen, so dass die Klassifikation weniger zögerlich und mit mehr Zutrauen habe durchgeführt werden können, obwohl die Übereinstimmung niedriger gewesen sei.

### Pleura auf CT-Bildern

Auch bei der Beurteilung der Pleura auf CT-Aufnahmen gab es keine einheitliche Kategorisierung.

Copley et al. untersuchten mittels einfachem CT und umfassendem CT (Copley et al. 2001). Beim einfachen CT wurden auf fünf Höhen 1) Pleuraplaques (Anzahl der betroffenen Quadranten) ( $\kappa=0,87$ ), 2) Rundatelektasen (Vorhandensein und Anzahl) (0,74), 3) diffuse Pleuraverdickung (durchschnittliche Dicke)(0,59) und 4) der Umfang der Pleuraverdickung (in % des Thorax angegeben) (0,82) beurteilt. Im umfassenden CT wurde auf die Kategorisierung von Aberle et al. und Jarad et al. zurückgegriffen. Hierbei wurde der Thorax in horizontale Drittel unterteilt und zudem jeder Hemithorax einzeln betrachtet. In jedem Drittel wurden 1) Max., Min. und durchschnittliche Dicke der diffusen Pleuraverdickung (0,78), 2) max. Umfang der Pleuraplaques (0,74) (größter Umfang: < 25 %, 25-50 %, 51-75 % und > 75 %), 3) Anwesenheit von Kalzifizierungen (wenn ja: lokalisiert gegen ausgedehnt) (0,63), 4) Rundatelektasen (0,74)(1 = Krähenfüße, 2 = < 4 cm im Durchmesser, 3 = > 4 cm im Durchmesser), und 5) das Maß der Spaltenverdickungen (0,57)(< 50 %, 50-99 % und 100 %). Sie konnten herausstellen, dass das einfache CT das Verfahren der Wahl war.

De Raeve et al. wählten fünf Kategorien (0 = normal 1 = anormal, aber kein Plaque, 2 = anormal mit geringer Wahrscheinlichkeit für Plaques, 3 = anormal mit mittelmäßiger Wahrscheinlichkeit für Plaques und 4 = anormal mit hoher Wahrscheinlichkeit für Plaques), um eine allgemeine Veränderung basierend auf Anwesenheit, Ausdehnung, Verteilung und Erscheinung der Pleuraveränderungen festzuhalten (De Raeve et al. 2001). Bei 100 Personen, die mind. zehn Jahre asbestexponiert waren, wurde ein durchschnittliches  $\kappa$  von 0,38 erreicht (3 Paare: 0,26, 0,43, 0,45). Sie konnten herausstellen, dass beim Konsensuslesen fünf zusätzliche Fälle minimaler Pleuraläsionen festgestellt wurden, was bedeute, dass eine Interpretation durch drei Radiologen die Entdeckungsrate kleiner pleuraler Läsionen verglichen mit der durch einen Radiologen anhebe.

Tiitola et al. beurteilten 602 Bauarbeiter mit asbestbedingten Veränderungen. Die durchschnittliche und größte Dicke der pleuralen Veränderungen wurden nach ILO jeweils mit 3 Kategorien beurteilt ( $\kappa$  jeweils 0,60)(Tiitola et al. 2002). Bei der Beurteilung der Kalzifikation (keine Angabe zu Anzahl der Kategorien, wahrscheinlich nach ILO zwei) wurde ein  $\kappa$  von 0,81 erreicht. Für die Beurteilung der Adhärenzen und Interlobarpleuraveränderungen wurden vier willkürliche Kategorien gewählt (0 = normal, 1 = schwache, 2 = mittelmäßige, 3 = starke Veränderungen). Die  $\kappa$ -Werte betragen für konvergierende Adhärenzen: 0,38, diaphragmatische Adhärenzen: 0,31, andere Adhären-

zen: 0,23, Interlobarpleuraplaques: 0,49 und Interlobarpleuradicke: 0,24. Es zeigte sich, dass bei der Beurteilung von Veränderungen der viszeralen Pleura (Adhärenzen und Interlobarverdickungen) eine schlechte Übereinstimmung erzielt werden konnte. Tiitola et al. erarbeiteten, dass Plaqueausdehnung, -kalkifikation und -dicke verlässliche Indikatoren für benigne Pleuraläsionen seien und dass sie in zukünftigen Klassifikationsystemen berücksichtigt werden sollten.

### 3. Hypothesen

In der vorliegenden Studie wurden Röntgen- und CT-Aufnahmen von 636 Patienten von jeweils zwei Beurteilern klassifiziert. Insgesamt gab es 7 Beurteiler, die zu 10 Paaren kombiniert wurden. Die Befunde der Röntgenaufnahmen wurden in ILO-1980-Bögen und die Befunde der CT-Aufnahmen in einen selbstentwickelten Bogen eingetragen.

Folgende Hypothesen sind zu prüfen:

- 1) Bei der Interpretation von Röntgenbildern wird die Interreadervarianz größer sein als bei der Interpretation von CT-Aufnahmen. CT-Aufnahmen sind leichter zu lesen und führen deshalb zu übereinstimmenderen Aussagen, obwohl bei der Interpretation von Röntgenbildern bezüglich Pneumokoniosen viel mehr Übung besteht.
- 2) Bei Röntgen-Aufnahmen wird die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen größer sein als bei pleuralen Veränderungen. Allgemein lassen sich pneumokoniotische Veränderungen besser im Pleurabereich erkennen als im Parenchym.
- 3) Bei CT-Aufnahmen wird die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen größer sein als bei pleuralen Veränderungen mit der gleichen Begründung wie 2).
- 4) Im zeitlichen Verlauf wird die Interreadervarianz bei der Beurteilung von Röntgenbildern sinken, da ein Lernprozess stattfinden wird.
- 5) Im zeitlichen Verlauf wird die Interreadervarianz bei der Beurteilung von CT-Bildern sinken, da ein Lernprozess stattfinden wird.

## 4. Methoden

### 4.1. Kollektiv und Untersuchungsmethoden

#### 4.1.1. Untersuchungskollektiv

Die Rekrutierung des Untersuchungskollektiv erfolgte aus den bei der Zentralstelle asbeststaubgefährdeter Arbeitnehmer (ZAs) zum Stichtag 30. Januar 1992 gemeldeten Personen. Dies waren 109.932 Personen. 35.106 mussten ausgeschlossen werden, weil bei ihnen das Erkrankungsrisiko aufgrund unzureichender Expositionsdaten nicht berechenbar war. Übrig blieben 74.106 Personen. Aus ihnen stellte das Berufsgenossenschaftliche Institut für Arbeitssicherheit (BIA) ein Hochrisikokollektiv von 3.000 Personen zusammen. Nach dem mathematischen Modell von Coenen und Schenk (Coenen und Schenk 1990) wurde für jede Person das Risiko berechnet, zwischen 1991 und 1995 entweder an einem Bronchialkarzinom oder an einem Mesotheliom zu erkranken. Es wurden jeweils die 3.000 Personen mit dem höchsten Risiko ausgewählt und dann aus beiden Risiko-Gruppen eine Schnittmenge gebildet. Diese umfasste 4.714 Personen, von denen wiederum die 3.000 Personen mit dem höchsten Gesamtrisiko als Hochrisiko-Kollektiv ausgewählt wurden. Hiervon mussten 932 ausgeschlossen werden, weil sie am Stichtag 30.9.93 > 70 Jahre alt waren und man bei ihnen eine schlechtere Compliance fürchten musste, da die Anreise häufig recht weit war, 39 mussten ausgeschlossen werden, weil sie keine weiterführenden Untersuchungen mehr wünschten und 112 weil bei ihnen bereits ein BK-Verfahren eingeleitet worden war. Übrig blieben 1.917 Personen unter 70 Jahren mit höchstem Erkrankungsrisiko, bei denen bislang keine Erkrankung festgestellt worden war und die sich nicht grundsätzlich gegen Untersuchungen ausgesprochen hatten. Ihnen wurde das Angebot einer ergänzenden arbeitsmedizinischen Vorsorgeuntersuchungsserie gemacht, die über 4 Jahre hinweg jährlich eine gewöhnliche Vorsorgeuntersuchung mit einer zusätzlichen HRCT-Untersuchungen vorsah. Von 1.166 Personen wurde keine Einverständniserklärung erhalten. Diese Personen wurden als Vergleichskollektiv betrachtet, wenn ein Vergleich aufgrund des Datenschutzes auch nur eingeschränkt möglich war. 751 waren mit den Zusatzuntersuchungen einverstanden, 115 widerriefen ihre Einverständniserklärung jedoch vor der ersten Untersuchung. Letztlich konnte die Studie mit einem Kollektiv von 636 Personen begonnen werden.

#### 4.1.2. Demographische Daten

Die Studie begann im September 1993 mit 636 Personen (52 Frauen und 584 Männer) im Alter von 53 bis 70 Jahren (Median: 63 Jahre). Die Probanden waren zwischen 1936 und 1991 mit einer medianen Dauer von 19 Jahren bzw. mittleren Dauer von 20 Jahren



Asbeststaub ausgesetzt (Interquartilabstand (IQ-A): 23, Standardabweichung: 13). Der Expositionsbeginn lag im Median bei 1950 (IQ-A: 7), das Ende bei 1970 (IQ-A: 17). Die mediane Latenzzeit betrug 43 Jahre (IQ-A: 7). 92 der Untersuchten (14,5 %) waren zum Untersuchungszeitpunkt Raucher, 348 (54,7 %) Exraucher und 196 (30,8 %) Nieraucher. Die beiden größten Gruppen hatten im Bereich der Isolation (26,9 %) oder mit Asbestzementprodukten (20,6 %) gearbeitet. Bei den meisten (86,3 %) stand eine Belastung mit Chrysotilasbest im Vordergrund, wobei gleichzeitig auch eine Exposition mit anderen Asbestsorten wie Krokydolith möglich war. Bei 18,2 % überwog das Krokydolith.

Als Vergleichskollektiv wurde die Gruppe der Personen aus dem Hochrisikokollektiv betrachtet, die keine Einverständniserklärung zu den angebotenen Zusatzuntersuchungen abgegeben hatte (1.166 Personen). Bei ihnen wurden die gewöhnlichen Vorsorgeuntersuchungen nach BG Grundsatz G 1.2 ganz normal fortgesetzt. Aufgrund des Datenschutzes war es nicht möglich, Angaben über die detaillierten Untersuchungsergebnisse zu erhalten, darum handelt es sich nicht um ein Kontrollkollektiv im strengeren Sinne. Die wichtigen Basisvariablen Alter, Geschlecht und die im Satz 1 des Untersuchungsboogens aufgeführten Parameter bezüglich Expositionsbeginn, -ende, Tätigkeitsart, -bereich, -schwerpunkt, Asbestcharakteristika, technische Hilfsmittel und Schutzmaßnahmen waren jedoch verfügbar, so dass eine Prüfung auf Strukturgleichheit möglich war. Informationen über Anzahl und Art der BK-Anzeigen und die Anzahl der Todesfälle (leider jedoch nicht die Todesursache) konnten auch in Erfahrung gebracht werden.

Es lag eine weitgehende Strukturgleichheit vor, so dass ein Vergleich der beiden Kollektive möglich war. Für nähere Betrachtungen verweise ich auf die Veröffentlichungen von Kraus et al. (1998).

### 4.1.3. Untersuchungsaufbau und Untersuchungsmethoden

Bei der vorliegenden Studie handelt es sich um eine Längsschnittstudie, die im Zeitraum von 1993-1997 durchgeführt wurde. Die durchschnittlichen Untersuchungsabstände der vier Untersuchungen pro Person lagen bei etwa zwölf Monaten (zwischen zehn und vierzehn Monaten). Die Untersuchungen wurden in den in Tab. 4.1 aufgeführten sechs Zentren durchgeführt. Bei Verdacht auf eine asbeststaubverursachte Berufskrankheit wurde eine BK-Anzeige erstattet, der Patient verblieb aber weiterhin im Untersuchungskollektiv.

*Tabelle 4.1.: Die Untersuchungszentren der Studie*

<b>Hamburg-Altona:</b>	Radiologische Abteilung des Allgemeinen Krankenhauses
<b>Heidelberg-Rohrbach:</b>	Radiologische Abteilung der Thoraxklinik
<b>Bochum:</b>	Institut für Radiologie der Ruhr-Universität
<b>Dortmund:</b>	Radiologische Abteilung des Knappschaftskrankenhauses
<b>Bad Reichenhall:</b>	Klinik für Berufskrankheiten
<b>Erlangen-Nürnberg:</b>	Institut für Arbeits-, Sozial- und Umweltmedizin der Universität

Bei der konventionellen Vorsorge sind Kriterien für die Erstattung einer BK-Anzeige festgelegt (s. Abs. 2.4), während keine Kriterien für CT oder HRCT-Befunde existierten. Darum wurden vor Studienbeginn die in Tab. 4.2 dargestellten Empfehlungen von einer interdisziplinären Arbeitsgruppe erarbeitet.

**Tabelle 4.2.:** Kriterien zur Erstattung einer BK-Anzeige anhand von CT-Befunden

Parenchym (Streuung)	Form	Auskultations- bzw. Lungenfunktionsbefund
0/1	s'/t'/u' und/oder v'/w'/x'	Knisterrasseln und/oder VKI $\leq$ 90 % von VKS nach EGKS Mindestsollwert unter BTPS-Bedingungen
$\geq$ 1	s'/t'/u' und/oder v'/w'/x'	auch wenn keine klinischen Einschränkungen und keine Einschränkungen der VKI messbar sind
Pleurabefund	Dicke/Ausdehnung	Anamnese
hyaline Pleuraplaques mit/ohne Verkalkung	unabhängig (auch einseitig)	anamnestische Hinweise auf Asbeststaubexposition
Hyalinosis complicata bzw. Pleuraerguss	unabhängig	
gleichförmige Pleuraverdickung	Ausdehnung $\geq$ 2cm; Dicke $\geq$ 3mm im Bereich beider Mittel-/Unterfelder	

Auch bei malignen Erkrankungen verblieb der Patient im Untersuchungskollektiv. Es erfolgte ein enger Informationsaustausch zwischen dem betreuenden Arzt und den Berufsgenossenschaften, so dass eine lückenlose Dokumentation möglich war. Bei Aussteigern versuchte man, den Grund zu erfahren, um mögliche Verzerrungen (wie gehäuftes Aussteigen durch maligne Erkrankungen) zu vermeiden, bei Verstorbenen wurde die Todesursache ermittelt.

### Untersuchungsprogramm

Bezüglich des detaillierten Untersuchungsprogramms darf auf Kraus et al. (1998) verwiesen werden. Die Röntgenuntersuchung wurde mittels Hartstrahltechnik in posterior-anterior(p.a.)-Position und seitlich links anliegend durchgeführt. In Einzelfällen wurde auf aktuelle Aufnahmen anderer Untersucher zurückgegriffen. Die Befunde wurden gemäß ILO-Staublungenklassifikation von 1980 auf Satz III des G-1.2-Bogens festgehalten. Die CT-Aufnahmen erfolgten in Hochauflösungstechnik (HRCT) mit den in Tab. 4.3 genannten Geräten unter standardisierten Untersuchungsbedingungen (Tab. 4.4).

**Tabelle 4.3.:** CT-Geräte der Studie

Picker PQ 2000 in Bad Reichenhall	Siemens Somatom plus in Hamburg
Philips Tomoscan SR 7000 in Bochum	Siemens Somatom plus 4 in Erlangen
Siemens Somatom DR in Heidelberg	Siemens Somatom ART in Dortmund

**Tabelle 4.4.:** Verwendete CT-Einstellparameter (W = Window, C = Center, H.Ü. = Hounsfield Units)

1. Rückenlage		2. Bauchlage	
Algorithmus	High	Algorithmus	Ultra High
Schichtdicke	1 mm	Schichtdicke	1 mm
Scan-Zeit	1 s	Scanzeit	1 s
Tischvorschub	10 mm	Schichten	obligatorisch: Lungenunterfeld, fakultativ: Lungenmittel- bzw. -oberfeld
Fenster	a) W = 500 H.Ü.; C = 40 H.Ü. b) W = 1.500 H.Ü.; C = -400 H.Ü.	Fenster	W = 2.000 H.Ü.; C = -400 H.Ü.

Somit erhielt man pro Proband etwa 24 Schichtaufnahmen, deren Beurteilung in einem EDV-kompatiblen Kodierungssystem vermerkt wurde, das in Anlehnung an die ILO-

Klassifikation entwickelt worden war und hier das erste Mal breitere Anwendung fand (Kraus et al. 1996).

Bei Malignomverdacht erfolgte eine Spiraltechnik-Untersuchung, mit der eine lückenlose Darstellung des Thoraxraumes möglich war. War eine malignomverdächtige Veränderung im zeitlichen Verlauf feststellbar, wurde sofort eine invasiv-diagnostische Abklärung durchgeführt.

### Ablauf und Compliance

Die Drop-Out-Rate lag zwischen 7,4 und 14,4 % pro Jahr mit großen Unterschieden zwischen den einzelnen Untersuchungszentren. Die insgesamt niedrigste Drop-out-Rate von 2,0 % gab es in Erlangen, die höchste mit 37,5 % in Bad Reichenhall. Es gab sowohl Probanden, die komplett ausstiegen als auch solche, die für einen oder zwei Termine (=Jahre) aussetzten. Bei komplett-Aussteigern wurden die Gründe erfragt, um Verzerrungen zu vermeiden. Es stellte sich heraus, dass 63 % von ihnen keine asbestassoziierten Veränderungen aufwiesen. Komplette Datensätze konnten von 380 Probanden erstellt werden. Damit war die Compliance mit 59,7 % trotz des großen Aufwands und den langen Anreisen überdurchschnittlich gut. In Tab. 4.5 ist die Anzahl der untersuchten Probanden im Zeitraum 1993 bis 1997 für jedes Zentrum aufgeführt.

*Tabelle 4.5.: Anzahl der untersuchten Probanden im Zeitraum von 1993 bis 1997*

Jahr	Zahl der untersuchten Personen			
	1. Querschnitt 1993/1994	2. Querschnitt 1994/1995	3. Querschnitt 1995/1996	4. Querschnitt 1996/1997
gesamt	621	543	484	448
Bochum	109	86	75	67
Bad Reichenhall	16	16	10	9
Dortmund	143	129	117	103
Erlangen	100	98	86	81
Heidelberg	110	89	79	74
Hamburg	143	125	117	114

#### 4.1.4. Zweitbeurteilung

Alle Röntgenbilder und CT-Aufnahmen eines Probanden (jährliche Röntgen- und CT-Aufnahmen) wurden von zwei Beurteilern ausgewertet, wobei dem Zweitauswerter das Ergebnis des Erstauswerter unbekannt war. Die CT/Rö-Aufnahmen wurden im Abstand von mehreren Wochen zu den Beurteilern geschickt, so dass sie sich nicht mehr erinnern konnten, wie sie die andere Aufnahme des Probanden beurteilt hatten. Nach zwei Untersuchungszyklen wechselte der Zweitbeurteiler, um systematische Fehler zu minimieren. Trat eine starke Diskrepanz bezüglich BK-relevanter Befunde oder einem Malignomverdacht auf, kam es zu einer Abstimmung der Befunde unter Beteiligung aller Beurteiler dieser Studie. In Tab. 4.6 sind alle Beurteiler dieser Studie aufgeführt. Sie hatten sich durch besonders großen Erfahrungsreichtum in der Interpretation von Pneumokoniosebildern qualifiziert und bildeten zehn Beurteilerpaare. In Tab. 4.7 ist die

**Tabelle 4.6.:** die Beurteiler der Studie mit Orts- und Qualifikationsangabe

Bochum	Dr. med. V. Wiebe (1)	Radiologe
Bad Reichenhall	Dr. med. W. Raab (2)	Arbeitsmediziner
Dortmund	Dr. med. K. G. Hering (3)	Radiologe
Erlangen	Dr. med. T. Kraus (4)	Arbeitsmediziner
Heidelberg	PD Dr. med. S. Tuengerthal (5)	Radiologe
Hamburg	Prof. Dr. med. O. Wegener (6)	Radiologe
Erlangen	Prof. Dr. med. Dr. rer. nat. H.-J. Raithel (7)	Arbeitsmediziner

Anzahl der befundeten Personen pro Beurteilerpaar und Jahr dargestellt. Einige Patienten mussten aufgrund fehlender Angaben aus dem auswertbaren Kollektiv entfernt werden (fehlende Angaben zu den Beurteilern, fehlende Bögen, etc.). Dies wird in der Tabelle in den letzten beiden Spalten dargestellt.

**Tabelle 4.7.:** Beurteilerpaare: Anzahl beurteilter Personen

Beurteiler	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	54	33		
1 + 6	54	50	88	
1 + 7			62	66
2 + 4	16	16	13	11
3 + 2			40	47
3 + 4	243	218	59	53
3 + 5			68	66
4 + 5	109	77		8
6 + 4	73	64	70	187
6 + 5	70	60		
$\kappa$ errechenbar von ... Personen	619	518	400	438
Zahl der insgesamt untersuchten Personen	621	543	484	448

#### 4.1.5. Röntgen versus HRCT

Im folgenden werden Ergebnisse der Langzeitstudie vorgestellt, die bereits vor Erstellen dieser Arbeit vorlagen und mit in die Methodenplanung einbezogen wurden.

##### Befunde in konventionellen Röntgenaufnahmen

Eindeutig asbestotische Veränderungen des Parenchyms (ILO-Kodierung  $s/t/u \geq 1/1$ ) fanden sich nur in 3,4 %, während grenzwertige Befunde ( $s/t/u$  von  $1/0$ ) in 10 % der Fälle diagnostiziert wurden (für Details siehe Kraus et al. 1998).

In drei Fällen wurde ein Pleuraerguss festgestellt, der eine weitergehende Abklärung veranlasste. Diffuse Pleuraverdickungen konnten bei 33 %, umschriebene Pleuraverdickungen in 25,4 % und Pleuraverkalkungen in 14,7 % nachgewiesen werden.

##### Computertomographische Befunde

Mittels Computertomographie wurden mehr sicher krankhafte Befunde erhoben. Eindeutig asbestotische Veränderungen des Parenchyms (CT-Bogen  $s',t',u' \geq 1/1$ ) fanden

sich in 35,4 %, während grenzwertige Befunde (s´,t´,u´ von 0/1) nur in 6,1 % der Fälle diagnostiziert wurden. Für v´,w´,x´ gelten ähnliche Zahlen, nur dass hier noch zusätzlich mehr grenzwertige Befunde festgestellt wurden. Die Parenchym- und Pleurabefunde sind in Tab. 4.8 und Tab. 4.9 dargestellt.

**Tabelle 4.8.:** Parenchymveränderungen CT in % (in Klammern Anzahl); p´q´r´ = kleine runde Fleckschatten, s´t´u´ = kleine irreguläre Fleckschatten, v´w´x´ = lineare Fleckschatten, vgl. Tab. 2.11

	0	0/1	1	2	3
p´q´r´	95,5 (593)	1,1 (7)	3,1 (19)	0,2 (1)	0,2 (1)
s´t´u´	58,5 (363)	6,1 (38)	30 (186)	2,7 (17)	2,7 (17)
v´w´x´	52,7 (327)	12,1 (75)	33,8 (210)	1,1 (7)	0,3 (2)

**Tabelle 4.9.:** Plaqueartige und gleichförmige Pleuraverdickungen in % (Anzahl in Klammern)

	Thoraxwand	Mediastinum	Zwerchfell
Plaques	56,8 (353)	8,7 (54)	29,3 (182)
gleichförmige Pleuraverdickungen	27,1 (168)	0,4 (3)	1,1 (7)

## Vergleich Röntgen und HRCT

Bei einem Vergleich der Röntgen- und CT-Befunde bezüglich der parenchymalen Veränderungen ließ sich eine Übereinstimmung von 73,2 % feststellen. Als „positiv“ war ein Befund definiert, der bei der ILO-Klassifikation der irregulären Schatten mit  $\geq 1/0$  gemäß ILO gewertet worden war. Auch computertomographisch galt ein Befund  $\geq 1$  als „positiv“. Bei 19,3 % der Patienten wurden mittels CT positive Befunde erlangt, die röntgenologisch als negativ bewertet wurden. 7,6 % der Röntgenbilder wurden falsch-positiv bewertet.

Bezüglich der Pleura fand man eine etwas schlechtere Übereinstimmung von 69,1 %. Als „positiv“ galt ein röntgenologischer Befund, wenn eine plaqueartige und/oder diffuse Pleuraverdickung oder Verkalkung diagnostiziert wurde. Computertomographisch wurden umschriebene und/oder gleichförmige Pleuraverdickungen mit oder ohne Verkalkungen als positiver Befund gewertet. Bei 18,4 % erhielt man CT-positive Befunde, wo sich im Röntgen nichts nachweisen ließ, röntgenologisch falsch-positive Befunde bei 12,5 %. Letztere waren bei 67 von 80 Probanden auf Fett, das röntgenmorphologisch schlecht von Plaques abzugrenzen ist, während dies bei CTs aufgrund der Hounsfield-Fensterungsmöglichkeiten klar differenziert werden kann, und bei 6 von 80 Probanden auf plaqueartige Zwerchfellverdickungen zurückzuführen.

### Vergleich von Röntgen und HRCT bezüglich der Erstattung von BK-Anzeigen

Nach den für Röntgenbefunde bestehenden Kriterien (s.o.) hätten nach dem ersten Untersuchungszyklus bei 89 von 621 Patienten aufgrund parenchymaler Veränderungen und bei 189 von 621 aufgrund pleuraler Veränderungen BK-Anzeigen erstattet werden müssen. Insgesamt hätten nach dem 1. Zyklus 189 Anzeigen (30,4 %) schon allein aufgrund röntgenologischer Befunde in Kombination mit klinischen Befunden erstattet werden müssen.

Computertomographisch wurden mehr pathologische Befunde erhoben (s.o.), so dass es nach den selbst festgelegten Kriterien (s.o.) bei 304 der 621 Probanden (48,9 %) zu einer BK-Anzeige kommen musste.

Insgesamt wurden im Laufe der vierjährigen Studie bei 52,4 % BK-Anzeigen erstattet, womit die Dunkelziffer gegenüber der konventionellen Methode deutlich reduziert werden konnte. Zudem konnte bei 7,2 % eine BK-Anzeige aufgrund falsch-positiver Röntgenbefunde verhindert werden, was nicht nur Zusatzuntersuchungen eingespart hat, sondern auch von großem Wert für die betroffenen Personen war, die so vor einem „Fehlalarm“ bewahrt wurden.

Hinsichtlich BK-relevanter Befunde weisen die Röntgenaufnahmen gegenüber den CT-Aufnahmen eine Sensitivität (Kranke werden als krank erkannt) von 47,4 %, eine Spezifität (Gesunde werden als gesund erkannt) von 85,8 %, einen positiven prädiktiven Wert von 76,2 % und einen negativen prädiktiven Wert von 63 % auf (positiver prädiktiver Wert: Anteil der tatsächlich Kranken unter den als krank befundeten, negativer prädiktiver Wert: Anteil der tatsächlich Gesunden unter den als gesund befundeten).

### Effektivität des HRCT-Einsatzes bezüglich BK-relevanter Befunde

Bei einem Vergleich beider Kollektive zeigte sich, dass in der CT-Kohorte die BK-Anzeigen in Bezug auf die BK-Nr. 4103 7,5 mal, auf die BK-Nr. 4104 4,3 mal und auf die BK-Nr. 4105 5,5 mal höher lagen. Insgesamt wurden im Vergleichskollektiv nur bei 6,6 % der Personen BK-Anzeigen erstattet.

## 4.2. Kategorisierung

Zur Befundung der Röntgen- und CT-Aufnahmen wurden zwei verschiedene Fragebögen verwendet. Während für die Befundung der Röntgen-Aufnahmen die altbewährten ILO-1980-Untersuchungsbögen verwendet wurden (s. Abs. 2.4.1), wurden die CT-Kodierungsbögen von einem interdisziplinären Team in Anlehnung an die ILO-Bögen neu entwickelt (s. Abs. 2.4.3).

Ziel dieser Arbeit war es, die Übereinstimmung der Beurteilung eines Bildes durch zwei Beurteiler zu ermitteln. Dafür wurde der  $\kappa$ -Wert ermittelt (s. Abs. 2.5.1). Beide Fragebögen enthalten ca. 300 Variablen. Die Berechnung von  $\kappa$  ist sinnlos, wenn die Anzahl der Merkmalausprägungen zu hoch ist. Aus diesem Grunde ist es nötig, Kategorien zu bilden, die gleichwertige Aussagen zusammenfassen und deren Auswertung zu übersichtlichen und aussagekräftigen Ergebnissen führt. Für die Kategorienbildung gibt

es keine internationalen Standards oder Richtlinien (s. Abs. 2.5.2). Bisher wurde die Kategorienbildung durch verschiedene Arbeitsgruppen auf sehr unterschiedliche Arten durchgeführt. In der vorliegenden Studie war es wichtig, für beide Verfahren möglichst ähnliche Kategorien zu bilden, um die beiden Verfahren miteinander vergleichen zu können. Dies wurde dadurch erschwert, dass der CT-Fragebogen zwar in Anlehnung an den Röntgenbogen entwickelt wurde, aber deutlich detaillierter und genauer ist. Die Filmqualität wurde bei der Berechnung nicht berücksichtigt, da nicht davon ausgegangen werden konnte, dass die Qualität der Aufnahmen Einfluss auf die Übereinstimmung der Befundung hatte, da durchweg hochwertige Geräte eingesetzt worden waren und davon auszugehen war, dass die Qualität der Bilder sehr gut war.

Im Folgenden wird die Kategorienbildung erläutert und begründet, wobei auf die Erklärung der Fragebögen in den Abschnitten 2.4.1 und 2.4.3 verwiesen wird.

### 4.2.1. Röntgen-Kodierungsbogen

#### Parenchymveränderungen

Das Auswahlverfahren, bei dem nur Probanden zugelassen wurden, für die noch kein BK-Verfahren eingeleitet worden war, hatte zur Folge, dass ein Subkollektiv betrachtet wurde, bei dem sich überwiegend keine bis geringe Veränderungen fanden. Große Schatten wurden fast nie diagnostiziert, weshalb diese von einer Berechnung des  $\kappa$ -Wertes ausgeschlossen wurden. Dementsprechend wurden  $\kappa$ -Werte für die kleinen rundlichen Fleckschatten und für die unregelmäßigen Fleckschatten ermittelt, für die es jeweils Angaben über Größe, Streuung und Lokalisation gab. Die drei möglichen Größenangaben wurden zu einer positiven Aussage zusammengefasst, so dass dichotom ausgesagt werden konnte, ob die betreffende Veränderung aufgetreten war oder nicht.

Im ILO-Bogen gab es zwölf mögliche Streuungsdichtegrade, im CT-Bogen lediglich fünf. Die ILO formuliert eine präzise Definition der vier Hauptkategorien. Die Einteilungen dazwischen werden nicht definiert, sondern als fließende Übergänge betrachtet und vermehren nur die Unterscheidungsmöglichkeiten. Dementsprechend sprach nichts dagegen, die Anzahl der Streuungsvariablen wieder auf die Hauptkategorien zu dezimieren.

Der CT-Bogen macht eine zusätzlich Unterscheidung zwischen dem Befund 0 und dem Befund 0/1. Diese Einteilung spiegelte die relevanten Aspekte für die Erstattung einer BK-Anzeige wieder. Ein Befund von  $\geq 1$  galt als definitiv krank und musste angezeigt werden. Ein Befund von 1/0 (Rö) bzw. 0/1 (CT) galt als vielleicht krank und musste nur in Zusammenhang mit bestimmten körperlichen Befunden angezeigt werden. Die Entscheidung, ob ein Befund als 0/1 (bzw. 1/0) oder 1 eingestuft wurde, konnte demnach im Einzelfall von großer Relevanz sein. Deshalb war es nach Meinung der Autoren nicht angebracht, den Befund 0/1 der Kategorie 0 zuzuordnen, wie es in vielen Studien üblich war (2.5.2). Die Streuungs-Kategorien des CT-Bogens wurden auf den Röntgenbogen übertragen:

- Kategorie 0: 0/- und 0/0
- Kategorie 0/1

- Kategorie 1: 1/0, 1/1 und 1/2
- Kategorie 2: 2/1, 2/2 und 2/3
- Kategorie 3: 3/2, 3/3 und 3/+.

Auf dem CT-Bogen wurde jedes Lungenfeld einzeln betrachtet, während auf dem Röntgenbogen ein Integrationsprozess über die ganze Lunge gefordert wurde. Bei zu großen Differenzen sollten die Felder mit den geringsten Veränderungen ignoriert werden. Dementsprechend konnte kein Rückschluss auf die genaue Lokalisation der Veränderungen auf den Röntgenbildern gezogen werden, weshalb die Lokalisationsangaben in beiden Verfahren nicht ausgewertet werden konnten.

### **Pleuraveränderungen**

Bezüglich der Pleura wurden folgende Aspekte ausgewertet: die Verbreitung und Dicke von diffuser oder umschriebener Pleuraverdickung und die Verbreitung von Pleuraverkalkungen. Die Einteilungen des CT-Bogens entsprachen denen des ILO-Bogens. Lediglich die Aussagen zur Pleuraverkalkung waren auf dem CT-Bogen nur dichotom, so dass die entsprechenden positiven Angaben auf dem ILO-Bogen (1-3) zusammengefasst wurden. Für umschriebene und diffuse Pleuraverdickungen gab es drei Größenangaben (Verbreitung (1-3) und Dicke (a-c)), die nicht weiter zusammengefasst wurden. Lokalisationsangaben konnten aus den selben Gründen wie bei der Beurteilung des Parenchyms nicht berücksichtigt werden.

## **4.2.2. CT-Kodierungsbogen**

### **Parenchymveränderungen**

Die Kategorienbildung des CT-Bogens wurde mit der des ILO-Bogens abgestimmt. Ausgewertet wurden die Übereinstimmungen bei:

- kleine runde Fleckschatten
- kleine irreguläre Fleckschatten
- lineare Fleckschatten
- Honeycombing und
- Emphysem.

Die Größenangaben wurden wie beim Röntgen daraufhin analysiert, ob eine Veränderung vorhanden war (mind. 1 Kreuz) oder nicht (kein Kreuz). Die Streuungsangaben wurden nicht weiter zu Kategorien zusammengefasst und entsprachen somit den oben gebildeten Röntgen-Streuungskategorien.



Ein Problem entstand dadurch, dass auf dem CT-Bogen für jedes Lungenfeld eine eigene Angabe zur Streuung existierte. Hier fand kein „Integrationsprozess“ statt, wie er bei den ILO-Bögen gefordert wurde. Um die beiden Verfahren vergleichbar zu machen, mussten die Streuungsangaben von den Lokalisationsangaben getrennt werden. Der Integrationsprozess wurde nachempfunden, indem die stärkste angegebene Streuung gewählt wurde.

### **Pleuraveränderungen**

Im CT lassen sich, im Gegensatz zum Röntgen, Pleuraverdickungen gut voneinander abgrenzen. Dabei stellte sich heraus, dass gleichförmige Pleuraverdickungen selten kodiert wurden. Sowohl die umschriebenen als auch die gleichförmigen (entspricht „diffus“ auf dem Röntgenbogen) Pleuraverdickungen wurden bezüglich Ausdehnung, Dicke und Verkalkung gemäß dem ILO-Bogen in drei Größenangaben bewertet. Die beiden „Kalk“-Rubriken unter „umschrieben“ und „gleichförmig“ wurden zusammengefasst, um entsprechend dem ILO-Bogen lediglich eine (dichotome) Rubrik „Kalk“ zu erhalten. Wie beim Parenchym wurde die stärkste genannte Veränderung ermittelt und ausgewertet, um den Integrationsprozess zu imitieren.

## **4.3. Datenerfassung und -auswertung**

Die Röntgen-Befunde wurden auf den ILO-Bögen notiert und direkt von dort durch Dokumentationsassistentinnen in ein spezielles EDV-Datenerfassungssystem übertragen. Die CT-Ergebnisse wurden auf den oben erwähnten Befundungsbögen notiert und entweder durch den Arzt selbst oder durch Dokumentationsassistentinnen in ein selbstentwickeltes EDV-System übertragen. Die Einträge wurden daraufhin umfangreichen Plausibilitätskontrollen unterworfen, die teilweise in die EDV-Programme integriert waren. Traten unplausible Daten auf, kontrollierte der Arzt die Originaldaten. War eine Korrektur nicht möglich, wurde der komplette Datensatz entfernt.

Die statistische Auswertung erfolgte in Zusammenarbeit mit den Mitarbeitern des Instituts für Medizinische Statistik der Medizinischen Fakultät der RWTH Aachen und mit Hilfe von SPSS-Software, Excel und MedCalc. Die Graphen wurden mit Hilfe von Gnuplot erstellt, die Arbeit mittels MiKTeX verfasst. Die Daten lagen im SPSS-Format vor. Nach oben genannten Kategorisierungsvorstellungen wurden die Daten mit Hilfe von SPSS verdichtet.

Dort, wo Beurteilerangaben fehlten oder keine Bögen ausgefüllt waren (von einem oder von beiden Beurteilern), wurden die Patientendaten komplett aus der statistischen Analyse ausgeschlossen. Lücken im Bogen wurden folgendermaßen ausgefüllt: ließ sich im Kontext feststellen, dass lediglich vergessen worden war, eine 0 anzukreuzen, wurde diese eingesetzt. Widersprachen sich Angaben, wurde dieser Aspekt bei der betroffenen Person bei beiden Beobachtern gestrichen. Wenn zu einem Aspekt generell nichts angekreuzt wurde, konnte man davon ausgehen, dass der Beurteiler vergessen hatte, eine 0 zu setzen. Wenn allerdings an einer Stelle nichts angekreuzt wurde (z.B. Größe der Ver-

änderung), aber an einer anderen Stelle (z.B. Streuung dieser Veränderung) ein Kreuz gesetzt wurde, widersprachen sich diese Angaben und man musste die Fragestellung mit fehlender Angabe entfernen.

Die verdichteten Daten wurden in Excel umgeschrieben und anschließend in MedCalc eingelesen. MedCalc errechnete die  $\kappa$ -Werte. In einzelnen Fällen kam es vor, dass ein Beurteilerpaar nur einen oder zwei Patienten beurteilte. Hier ließ sich kein  $\kappa$  ermitteln. Diese Paare wurden als nicht existent betrachtet. Das Paar mit den wenigsten Patienten war das Paar 4+5, das im 4. Jahr lediglich 8 Patienten beurteilt hat. In einem Sonderfall wurden die Daten modelliert: haben beide Beurteiler nur '0' angekreuzt, so konnte das Computerprogramm keinen  $\kappa$ -Wert ermitteln. Dennoch war die Übereinstimmung hier perfekt, weshalb in diesen Fällen ein  $\kappa$  von 1 eingesetzt wurde. Für jedes Jahr wurde über alle Beurteilerpaare das Minimum, das Maximum, der Median, der Mittelwert, der Interquartilabstand (IQ-Abstand) und die Standardabweichung (s) errechnet.

1) **Maximal-Wert**: die beste Übereinstimmung. Zum Verfahrensvergleich kam dieses Vorgehen nicht in Betracht, da es zu beinahe jeder Fragestellung ein Paar mit perfekter Übereinstimmung gab. Damit hatten alle Fragestellungen den gleichen Maximalwert und unterschieden sich nach diesem Kriterium nicht. Zudem ging es in dieser Studie um die Verteilung der Übereinstimmungen, die so nicht berücksichtigt würden.

2) **Median**: ist der Wert, der bei aufsteigender Anordnung in mittlerer Position liegt. Das heißt, wenn es fünf Größenangaben gibt, stellt die dritte den Median dar. Gibt es sechs, wird das arithmetische Mittel aus der dritten und der vierten gebildet. Somit wird verhindert, dass Größenangaben, die stark aus der Reihe fallen, das Ergebnis zu sehr in ihre Richtung ziehen. Nachdem durch die Medianermittlung erst einmal der Einfluss der Extremwerte eliminiert worden war, wurde fortan bei weiterer Mittlung der statistische Mittelwert ermittelt.

3) **Mittelwert**: das arithmetische Mittel. Ein stark abweichender Wert verzieht den Mittelwert stark in seine Richtung. In der vorliegenden Studie war es interessant, auch die extremeren Werte mit zu betrachten, weil sie auf eine Bewegung hindeuten konnten.

4) **Interquartil-(IQ)-Abstand**: Der Interquartilabstand ist der Abstand, der zwischen dem 25 %- und dem 75 %-Quartil (also bei fünf Werten der 2. und der 4. Wert) besteht. Der Wert sagt nichts über die Qualität der Beurteilungen aus, lediglich über ihre Streuung.

5) **Standardabweichung (s)**: Die Standardabweichung (s) wird von der Varianz (VAR), dem arithmetischen Mittelwert aller quadrierten Abweichungen der Daten von ihrem arithmetischen Mittelwert, abgeleitet. Sie ist als die Quadratwurzel aus der Varianz definiert. Damit macht die Standardabweichung Aussagen über das Ausmaß, in dem die Werte um den Mittelwert streuen. Normalverteilung vorliegen. Dies ist hier nicht der Fall. einen größeren Einfluss auf die Standardabweichung als auf die durchschnittliche Abweichung oder den Quartil-Abstand.

# 5. Ergebnisse, Diskussion und Ausblick

## 5.1. $\kappa$ -Werte bezogen auf Paar und Fragestellung

Für jedes Jahr wurde über alle Beurteilerpaare das Minimum, das Maximum, der Median, der Mittelwert und die 25- und 75 %-Perzentile der  $\kappa$ -Werte errechnet. Die Ergebnisse sind in den Tabellen 5.0(a) bis 5.1(o) dargestellt. Tabelle 5.3 stellt die mittleren  $\kappa$ -Werte der einzelnen Paare dar.

**Tabelle 5.1.:**  $\kappa$ -Werte: Röntgen ( $s$  = Standardabweichung)

(a) Rö: kleine runde Fleckschatten vorhanden/nicht vorhanden

Beurteiler	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	-0,03	0,00		
1 + 6	0,26	0,00	0,00	
1 + 7			0,00	1,00
2 + 4	1,00	0,00	1,00	1,00
3 + 2			0,00	-0,02
3 + 4	0,04	0,05	0,00	0,00
3 + 5			-0,02	0,00
4 + 5	1,00	1,00		1,00
6 + 4	0,00	0,00	-0,02	0,04
6 + 5	0,00	-0,03		
Minimum	-0,03	-0,03	-0,02	-0,02
25 %-Perz.	0,00	0,00	-0,02	0,00
<b>Median</b>	<b>0,04</b>	<b>0,00</b>	<b>0,00</b>	<b>0,04</b>
75 %-Perz.	1,00	0,05	0,00	1,00
Maximum	1,00	1,00	1,00	1,00
IQ-Abstand	1,00	0,05	0,02	1,00
Mittelwert	0,32	0,15	0,14	0,43
s	0,47	0,38	0,38	0,53

(b) Rö: Streuung kleiner runder Fleckschatten

Beurteiler	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	-0,03	0,00		
1 + 6	0,15	0,00	0,00	
1 + 7			0,00	1,00
2 + 4	1,00	0,00	1,00	1,00
3 + 2			0,00	-0,02
3 + 4	0,06	0,07	0,00	0,00
3 + 5			-0,02	0,00
4 + 5	1,00	1,00		1,00
6 + 4	0,00	0,00	-0,02	0,02
6 + 5	0,00	-0,02		
Minimum	-0,03	-0,02	-0,02	-0,02
25 %-Perz.	0,00	0,00	-0,02	0,00
<b>Median</b>	<b>0,06</b>	<b>0,00</b>	<b>0,00</b>	<b>0,02</b>
75 %-Perz.	1,00	0,07	0,00	1,00
Maximum	1,00	1,00	1,00	1,00
IQ-Abstand	1,00	0,07	0,02	1,00
Mittelwert	0,31	0,15	0,14	0,43
s	0,47	0,38	0,38	0,53

(c) R : kleine unregelm ssige Fleckschatten vor-  
handen/nicht vorhanden

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	-0,05	0,08		
1 + 6	0,16	0,26	0,35	
1 + 7			0,13	0,13
2 + 4	0,74	0,88	0,49	1,00
3 + 2			0,17	0,19
3 + 4	0,36	-0,02	0,18	0,38
3 + 5			0,17	0,16
4 + 5	0,41	0,23		0,39
6 + 4	0,10	0,16	0,22	0,14
6 + 5	0,08	0,17		
Minimum	-0,05	-0,02	0,13	0,13
25 %-Perz.	0,08	0,08	0,17	0,15
<b>Median</b>	<b>0,16</b>	<b>0,17</b>	<b>0,18</b>	<b>0,19</b>
75 %-Perz.	0,41	0,26	0,35	0,39
Maximum	0,74	0,88	0,49	1,00
IQ-Abstand	0,33	0,18	0,18	0,24
Mittelwert	0,26	0,25	0,24	0,34
s	0,27	0,29	0,13	0,31

(d) R : Streuung kleiner unregelm ssiger Fleckschatten

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	-0,04	0,12		
1 + 6	0,16	0,29	0,31	
1 + 7			0,22	0,15
2 + 4	0,68	0,81	0,59	0,91
3 + 2			0,24	0,14
3 + 4	0,45	0,80	0,23	0,46
3 + 5			0,26	0,26
4 + 5	0,48	0,54		0,31
6 + 4	0,13	0,54	0,28	0,20
6 + 5	0,14	0,10		
Minimum	-0,04	0,10	0,22	0,14
25 %-Perz.	0,13	0,12	0,23	0,15
<b>Median</b>	<b>0,16</b>	<b>0,54</b>	<b>0,26</b>	<b>0,26</b>
75 %-Perz.	0,48	0,80	0,31	0,46
Maximum	0,68	0,81	0,59	0,91
IQ-Abstand	0,35	0,68	0,08	0,31
Mittelwert	0,29	0,46	0,30	0,35
s	0,25	0,30	0,13	0,27

(e) R : diffuse Pleuraverdickung: Dicke

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,00	0,00		
1 + 6	0,00	0,00	0,00	
1 + 7			0,00	0,00
2 + 4	0,90	0,59	0,71	0,56
3 + 2			0,18	0,21
3 + 4	0,44	0,22	-0,01	0,36
3 + 5			0,33	0,29
4 + 5	0,41	0,13		0,69
6 + 4	0,31	0,25	0,17	0,23
6 + 5	0,04	0,10		
Minimum	0,00	0,00	-0,01	0,00
25 %-Perz.	0,00	0,00	0,00	0,21
<b>Median</b>	<b>0,31</b>	<b>0,13</b>	<b>0,17</b>	<b>0,29</b>
75 %-Perz.	0,44	0,25	0,33	0,56
Maximum	0,90	0,59	0,71	0,69
IQ-Abstand	0,44	0,25	0,33	0,35
Mittelwert	0,30	0,18	0,20	0,33
s	0,33	0,20	0,26	0,23

(f) R : diffuse Pleuraverdickung: Ausdehnung

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,00	0,00		
1 + 6	0,00	0,00	0,00	
1 + 7			0,00	0,00
2 + 4	0,95	0,88	0,78	0,64
3 + 2			0,36	1,00
3 + 4	0,44	0,25	-0,05	0,27
3 + 5			0,36	0,42
4 + 5	0,42	0,06		0,81
6 + 4	0,44	0,25	0,31	0,26
6 + 5	-0,01	0,04		
Minimum	-0,01	0,00	-0,05	0,00
25 %-Perz.	0,00	0,00	0,00	0,26
<b>Median</b>	<b>0,42</b>	<b>0,06</b>	<b>0,31</b>	<b>0,42</b>
75 %-Perz.	0,44	0,25	0,36	0,81
Maximum	0,95	0,88	0,78	1,00
IQ-Abstand	0,44	0,25	0,36	0,55
Mittelwert	0,32	0,21	0,25	0,48
s	0,35	0,31	0,30	0,35

(g) R : umschriebene Pleuradickungen: Dicke

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,49	0,21		
1 + 6	0,43	0,48	0,17	
1 + 7			0,40	0,37
2 + 4	0,87	0,76	0,91	0,60
3 + 2			-0,03	0,29
3 + 4	0,51	0,47	0,09	0,53
3 + 5			0,27	0,45
4 + 5	0,51	0,78		0,28
6 + 4	0,32	0,51	0,40	0,42
6 + 5	0,12	0,39		
Minimum	0,12	0,21	-0,03	0,28
25 %-Perz.	0,32	0,39	0,09	0,29
<b>Median</b>	<b>0,49</b>	<b>0,48</b>	<b>0,27</b>	<b>0,42</b>
75 %-Perz.	0,51	0,76	0,40	0,53
Maximum	0,87	0,78	0,91	0,60
IQ-Abstand	0,19	0,37	0,31	0,24
Mittelwert	0,46	0,51	0,32	0,42
s	0,23	0,20	0,31	0,12

(h) R : umschriebene Pleuravedickungen: Ausdehnung

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,66	0,25		
1 + 6	0,37	0,53	0,30	
1 + 7			0,36	0,45
2 + 4	1,00	0,90	1,00	0,57
3 + 2			-0,03	0,50
3 + 4	0,59	0,41	0,00	0,42
3 + 5			0,33	0,29
4 + 5	0,49	0,86		0,28
6 + 4	0,41	0,62	0,33	0,53
6 + 5	0,09	0,42		
Minimum	0,09	0,25	-0,03	0,28
25 %-Perz.	0,37	0,41	0,00	0,29
<b>Median</b>	<b>0,49</b>	<b>0,53</b>	<b>0,33</b>	<b>0,45</b>
75 %-Perz.	0,66	0,86	0,36	0,53
Maximum	1,00	0,90	1,00	0,57
IQ-Abstand	0,29	0,45	0,36	0,24
Mittelwert	0,52	0,57	0,33	0,43
s	0,28	0,24	0,34	0,11

(i) R : Kalk vorhanden/nicht vorhanden

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,00	0,65		
1 + 6	0,84	0,73	0,52	
1 + 7			0,76	0,65
2 + 4	1,00	0,60	1,00	0,56
3 + 2			0,47	0,56
3 + 4	0,44	0,39	0,24	0,42
3 + 5			0,46	0,54
4 + 5	0,83	0,93		0,60
6 + 4	1,00	0,70	0,47	1,00
6 + 5	0,50	0,61		
Minimum	0,00	0,39	0,24	0,42
25 %-Perz.	0,44	0,60	0,46	0,54
<b>Median</b>	<b>0,83</b>	<b>0,65</b>	<b>0,47</b>	<b>0,56</b>
75 %-Perz.	1,00	0,73	0,76	0,65
Maximum	1,00	0,93	1,00	1,00
IQ-Abstand	0,66	0,13	0,30	0,11
Mittelwert	0,66	0,66	0,56	0,62
s	0,37	0,16	0,25	0,18

**Tabelle 5.2.:**  $\kappa$ -Werte: Computertomographie (CT) ( $s = \text{Standardabweichung}$ )(a) CT: kleine runde Fleckschatten vorhanden/  
nicht vorhanden

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	-0,04	1,00		
1 + 6	-0,03	1,00	0,00	
1 + 7			0,00	1,00
2 + 4	1,00	0,00	1,00	1,00
3 + 2			0,00	0,00
3 + 4	0,08	0,06	-0,05	-0,05
3 + 5			0,32	0,30
4 + 5	0,47	0,85		0,00
6 + 4	-0,05	0,18	0,12	0,09
6 + 5	0,09	0,09		
Minimum	-0,05	0,00	-0,05	-0,05
25 %-Perz.	-0,04	0,06	0,00	0,00
<b>Median</b>	<b>0,08</b>	<b>0,18</b>	<b>0,00</b>	<b>0,09</b>
75 %-Perz.	0,47	1,00	0,32	1,00
Maximum	1,00	1,00	1,00	1,00
IQ-Abstand	0,51	0,94	0,32	1,00
Mittelwert	0,22	0,45	0,20	0,33
s	0,39	0,47	0,37	0,47

(b) CT: Streuung kleiner runder Fleckschatten

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	-0,03	1,00		
1 + 6	-0,02	1,00	0,00	
1 + 7			0,00	1,00
2 + 4	1,00	1,00	1,00	1,00
3 + 2			0,00	0,00
3 + 4	0,12	0,24	-0,05	-0,05
3 + 5			0,27	0,31
4 + 5	0,54	0,85		0,00
6 + 4	-0,04	0,68	0,15	0,10
6 + 5	0,05	0,40		
Minimum	-0,04	0,24	-0,05	-0,05
25 %-Perz.	-0,03	0,40	0,00	0,00
<b>Median</b>	<b>0,05</b>	<b>0,85</b>	<b>0,00</b>	<b>0,10</b>
75 %-Perz.	0,54	1,00	0,27	1,00
Maximum	1,00	1,00	1,00	1,00
IQ-Abstand	0,57	0,60	1,00	1,00
Mittelwert	0,23	0,74	0,20	0,34
s	0,40	0,31	0,37	0,47

(c) CT: kleine unregelmäßige Fleckschatten vor-  
handen/nicht vorhanden

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,12	0,46		
1 + 6	0,10	0,10	0,07	
1 + 7			0,30	0,14
2 + 4	0,85	0,54	1,00	0,07
3 + 2			0,10	0,03
3 + 4	0,49	0,43	0,29	-0,02
3 + 5			0,07	-0,03
4 + 5	0,28	0,33		0,00
6 + 4	0,40	0,20	0,44	0,15
6 + 5	0,06	0,09		
Minimum	0,06	0,09	0,07	-0,03
25 %-Perz.	0,10	0,10	0,07	-0,02
<b>Median</b>	<b>0,28</b>	<b>0,33</b>	<b>0,29</b>	<b>0,03</b>
75 %-Perz.	0,49	0,46	0,44	0,15
Maximum	0,85	0,54	1,00	0,15
IQ-Abstand	0,39	0,36	0,37	0,17
Mittelwert	0,33	0,31	0,32	0,05
s	0,28	0,18	0,33	0,07

(d) CT: Streuung kleiner unregelmäßiger Fleck-  
schatten

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,10	0,24		
1 + 6	0,08	0,14	0,07	
1 + 7			0,32	0,17
2 + 4	0,80	0,70	1,00	0,33
3 + 2			0,10	0,01
3 + 4	0,50	0,45	0,39	0,26
3 + 5			0,09	0,01
4 + 5	0,35	0,35		0,00
6 + 4	0,05	0,14	0,41	0,11
6 + 5	0,05	0,07		
Minimum	0,05	0,07	0,07	0,00
25 %-Perz.	0,05	0,14	0,09	0,01
<b>Median</b>	<b>0,10</b>	<b>0,24</b>	<b>0,32</b>	<b>0,11</b>
75 %-Perz.	0,50	0,45	0,41	0,26
Maximum	0,80	0,70	1,00	0,33
IQ-Abstand	0,45	0,31	0,32	0,25
Mittelwert	0,28	0,30	0,34	0,13
s	0,29	0,22	0,33	0,13

(e) CT: lineare Fleckschatten vorhanden/nicht vorhanden

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,48	0,00		
1 + 6	0,04	0,00	0,02	
1 + 7			0,00	0,00
2 + 4	0,77	0,75	0,56	1,00
3 + 2			0,08	0,03
3 + 4	0,58	0,42	0,26	0,06
3 + 5			0,44	0,37
4 + 5	0,43	0,36		-0,33
6 + 4	0,05	0,36	0,49	0,11
6 + 5	0,26	0,09		
Minimum	0,04	0,00	0,00	-0,33
25 %-Perz.	0,05	0,00	0,02	0,00
<b>Median</b>	<b>0,43</b>	<b>0,36</b>	<b>0,26</b>	<b>0,06</b>
75 %-Perz.	0,58	0,42	0,49	0,37
Maximum	0,77	0,75	0,56	1,00
IQ-Abstand	0,53	0,42	0,47	0,37
Mittelwert	0,37	0,28	0,26	0,18
s	0,27	0,27	0,24	0,47

(f) CT: Streuung linearer Fleckschatten

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,38	0,00		
1 + 6	0,05	0,00	0,03	
1 + 7			0,00	0,00
2 + 4	0,62	0,70	0,74	0,82
3 + 2			0,40	0,02
3 + 4	0,57	0,47	0,36	0,06
3 + 5			0,34	0,30
4 + 5	0,52	0,45		-0,18
6 + 4	0,09	0,23	0,43	0,12
6 + 5	0,22	0,12		
Minimum	0,05	0,00	0,00	-0,18
25 %-Perz.	0,09	0,00	0,03	0,00
<b>Median</b>	<b>0,38</b>	<b>0,23</b>	<b>0,36</b>	<b>0,06</b>
75 %-Perz.	0,57	0,47	0,43	0,30
Maximum	0,62	0,70	0,74	0,82
IQ-Abstand	0,48	0,47	0,40	0,30
Mittelwert	0,35	0,28	0,33	0,16
s	0,23	0,27	0,25	0,32

(g) CT: Emphysem vorhanden/nicht vorhanden

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,35	0,48		
1 + 6	0,16	-0,03	0,45	
1 + 7			0,36	0,88
2 + 4	-0,07	1,00	0,63	0,39
3 + 2			0,45	0,55
3 + 4	0,44	0,31	0,19	0,33
3 + 5			0,47	0,47
4 + 5	0,70	0,73		1,00
6 + 4	0,42	0,63	0,38	0,23
6 + 5	0,67	0,36		
Minimum	-0,07	-0,03	0,19	0,23
25 %-Perz.	0,16	0,31	0,36	0,33
<b>Median</b>	<b>0,42</b>	<b>0,48</b>	<b>0,45</b>	<b>0,47</b>
75 %-Perz.	0,67	0,73	0,47	0,88
Maximum	0,70	1,00	0,63	1,00
IQ-Abstand	0,51	0,42	0,11	0,55
Mittelwert	0,38	0,50	0,42	0,55
s	0,27	0,33	0,13	0,29

(h) CT: Streuung Emphyseme

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,38	0,38		
1 + 6	0,06	-0,03	0,48	
1 + 7			0,47	0,76
2 + 4	-0,04	1,00	0,78	0,49
3 + 2			0,51	0,58
3 + 4	0,48	0,36	0,22	0,29
3 + 5			0,52	0,42
4 + 5	0,66	0,75		0,64
6 + 4	0,40	0,66	0,39	0,24
6 + 5	0,54	0,30		
Minimum	-0,04	-0,03	0,22	0,24
25 %-Perz.	0,06	0,30	0,39	0,29
<b>Median</b>	<b>0,40</b>	<b>0,38</b>	<b>0,48</b>	<b>0,49</b>
75 %-Perz.	0,54	0,75	0,52	0,64
Maximum	0,66	1,00	0,78	0,76
IQ-Abstand	0,48	0,45	0,13	0,35
Mittelwert	0,35	0,49	0,48	0,49
s	0,25	0,34	0,17	0,19

(i) CT: Honeycombing vorhanden/nicht vorhanden

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,00	0,00		
1 + 6	0,00	1,00	1,00	
1 + 7			0,00	1,00
2 + 4	0,00	0,00	1,00	1,00
3 + 2			0,00	0,00
3 + 4	0,58	0,49	0,48	0,56
3 + 5			0,38	0,74
4 + 5	0,32	0,32		1,00
6 + 4	0,42	0,00	0,00	0,00
6 + 5	0,00	0,00		
Minimum	0,00	0,00	0,00	0,00
25 %-Perz.	0,00	0,00	0,00	0,00
<b>Median</b>	<b>0,00</b>	<b>0,00</b>	<b>0,38</b>	<b>0,74</b>
75 %-Perz.	0,42	0,49	1,00	1,00
Maximum	0,70	1,00	1,00	1,00
IQ-Abstand	0,42	0,49	1,00	1,00
Mittelwert	0,19	0,26	0,41	0,61
s	0,25	0,38	0,45	0,45

(j) CT: Streuung Honeycombing

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,00	0,00		
1 + 6	0,00	1,00	0,00	
1 + 7			0,00	1,00
2 + 4	0,00	0,00	1,00	1,00
3 + 2			0,00	0,00
3 + 4	0,66	0,58	0,30	0,58
3 + 5			0,38	0,53
4 + 5	0,21	0,19		1,00
6 + 4	0,22	0,00	1,00	0,00
6 + 5	0,00	0,00		
Minimum	0,00	0,00	0,00	0,00
25 %-Perz.	0,00	0,00	0,00	0,00
<b>Median</b>	<b>0,00</b>	<b>0,00</b>	<b>0,00</b>	<b>0,58</b>
75 %-Perz.	0,22	0,58	0,38	1,00
Maximum	0,66	1,00	1,00	1,00
IQ-Abstand	0,22	0,58	0,38	1,00
Mittelwert	0,16	0,25	0,24	0,59
s	0,24	0,39	0,45	0,45

(k) CT: gleichförmige Pleuraverdickung: Dicke

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,00	0,00		
1 + 6	-0,04	0,00	0,04	
1 + 7			-0,03	-0,02
2 + 4	0,75	0,24	0,63	0,00
3 + 2			0,00	0,06
3 + 4	0,30	0,17	0,30	0,24
3 + 5			0,12	0,07
4 + 5	0,23	0,17		0,53
6 + 4	0,09	0,23	0,08	0,05
6 + 5	-0,10	0,14		
Minimum	-0,10	0,00	-0,03	-0,02
25 %-Perz.	-0,04	0,00	0,00	0,00
<b>Median</b>	<b>0,09</b>	<b>0,17</b>	<b>0,08</b>	<b>0,06</b>
75 %-Perz.	0,30	0,23	0,30	0,24
Maximum	0,75	0,24	0,63	0,53
IQ-Abstand	0,34	0,23	0,30	0,24
Mittelwert	0,18	0,14	0,16	0,13
s	0,29	0,10	0,23	0,19

(l) CT: gleichförmige Pleuraverdickung: Ausdehnung

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,00	0,00		
1 + 6	-0,04	0,00	0,05	
1 + 7			-0,02	-0,02
2 + 4	0,58	0,20	0,78	0,00
3 + 2			0,00	0,04
3 + 4	0,27	0,11	0,28	0,14
3 + 5			0,10	0,07
4 + 5	0,14	0,17		0,43
6 + 4	0,03	0,20	0,07	0,04
6 + 5	0,02	0,19		
Minimum	-0,04	0,00	-0,02	-0,02
25 %-Perz.	0,00	0,00	0,00	0,00
<b>Median</b>	<b>0,03</b>	<b>0,17</b>	<b>0,07</b>	<b>0,04</b>
75 %-Perz.	0,27	0,20	0,28	0,14
Maximum	0,58	0,20	0,78	0,43
IQ-Abstand	0,27	0,20	0,28	0,14
Mittelwert	0,14	0,12	0,18	0,10
s	0,22	0,09	0,28	0,15



(m) CT: umschriebene Pleuraverdickungen: Di- (n) CT: umschriebene Pleuraverdickungen: Ausdehnung

Beurteilerpaar	$\kappa$ -Wert				Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4		Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,38	0,32			1 + 2	0,29	0,27		
1 + 6	0,47	0,45	0,38		1 + 6	0,51	0,42	0,42	
1 + 7			0,59	0,43	1 + 7			0,54	0,47
2 + 4	0,72	0,63	0,42	0,42	2 + 4	0,68	0,56	0,63	0,55
3 + 2			0,41	0,25	3 + 2			0,40	0,53
3 + 4	0,64	0,55	0,43	0,38	3 + 4	0,63	0,52	0,33	0,40
3 + 5			0,55	0,58	3 + 5			0,42	0,53
4 + 5	0,58	0,65		0,61	4 + 5	0,61	0,64		0,41
6 + 4	0,31	0,56	0,45	0,33	6 + 4	0,38	0,55	0,61	0,56
6 + 5	0,05	0,34			6 + 5	0,26	0,30		
Minimum	0,05	0,32	0,38	0,25	Minimum	0,26	0,27	0,33	0,40
25 %-Perz.	0,31	0,34	0,41	0,33	25 %-Perz.	0,29	0,30	0,40	0,41
<b>Median</b>	<b>0,47</b>	<b>0,55</b>	<b>0,43</b>	<b>0,42</b>	<b>Median</b>	<b>0,51</b>	<b>0,52</b>	<b>0,42</b>	<b>0,53</b>
75 %-Perz.	0,64	0,63	0,55	0,58	75 %-Perz.	0,63	0,56	0,61	0,55
Maximum	0,72	0,65	0,59	0,61	Maximum	0,68	0,64	0,63	0,56
IQ-Abstand	0,33	0,29	0,14	0,25	IQ-Abstand	0,34	0,26	0,21	0,14
Mittelwert	0,45	0,50	0,46	0,43	Mittelwert	0,48	0,47	0,48	0,49
s	0,23	0,13	0,08	0,13	s	0,17	0,14	0,11	0,07

(o) CT: Kalk vorhanden/ nicht vorhanden

Beurteilerpaar	$\kappa$ -Wert			
	Jahr 1	Jahr 2	Jahr 3	Jahr 4
1 + 2	0,38	0,54		
1 + 6	0,64	0,49	0,52	
1 + 7			0,78	0,74
2 + 4	0,88	0,74	1,00	1,00
3 + 2			0,67	0,43
3 + 4	0,73	0,75	0,64	0,54
3 + 5			0,50	0,70
4 + 5	0,84	0,82		1,00
6 + 4	0,37	0,36	0,57	0,49
6 + 5	0,00	0,71		
Minimum	0,00	0,36	0,50	0,43
25 %-Perz.	0,37	0,49	0,52	0,49
<b>Median</b>	<b>0,64</b>	<b>0,71</b>	<b>0,64</b>	<b>0,70</b>
75 %-Perz.	0,84	0,75	0,78	1,00
Maximum	0,88	0,82	1,00	1,00
IQ-Abstand	0,47	0,26	0,26	0,51
Mittelwert	0,55	0,63	0,67	0,70
s	0,32	0,17	0,17	0,23

**Tabelle 5.3.:** Durchschnittliches  $\kappa$  der einzelnen Paare; ermittelt als Mittelwert sämtlicher  $\kappa$ -Werte dieses Paares

Beurteilerpaar	1 + 2	1 + 6	1 + 7	2 + 4	3 + 2	3 + 4	3 + 5	4 + 5	6 + 4	6 + 5
$\kappa$ -MW	0,20	0,24	0,34	0,68	0,21	0,32	0,31	0,51	0,28	0,17
Standardabweichung	0,27	0,30	0,36	0,33	0,24	0,22	0,20	0,32	0,24	0,20

## 5.2. Ergebnisse: Betrachtung der einzelnen Hypothesen

### 5.2.1. Hypothese 1: Interreadervarianz beim Röntgen größer als bei CT?

1) Bei der Interpretation von Röntgenbildern wird die Interreadervarianz größer sein als bei der Interpretation von CT-Aufnahmen. CT-Aufnahmen sind leichter zu lesen und führen deshalb zu übereinstimmenderen Aussagen, obwohl bei der Interpretation von Röntgenbildern bezüglich Pneumokoniosen viel mehr Übung besteht.

Zur Überprüfung dieser Hypothese wird für jede Fragestellung, die auf beiden Bögen beurteilt werden musste, in jedem Jahr einzeln verglichen, welches der beiden Verfahren die höheren  $\kappa$ -Werte erreicht. Zwei  $\kappa$ -Werte werden nach eigener Definition als gleich angesehen, wenn sie sich um weniger als 0,05 unterscheiden. Tabelle 5.4 stellt dar, welches Verfahren in welchem Jahr einen höheren  $\kappa$ -Wert erlangt.

a) Ermittelt man für jede Zeile der Tab. 5.4, welches Verfahren häufiger genannt wird, so erfährt man für alle neun Fragestellungen, welches Verfahren über die vier Studienjahre hinweg öfter zu besseren Ergebnissen führte. Dies ist in der rechten Spalte der Tabelle aufgeführt. Die Anzahl der Fragestellungen, bei denen ein Verfahren überlegen war, sind in den Abbildungen 5.1 (Median) und 5.2 (Mittelwert) dargestellt. Die Computertomographie erzielte sowohl bezüglich des Mittelwertes als auch des Medians durchschnittlich bessere Ergebnisse.

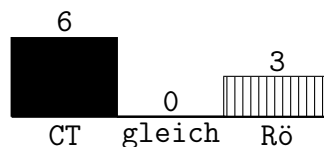


Abbildung 5.1.: Median: Anzahl der Fragestellungen, bei denen das genannte Verfahren höhere Übereinstimmungen erlangt

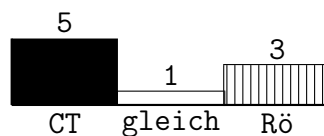


Abbildung 5.2.: MW: Anzahl der Fragestellungen, bei denen das genannte Verfahren höhere Übereinstimmungen erlangt.

Legende zu Abb. 5.1 und Abb. 5.2:

- CT besser
- beide gleich
- Röntgen besser

**Tabelle 5.4.:** Direkter Vergleich der Verfahren für eine Fragestellung für jedes einzelne Jahr; bei jeder Fragestellung wird in jedem Jahr festgestellt, welches Verfahren den höheren Median und Mittelwert hat (das Verfahren mit höherem Wert wird aufgeführt, „-“: gleiches  $\kappa$  bei beiden Verfahren); in Klammern: Anzahl der Fragestellungen, bei denen das Verfahren höhere  $\kappa$ -Werte erlangt

Frage- Stellung	Med/ MW	Jahr 1	Jahr 2	Jahr 3	Jahr 4	bei dieser Frage wurde häufiger ein höherer $\kappa$ - Wert erlangt von:
kl. runde Fleckschatten vorhanden?	Med	-	CT	-	CT	CT
	MW	Rö	CT	CT	CT	CT
kl. runde Fleckschatten Streuung	Med	-	CT	-	CT	CT
	MW	Rö	CT	CT	Rö	-
kl. unregelm. Fleckschatten vorhanden?	Med	CT	CT	CT	Rö	CT
	MW	CT	CT	CT	Rö	CT
kl. unregelm. Fleckschatten Streuung	Med	Rö	Rö	CT	Rö	Rö
	MW	-	Rö	-	Rö	Rö
diffuse Pleuraverdickung Dicke	Med	Rö	-	Rö	Rö	Rö
	MW	Rö	-	-	Rö	Rö
diffuse Pleuraverdickung Ausdehnung	Med	Rö	CT	Rö	Rö	Rö
	MW	Rö	Rö	Rö	Rö	Rö
umschr. Pleuraverdickung Dicke	Med	-	CT	CT	-	CT
	MW	-	-	CT	-	CT
umschr. Pleuraverdickung Ausdehnung	Med	-	-	CT	CT	CT
	MW	-	Rö	CT	CT	CT
Kalk vorhanden?	Med	Rö	CT	CT	CT	CT
	MW	Rö	-	CT	CT	CT
in diesem Jahr überlegen	Med	Rö(4); CT(1)	CT(6); Rö(1)	CT(5); Rö(2)	CT(4); Rö(4)	
	MW	Rö(5); CT(1)	Rö(3); CT(3)	CT(6); Rö(1)	Rö(5); CT(3)	

Bei folgenden Fragestellungen erlangte die Computertomographie deutlich bessere Übereinstimmungen (Nennung in der rechten Spalte sowohl für MW als auch für Median):

- Vorhandensein kleiner runder Fleckschatten,
- Vorhandensein kleiner unregelmäßiger Fleckschatten
- Ausdehnung und Dicke umschriebener Pleuraveränderungen und
- Pleuraverkalkungen

und bessere Übereinstimmungen (Median höher; MW ausgeglichen) bei:

- Streuung kleiner runder Fleckschatten.

Dagegen erlangte das Röntgenverfahren höhere Übereinstimmungen bei:

- Streuung kleiner unregelmäßiger Fleckschatten und
- Ausdehnung und Dicke diffuser Pleuraverdickungen.

b) Vergleicht man die Anzahl der Nennungen der Verfahren in den Spalten der Tabelle 5.4, so kann man ablesen, welches Verfahren im betrachteten Jahr eine höhere Übereinstimmung hatte. Dabei ist eine zeitliche Entwicklung festzustellen: mit der Zeit nimmt die Häufigkeit der Nennung des Röntgenverfahrens ab und des CT-Verfahrens zu.

a) bestätigt die Hypothese, dass die Interreadervarianz bei der Interpretation von Röntgenbildern höher ist.

b) zeigt, dass der Lerneffekt bei der Interpretation von CT-Bildern größer ist als der Lerneffekt bei der Interpretation von Röntgenbildern.

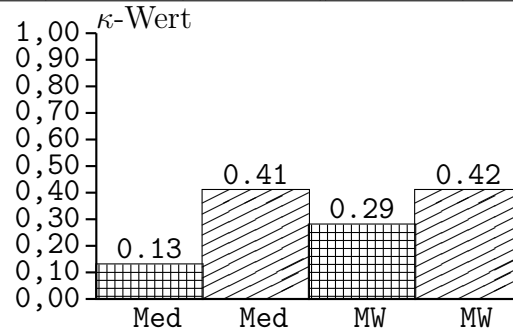
### 5.2.2. Hypothese 2: Rö: Interreadervarianz bei der Beurteilung des Parenchyms größer als bei der Pleurabeurteilung?

2) *Bei Röntgen-Aufnahmen wird die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen größer sein als bei pleuralen Veränderungen. Allgemein lassen sich pneumokoniotische Veränderungen besser im Pleurabereich erkennen als im Parenchym.*

Tab. 5.5 stellt die durchschnittlich erreichten  $\kappa$ -Werte der Pleura- und der Parenchymfragestellungen bei der Interpretation von Röntgenbildern für jedes Jahr einzeln dar. Es wurden die mittleren Mittelwerte und die mittleren Mediane über alle Parenchym- (kleine runde/kleine unregelmäßige Fleckschatten, jeweils vorhanden/nicht vorhanden und Streuung; also vier Fragestellungen) und alle Pleurafragestellungen (diffuse/umschriebene Pleuraverdickung, jeweils Dicke und Ausdehnung und Kalk vorhanden/nicht vorhanden; also fünf Fragestellungen) eines Jahres ermittelt und in Tab. 5.5 dargestellt. Anschließend wurden zur Überprüfung der zweiten Hypothese die Mittelwerte der Mittelwertspalten und die Mittelwerte der Medianspalten der Tab. 5.5 ermittelt, um die mittleren Mediane bzw. Mittelwerte über den gesamten Zeitraum zu erhalten. Die Ergebnisse sind in Abb. 5.3 dargestellt.

**Tabelle 5.5.:** R $\ddot{o}$ : Die Mediane und Mittelwerte der Parenchym- und Pleurafragestellungen der einzelnen Jahre (Med: mittlerer Median, MW: Mittelwert, in Klammern: die Standardabweichungen)


Jahr	Parenchym		Pleura	
	Med	MW	Med	MW
1	0,11 (0,06)	0,30 (0,03)	0,51 (0,19)	0,45 (0,15)
2	0,18 (0,25)	0,25 (0,15)	0,37 (0,26)	0,43 (0,22)
3	0,11 (0,13)	0,21 (0,08)	0,31 (0,11)	0,33 (0,14)
4	0,13 (0,12)	0,39 (0,05)	0,43 (0,10)	0,46 (0,11)

**Abbildung 5.3.:** R $\ddot{o}$ :  $\kappa$  für Parenchym- und Pleurafragestellungen (ermittelt für alle Fragestellungen im gesamten Zeitraum) Standardabweichungen: mittl. Median: Parenchym 0,03, Pleura 0,09; MW Parenchym 0,08, Pleura 0,16.

Legende:

Med: mittlerer Median

MW: mittlerer Mittelwert

 Parenchym

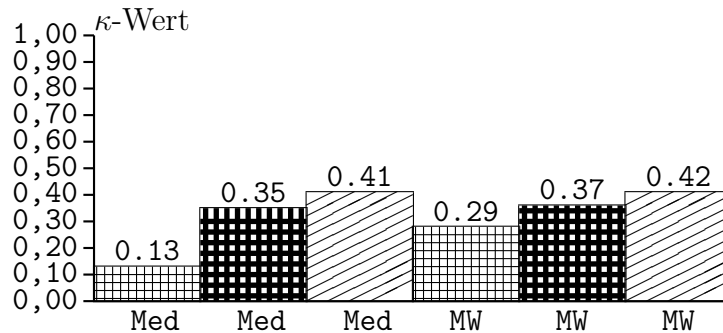
 Pleura

Es zeigt sich, dass die Übereinstimmung bei der Beurteilung von pleuralen Veränderungen deutlich besser war als bei der Beurteilung von parenchymalen Veränderungen. Der Pleuramedian war um den Faktor 3,2, der Pleuramittelwert um den Faktor 1,5 höher als der Parenchymmedian bzw. -mittelwert. Damit ist die zweite Hypothese belegt.

In Tabelle 5.6 und Abbildung 5.4 werden differenziertere Betrachtungen angestellt, die Ergebnisse werden im Diskussionsteil diskutiert.

**Tabelle 5.6.:** R $\ddot{o}$ : Die Mediane und Mittelwerte der Parenchym- und Pleurafragestellungen (+/- Kalk) der einzelnen Jahre (Med: mittlerer Median, MW: mittlerer Mittelwert, in Klammern: die Standardabweichungen)

Jahr	Parenchym		Pleura			
	Med	MW	ohne Kalk		mit Kalk	
	Med	MW	Med	MW	Med	MW
1	0,11 (0,06)	0,30 (0,03)	0,43 (0,09)	0,40 (0,11)	0,51 (0,19)	0,45 (0,15)
2	0,18 (0,25)	0,25 (0,15)	0,30 (0,24)	0,37 (0,20)	0,37 (0,26)	0,43 (0,22)
3	0,11 (0,13)	0,21 (0,08)	0,27 (0,07)	0,28 (0,06)	0,31 (0,11)	0,33 (0,14)
4	0,13 (0,12)	0,39 (0,05)	0,40 (0,07)	0,42 (0,06)	0,43 (0,10)	0,46 (0,11)





**Abbildung 5.4.:** R $\ddot{o}$ :  $\kappa$  f $\ddot{u}$ r Parenchym- und Pleurafragestellungen (ermittelt f $\ddot{u}$ r alle Fragestellungen im gesamten Zeitraum) Standardabweichungen: mittl. Median: Parenchym 0,03, Pleura ohne Verkalkungen 0,08, Pleura mit Verkalkungen 0,09; MW Parenchym 0,08, Pleura ohne Verkalkungen 0,06, Pleura mit Verkalkungen 0,16.


Legende:

Med: mittlerer Median

MW: mittlerer Mittelwert

 Parenchym

 Pleura ohne Verkalkungen

 Pleura mit Verkalkungen

### 5.2.3. Hypothese 3: CT: Interreadervarianz bei der Beurteilung des Parenchyms gr $\ddot{o}$ ber als bei der Pleurabeurteilung?

3) Bei CT-Aufnahmen wird die Interreadervarianz bei der Beurteilung von parenchymalen Ver $\ddot{a}$ nderungen gr $\ddot{o}$ ber sein als bei pleuralen Ver $\ddot{a}$ nderungen. Allgemein lassen sich pneumokoniotische Ver $\ddot{a}$ nderungen besser im Pleurabereich erkennen als im Parenchym.

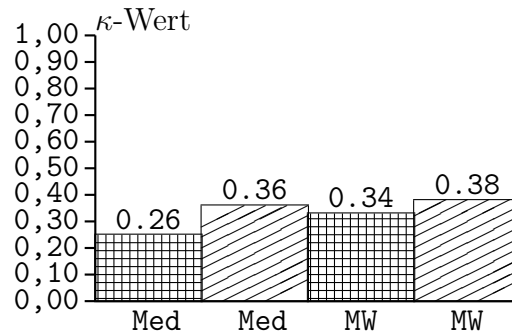
Hier wurde f $\ddot{u}$ r die  $\kappa$ -Werte des CT-Verfahrens genau so vorgegangen wie bei der Beurteilung der zweiten Hypothese: es wurden die Mediane und Mittelwerte aller Parenchym- und Pleurafragestellungen ermittelt.

**Tabelle 5.7.:** CT: Mediane und Mittelwerte der Parenchym- und Pleurafragestellungen der einzelnen Jahre (Med: mittl. Median, MW: Mittelwert)

Jahr	Parenchym		Pleura	
	Med	MW	Med	MW
1	0,21 (0,18)	0,29 (0,08)	0,35 (0,27)	0,36 (0,19)
2	0,31 (0,25)	0,39 (0,16)	0,42 (0,24)	0,37 (0,22)
3	0,25 (0,19)	0,32 (0,10)	0,33 (0,25)	0,39 (0,22)
4	0,27 (0,27)	0,34 (0,21)	0,35 (0,29)	0,37 (0,25)

Es l $\ddot{a}$ sst sich eine bessere  $\ddot{U}$ bereinstimmung bei der Beurteilung der Pleura als bei der Beurteilung des Parenchyms nachweisen. Der Pleuramedian war um den Faktor 1,4, der Pleuramittelwert um den Faktor 1,1 h $\ddot{o}$ her als der Parenchymmedian bzw. -mittelwert. Damit ist die dritte Hypothese best $\ddot{a}$ tigt, wenn auch nicht ganz so eindr $\ddot{u}$ cklich wie beim R $\ddot{o}$ ntgen.

Auch hier wurden differenziertere Betrachtungen angestellt (Tab. 5.8 und Abb. 5.6), die im Diskussionsteil diskutiert werden.




**Abbildung 5.5.:** CT:  $\kappa$  für Parenchym- und Pleurafragestellungen (ermittelt für alle Fragestellungen im gesamten Zeitraum) Standardabweichungen: mittl. Median: Parenchym 0,04, Pleura 0,04; MW Parenchym 0,04, Pleura 0,01

Legende:

Med: mittlerer Median

MW: mittlerer Mittelwert

 Parenchym

 Pleura

**Tabelle 5.8.:** CT: Die Mediane und Mittelwerte der Parenchym- und Pleurafragestellungen (+/- Kalk) der einzelnen Jahre (Med: mittlerer Median, MW: mittlerer Mittelwert; in Klammern: Standardabweichung; CT(Rö): kleine runde und kleine unregelmäßige Fleckschatten, jeweils vorhanden/nicht vorhanden und Streuung; CT komplett: alle CT-Parenchymfragen inkl. lineare Fleckschatten, Emphysem, Honeycombing)

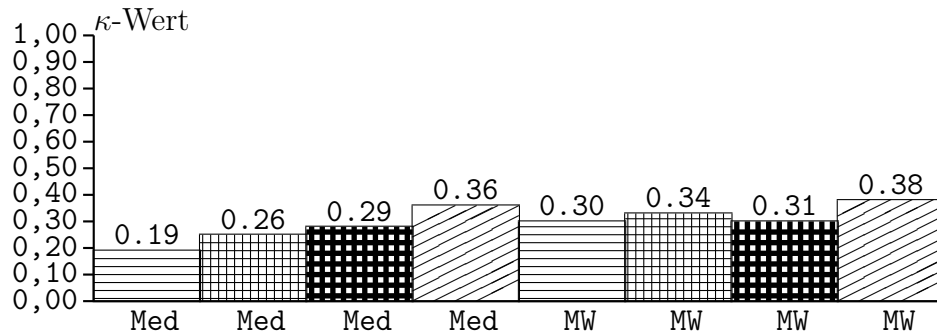
Jahr	Parenchym				Pleura			
	CT(Rö)		CT komplett		ohne Kalk		mit Kalk	
	Med	MW	Med	MW	Med	MW	Med	MW
1	0,13 (0,10)	0,27 (0,05)	0,21 (0,18)	0,29 (0,08)	0,28 (0,25)	0,31 (0,18)	0,35 (0,27)	0,36 (0,19)
2	0,40 (0,31)	0,45 (0,21)	0,31 (0,25)	0,39 (0,16)	0,35 (0,21)	0,31 (0,20)	0,42 (0,24)	0,37 (0,22)
3	0,15 (0,18)	0,27 (0,08)	0,25 (0,19)	0,32 (0,10)	0,25 (0,20)	0,32 (0,17)	0,33 (0,25)	0,39 (0,22)
4	0,08 (0,04)	0,21 (0,15)	0,27 (0,27)	0,34 (0,21)	0,26 (0,25)	0,29 (0,20)	0,35 (0,29)	0,37 (0,25)

#### 5.2.4. Hypothese 4: Rö: sinkende Interreadervarianz im zeitlichen Verlauf?

4) Im zeitlichen Verlauf wird die Interreadervarianz bei der Beurteilung von Röntgenbildern sinken, da hier ein Lernprozess stattfinden wird.

In den Abbildungen 5.7 und 5.8 sind die Median- bzw. Mittelwertverläufe aller neun Fragestellungen aufgeführt. Bezüglich des Röntgen-Medians findet sich ein uneinheitliches Bild bis auf das vierte Jahr, wo sich acht Verbesserungen und nur eine Verschlechterung findet. Beim Röntgen-Mittelwert überwiegen im 2. und 4. Jahr die Verbesserungen, wobei es im 4. Jahr lediglich Verbesserungen gibt.

Diese Betrachtung ist jedoch nur eingeschränkt aussagekräftig, da die Paare gewechselt haben und dadurch kein konstantes durchschnittliches Lernverhalten über die Jahre erwartet werden konnte. Deshalb wurde ein Verfahren entwickelt, um den durchschnittlichen Lernfortschritt der einzelnen Paare zu quantifizieren. Hierfür wurden die Differenzen zwischen den  $\kappa$ -Werten zweier Beurteilungstermine aufsummiert und durch die Anzahl dieser Differenzen geteilt (zum Beispiel gäbe es bei drei Terminen zwei Differenzen, dementsprechend müsste die Summe der zwei  $\kappa$ -Wertdifferenzen durch zwei geteilt





**Abbildung 5.6.:** CT:  $\kappa$  für Parenchym- und Pleurafragestellungen (ermittelt für alle Fragestellungen im gesamten Zeitraum) Standardabweichungen: mittl. Median: Parenchym -runde und irreguläre Fleckschatten: 0,14, Parenchym -alle Fragestellungen: 0,04, Pleura ohne Verkalkungen: 0,05, Pleura mit Verkalkungen: 0,04; MW: Parenchym -runde und irreguläre Fleckschatten: 0,10, Parenchym -alle Fragestellungen: 0,04, Pleura ohne Verkalkungen: 0,01, Pleura mit Verkalkungen: 0,01


Legende:


Med: mittlerer Median

MW: mittlerer Mittelwert

 Parenchym -runde und irreguläre Fleckschatten

 Parenchym -alle Fragestellungen: runde, irreguläre und lineare Fleckschatten, Emphysem, Honeycombing

 Pleura-ohne Verkalkungen

 Pleura-mit Verkalkungen

werden). Beispielhaft wird das Vorgehen anhand der Streuung kleiner runder Fleckschatten vorgeführt (s. Tab 5.9).

**Tabelle 5.9.:**  $\kappa$ -Wert-Differenz bei Rö: Streuung kleiner runder Fleckschatten

Beurteilerpaar	Jahr 1 zu Jahr 2	Jahr 2 zu Jahr 3	Jahr 3 zu Jahr 4
1 + 2	+ 0,03		
1 + 6	- 0,15	0,00	
1 + 7			+1,00
2 + 4	-1,00	+1,00	0,00
3 + 2			-0,02
3 + 4	+0,01	-0,07	0,00
3 + 5			+0,02
4 + 5	0,00		0,00
6 + 4	0,00	-0,02	+0,04
6 + 5	-0,02		

Beim Aufaddieren der Werte aus Tab. 5.9 erhält man einen Wert von 0,82. Dieser muss durch 18 dividiert werden, was 0,0456 ergibt. Dies entspricht einem Lernverhalten von 4,6 %. Die Ergebnisse dieses Vorgehens für jede Fragestellung zeigt Tabelle 5.10.

Einen deutlichen Lerneffekt (4,0-4,6 %) zeigten die Beurteilungen von

- Vorhandensein und
- Streuung kleiner runder Fleckschatten und
- Vorhandensein kleiner unregelmäßiger Fleckschatten.

Einen mittelmäßigen Lerneffekt (1,5-2,7 %) zeigten die Beurteilungen von

- Ausdehnung diffuser Pleuraverdickung und
- Streuung kleiner unregelmäßiger Fleckschatten.



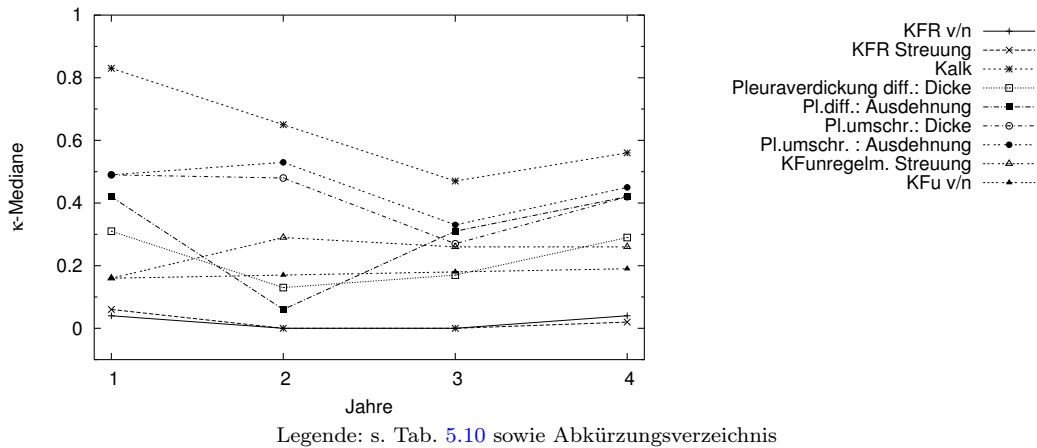


Abbildung 5.7.: Rö: Medianverläufe der einzelnen Fragestellungen

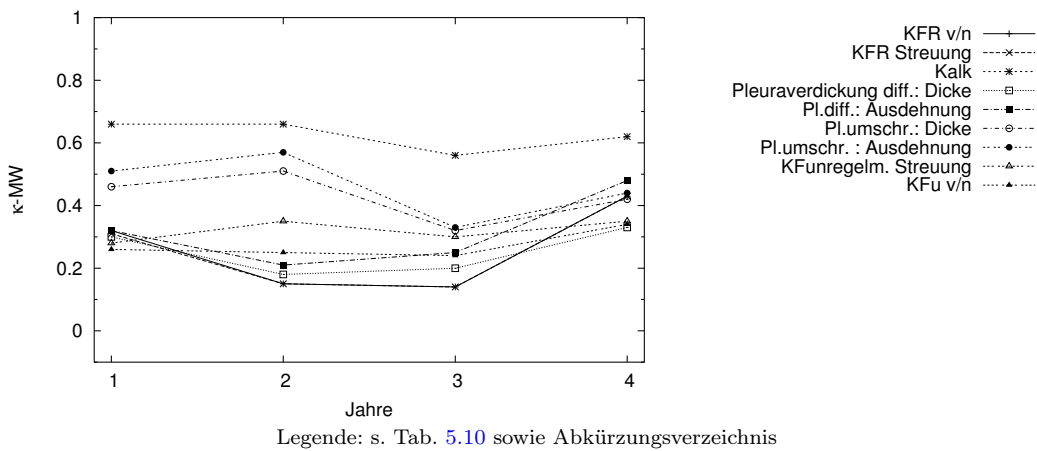


Abbildung 5.8.: Rö: Mittelwertverläufe der einzelnen Fragestellungen

Eine Verschlechterung (-1,5 bis -0,5 %) zeigten die Beurteilungen von

- Vorhandensein von Kalk
- Ausdehnung und
- Dicke umschriebener Pleuraverdickungen und
- Dicke diffuser Pleuraverdickungen.

Im Durchschnitt zeigt sich ein Lernverhalten von 1,4 %, womit die vierte Hypothese belegt werden konnte.

**Tabelle 5.10.: Lerneffekt beim Röntgen**

Fragest.	KFR v/n	KFR S	Kalk v/n	PdD	PdA
Lerneffekt	4,1 %	4,6 %	-1,1 %	-0,9 %	2,7 %
Fragest.	PuD	PuA	KFu S	KFu v/n	∅
Lerneffekt	-1,0 %	-1,5 %	1,5 %	4,0 %	1,4 %

Legende: KFR v/n: kleine runde Fleckschatten vorhanden/nicht vorhanden; KFR S: Streuung kleiner runder Fleckschatten; Kalk v/n: Kalk vorhanden/nicht vorhanden; PdD: Dicke diffuser Pleuraverdickungen; PdA: Ausbreitung diffuser Pleuraverdickungen; PuD: Dicke umschriebener Pleuraverdickungen; PuA: Ausbreitung umschriebener Pleuraverdickungen; KFu S: Streuung kleiner unregelmäßiger Fleckschatten; KFu v/n: kleine unregelmäßige Fleckschatten vorhanden/nicht vorhanden ; ∅: durchschnittlicher Lernfortschritt über alle Fragestellungen.

### 5.2.5. Hypothese 5: CT: sinkende Interreadervarianz im zeitlichen Verlauf?

5) *Im zeitlichen Verlauf wird die Interreadervarianz bei der Beurteilung von CT-Bildern sinken, da ein Lernprozess stattfinden wird.*

Hier wurde wie bei der Überprüfung des Lernverhaltens bei der Röntgen-Interpretation (Hypothese 4) vorgegangen. Die Abb. 5.9, 5.10, 5.11 und 5.12 stellen die MW- bzw. Medianverläufe der Fragestellungen dar. Die Abb. 5.9 und 5.10 zeigen das Lernverhalten bei Fragestellungen, die sowohl auf dem Röntgen(ILO)- als auch auf dem CT-Bogen vorkommen. Die Abb. 5.10 und 5.11 zeigen Fragestellungen, die nur auf dem CT-Bogen vorkommen.

Die Verläufe zeigen keine deutliche Tendenz. Dennoch soll im Folgenden auf besondere Auffälligkeiten aufmerksam gemacht werden. Bei den Medianen der Fragestellungen, die auf beiden Bögen vorkommen (Abb.5.9), fanden sich meist zwischen erstem und zweiten Termin Verbesserungen, danach Verschlechterungen und im letzten Jahr dann wieder einige Verbesserungen. Bei den MW dieser Fragestellungen wirkte das Bild etwas ruhiger (Abb.5.10). Ins Auge fiel eine besonders hohe Übereinstimmung im zweiten Jahr bei der Frage nach der Streuung kleiner rundlicher Fleckschatten und steile Abfälle im vierten Jahr bei den Fragen nach kleinen unregelmäßigen Fleckschatten. Ein deutlicher Lerneffekt zeigte sich bei der Frage nach Verkalkungen. Bei den Fragen, die lediglich auf den CT-Bögen vorkommen (Abb. 5.11 und 5.12) zeigten sich deutliche Lerneffekte bei den Fragen nach Honeycombing. Deutliche Verschlechterungen sah man bei den Fragen nach linearen Fleckschatten.

Die durchschnittlichen Lernfortschritte zwischen zwei Terminen wurden gemäß den Lernfortschritten bei der Interpretation von Röntgenbildern ermittelt und werden in den Tabellen 5.11 und 5.12 dargestellt.

**Tabelle 5.11.: Der Lerneffekt bei der CT in Fragestellungen, die auf beiden Bögen vorkommen. Bezeichnung s. Tab. 5.10**

Fragest.	KFR v/n	KFR S	Kalk v/n	PdD	PdA
Lerneffekt	8,8 %	10,4 %	4,9 %	-1,7 %	-0,2 %
Fragest.	PuD	PuA	KFu v/n	KFu S	∅
Lerneffekt	-3,1 %	-1,4 %	-8,9 %	-6,5 %	0,3 %

Legende: s. Abb. 5.7

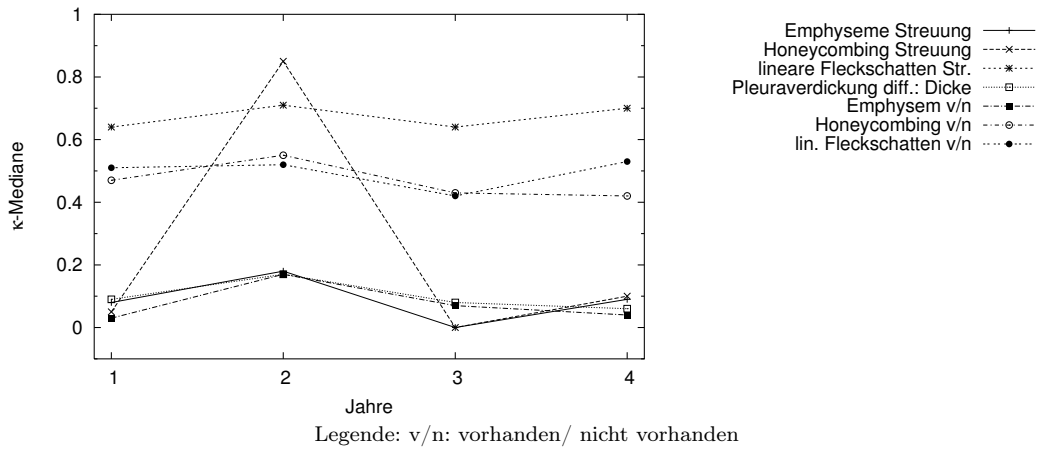


Abbildung 5.9.: CT: Medianverläufe der Fragestellungen, die auf beiden Bögen vorkommen

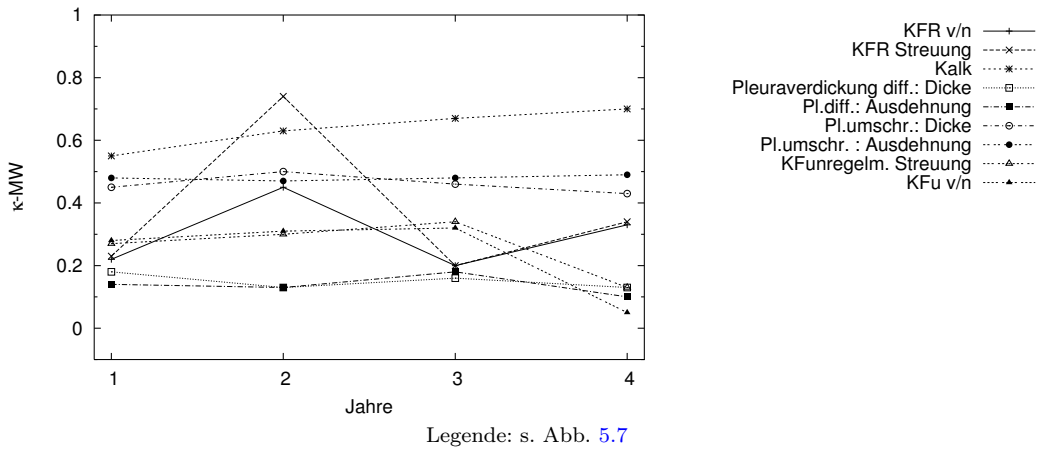


Abbildung 5.10.: CT: Mittelwertverläufe der Fragestellungen, die auf beiden Bögen vorkommen

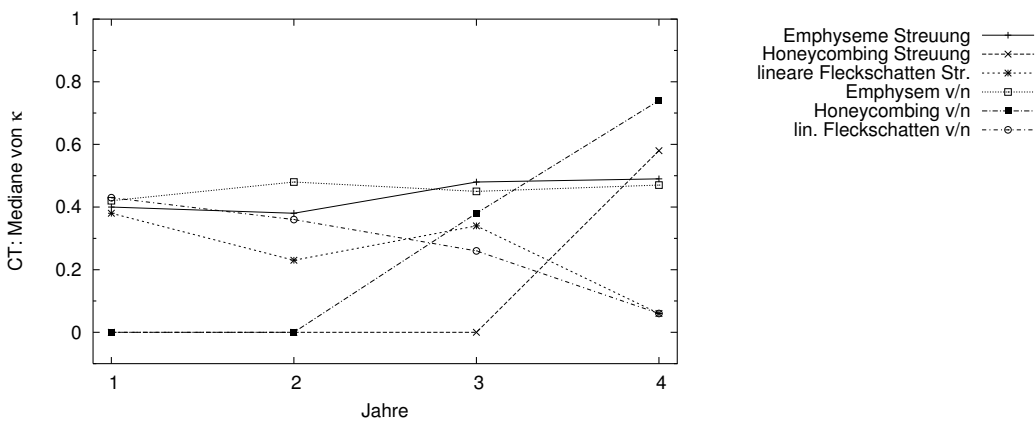


Abbildung 5.11.: CT: Medianverläufe der Fragestellungen, die es nur auf dem CT-Bogen gibt

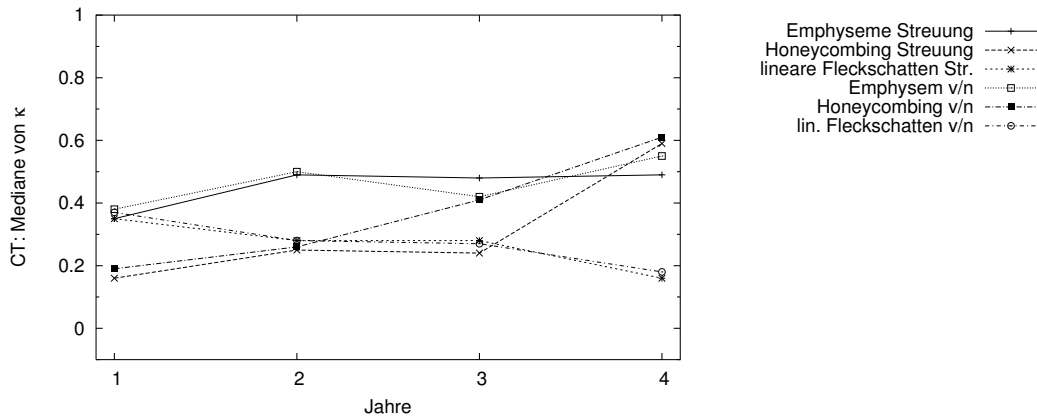


Abbildung 5.12.: CT: MW-Verläufe der Fragestellungen, die es nur auf dem CT-Bogen gibt

Tabelle 5.12.: Der Lerneffekt bei der CT in Fragestellungen, die nur auf dem CT-Bogen vorkommen

Fragest.	Em v/n	Em S	Hon v/n	Hon S	linFS v/n	linFS S	∅
Lerneff	6,6 %	3,3 %	20,0 %	14,6 %	-9,9 %	-10,6 %	4,0 %

scriptsize

Legende: Em S: Streuung von Emphysemen; Em v/n: Emphyseme vorhanden/nicht vorhanden; Hon S: Streuung von Honeycombing; Hon v/n: Honeycombing vorhanden/nicht vorhanden; linFS S: Streuung linearer Fleckschatten; linFS v/n: lineare Fleckschatten vorhanden/nicht vorhanden; ∅: durchschnittlicher Lernfortschritt über alle Fragestellungen

a) Bei den Fragestellungen, die auf beiden Bögen vorkommen, gibt es drei, bei denen dazugelernt wird. Hier ist das Lernverhalten deutlich ausgeprägter als bei Fragestellungen, bei denen bei der Interpretation von Röntgenbildern dazugelernt wird:

- Vorhandensein (8,8 %) und Streuung (10,4 %) kleiner runder Fleckschatten und
- Vorhandensein von Kalk (4,9 %).

Bei drei Fragestellungen ist die Verschlechterung ähnlich wie beim Röntgen:

- Ausdehnung (-0,2 %) und Dicke (-1,7 %) diffuser Pleuraverdickungen und
- Ausdehnung umschriebener Pleuraverdickungen (-1,4 %).

Bei drei Fragestellungen ist die Verschlechterung viel ausgeprägter als beim Röntgen:

- Dicke umschriebener Pleuraverdickungen (-3,1 %) und
- Vorhandensein (-8,9 %) und Streuung (-6,5 %) unregelmäßiger Fleckschatten.

Im Schnitt lässt sich ein Lerneffekt von 0,3 % errechnen. Insgesamt ist hier demnach kein deutlicher Lerneffekt festzustellen.

b) Bei den Fragestellungen, die nur auf dem CT-Bogen vorkommen, ist jedoch ein ausgeprägter Lerneffekt nachweisbar. Es zeigt sich ein sehr gutes Lernverhalten bei der Beurteilung von:

- Vorhandensein von Emphysemen (6,6 %) und
- Vorhandensein (20,0 %) und Streuung (14,6 %) von Honeycombing,

ein gutes Lernverhalten bei:

- Streuung von Emphysemen (3,3 %)

und eine starke Verschlechterung bei:

- Vorhandensein und Streuung linearer Fleckschatten (-9,9 bzw. -10,6 %).

Im Schnitt wurde ein Lerneffekt von 4,0 % erzielt.

a) und b) zeigen, dass insgesamt bei der Interpretation von CT-Bildern ein Lerneffekt nachzuweisen ist, womit die Hypothese belegt ist.

## 5.3. Diskussion und Schlussfolgerungen

Bei der vorliegenden Studie handelt es sich um eine Längsschnittstudie, die im Zeitraum von 1993-1997 in sechs Zentren durchgeführt wurde. Aus einem Hochrisikokollektiv von 3.000 Personen wurden nach festgelegten Ausschlusskriterien 1.917 bisher nicht erkrankte Personen unter 70 Jahren ausgewählt, die sich bislang nicht gegen eine weitergehende Untersuchung ausgesprochen hatten und ein hohes Risiko hatten, an einer asbestinduzierten Erkrankung zu erkranken. Ihnen wurde das Angebot einer ergänzenden arbeitsmedizinischen Vorsorgeuntersuchungsserie gemacht, die über 4 Jahre hinweg jährlich eine gewöhnliche Vorsorgeuntersuchung mit einer zusätzlichen HRCT-Untersuchungen vorsah. Die Studie begann 1993 mit 636 Personen, die im durchschnittlichen Abstand von zwölf Monaten mittels Röntgen-Aufnahmen und HRCT-Aufnahmen untersucht wurden. Die Röntgenbefunde wurden im ILO-1980-Bogen, die HRCT-Befunde auf einem in Anlehnung an den ILO-Bogen entwickelten, hier das erste Mal Anwendung findenden, Bogen festgehalten.

Die Compliance war mit 59,7 % überdurchschnittlich gut. Alle Röntgenbilder und CT-Aufnahmen eines Probanden (jährliche Röntgen- und CT-Aufnahmen) wurden von zwei Beurteilern ausgewertet, wobei dem Zweitauswerter das Ergebnis des Erstauswerter unbekannt war. Die Beurteiler hatten sich durch besonders grossen Erfahrungsreichtum in der Interpretation von Pneumokoniosebildern qualifiziert. Nach zwei Untersuchungszyklen wurden die Zweitbeurteiler gewechselt, um systematische Fehler zu vermeiden. Insgesamt gab es zehn Beurteilerpaare. Einige Patienten wurden aufgrund fehlender Angaben aus dem auswertbaren Kollektiv entfernt (fehlende Angaben zu den Beurteilern, fehlende Bögen, etc.).

Ziel dieser Arbeit ist es, die Übereinstimmung der Beurteilung eines Bildes durch zwei Beurteiler zu ermitteln. Aus den ca. 300 Variablen der Kodierungsbögen mussten Kategorien gebildet werden, die gleichartige Aussagen zusammenfassten. Bisher wurde die Kategorienbildung durch verschiedene Arbeitsgruppen auf sehr unterschiedliche Arten durchgeführt. In der vorliegenden Studie war es wichtig, für beide Verfahren möglichst ähnliche Kategorien zu bilden, um die beiden Verfahren miteinander vergleichen zu können.

Bezüglich Parenchymveränderungen auf Röntgen-Aufnahmen wurden  $\kappa$ -Werte für die kleinen rundlichen Fleckschatten und für die unregelmäßigen Fleckschatten ermittelt, für die es jeweils Angaben über Größe, Streuung und Lokalisation gab. Die Größenangaben wurden auf die dichotome Aussage reduziert, ob die genannte Veränderung vorhanden war oder nicht. Für Streuungsangaben wurden fünf Kategorien gebildet, entsprechend den vier ILO-Hauptkategorien und eine zusätzlichen Kategorie 0/1. Lokalisationsangaben konnten nicht ausgewertet werden.

Bezüglich Pleuraveränderungen wurden die Angaben zu Verbreitung und Dicke umschriebener und diffuser Pleuraverdickungen nicht zusammengefasst. Aussagen zur Pleuraverkalkung wurden entsprechend dem CT-Bogen dichotom (pos/neg) zusammengefasst. Lokalisationsangaben konnten nicht ausgewertet werden.

Die Kategorienbildung des CT-Bogens wurde mit der des ILO-Bogens abgestimmt. Ausgewertet wurden die Übereinstimmungen bei kleinen, rundlichen Fleckschatten, bei

kleinen irregulären Fleckschatten, bei linearen Fleckschatten, bei Honeycombing und bei Emphysemen. Die zugehörigen Größenangaben wurden entsprechend dem Röntgenverfahren zu dichotomen Aussagen (vorhanden/ nicht vorhanden) verdichtet. Die Streuungsangaben wurden nicht weiter zu Kategorien zusammengefasst und entsprachen somit den oben gebildeten Röntgen-Streuungskategorien. Lokalisationsangaben konnten nicht ausgewertet werden.

Sowohl die umschriebenen als auch die gleichförmigen Pleuraverdickungen wurden bezüglich Ausdehnung, Dicke und Verkalkung gemäß dem ILO-Bogen in drei Größenangaben bewertet. Die beiden „Kalk“-Rubriken unter „umschrieben“ und „gleichförmig“ wurden zusammengefasst, um entsprechend dem ILO-Bogen lediglich eine (dichotome) Rubrik „Kalk“ zu erhalten. Lokalisationsangaben konnten nicht ausgewertet werden.

Mit Hilfe des Programms MedCalc wurden die  $\kappa$ -Werte der einzelnen Beurteilerpaare errechnet. Für dichotome Fragestellungen wurde ein ungewichtetes  $\kappa$ , für nicht-dichotome ein gewichtetes  $\kappa$  ermittelt. Für jedes Jahr wurde über alle Beurteilerpaare das Minimum, das Maximum, der Median, der Mittelwert, die 25- und 75 %-Perzentile, der IQ-Abstand und die Standardabweichung (s) für jede Fragestellung errechnet.

Es wurden fünf Hypothesen geprüft. Im folgenden sollen die Ergebnisse diskutiert werden.

### 5.3.1. Diskussion des Studiendesigns

Die Ermittlung des  $\kappa$ -Wertes ist das derzeit anerkannteste Verfahren, Übereinstimmungen in der Form zu ermitteln wie es in der vorliegenden Studie nötig war. Mögliche Fehlerquellen, die zu einer Verzerrung der  $\kappa$ -Werte geführt haben können, sollen im folgenden diskutiert werden.

Da die Fragebögen verschiedene Antwortmöglichkeiten vorsehen, mussten - um die Vergleichbarkeit zu gewährleisten- mehrere Möglichkeiten zu Kategorien zusammengefasst werden. Durch die dadurch eingetretene Verminderung der Antwortmöglichkeiten kam es zu einer Verbesserung der Übereinstimmung und damit der  $\kappa$ -Werte.

Weiterhin ist zu berücksichtigen, dass die einzelnen Paare eine sehr unterschiedliche Anzahl an Probanden untersucht haben. Die Zahlen variieren zwischen 8 und 243. Paar 2 + 4 -mit beachtlicher Übereinstimmung- hat mit durchschnittlich 14 Probanden die geringste Probandenzahl beurteilt. Damit sind die  $\kappa$ -Werte der verschiedenen Paare unterschiedlich verlässlich. Die Verlässlichkeit steigt mit wachsender Größe des Kollektivs. Zudem ist zu überlegen, ob die Anzahl der beurteilten Probanden insofern einen Einfluss auf die  $\kappa$ -Werte hat, als Beurteiler von vielen Probanden vielleicht dazu neigen, die Bilder nicht so aufmerksam zu betrachten wie Beurteiler, die nur wenige Bilder beurteilen müssen.

Die Beurteiler haben sich einmal im Jahr zu einer Konferenz zusammengefunden und über die Definitionen der Veränderungen diskutiert. Theoretisch müsste ein allgemeiner Gruppenlernerneffekt nachzuweisen sein, der sich in einer Verbesserung der Übereinstimmungen der einzelnen Paare niederschlagen sollte. Es ist jedoch nicht auszuschließen, dass zu manchen Definitionen keine Einigung gefunden werden konnte und dementsprechend kein Gruppenlernerneffekt bezüglich dieser Fragestellungen nachzuweisen sein kann.

Desweiteren bleibt offen, wie lernbereit die einzelnen Beurteiler waren. Es ist möglich, dass sehr erfahrene Beurteiler weniger bereit waren, ihre Beurteilungsgewohnheiten zu ändern als unerfahrenere.

Es sollte beachtet werden, dass manche Beurteiler bereits vor Studienbeginn zusammen gearbeitet haben und sich somit in ihrer Befunderhebung vor Studienbeginn stärker aneinander angeglichen hatten, so dass ihre Übereinstimmung dementsprechend höher ist als bei anderen Paaren unabhängig von der Qualität ihrer Befunde.

Es fanden jährlich kleine Zusammensetzungsänderungen im Kollektiv statt, da nicht alle Probanden jedes Jahr zur Untersuchung erschienen. Besonders beachtet werden sollte, dass im dritten Jahr 80 Patienten aufgrund fehlender Zweitbeurteilerbögen gestrichen werden mussten. Dies entspricht einem Sechstel des Gesamtkollektivs des dritten Jahres. Damit sind die  $\kappa$ -Werte des dritten Jahres nicht so verlässlich wie die der anderen Jahre.

Allgemein beeinflusst wurden die Werte möglicherweise auch durch den Untersuchungsort, der die Kollektive bestimmte, gegebenenfalls gab es an verschiedenen Orten unterschiedlich leicht zu interpretierende Kollektive. Generell gilt, dass das gesamte Kollektiv nicht leicht zu interpretieren war, da sich in fast allen Fällen wenn überhaupt nur minimale Veränderungen finden ließen, die besonders schwer gleichartig zu beurteilen sind, weil hier die Definitionen besonders stark selbstbestimmt sind.

Trotz all der Kritikpunkte ist festzuhalten, dass in dieser Studie mit einem aussergewöhnlich großen Patientenkollektiv gearbeitet wurde. Durch das vierjährige Langzeitstudiendesign, wurden letztendlich Daten von etwa 2400 Untersuchungen festgehalten. Diese Datenmenge ist bislang einmalig. Gleichzeitig konnte durch das bisher einmalige Langzeitstudiendesign das Lernverhalten beurteilt werden. Es ist auch die einzige Studie, die Parenchym und Pleura auf Röntgen- und CT-Bildern beurteilt. Dabei wurden die Fragebogenkategorisierungen so aufeinander abgestimmt, dass erstmalig ein vorsichtiger Vergleich der Übereinstimmungen der beiden Verfahren möglich ist (die Verfahren sind aufgrund ihrer sehr unterschiedlichen Natur nur eingeschränkt vergleichbar). Damit hat diese Studie eine herausragende Position unter den bislang durchgeführten Studien zur Interreadervarianz bei der Beurteilung von Pneumokoniosen.

### 5.3.2. Allgemeine Diskussion durchschnittlicher $\kappa$ -Werte

Generell lässt sich feststellen, dass die durchschnittlichen  $\kappa$ -Werte der vorliegenden Studie meist in einem Bereich zwischen 0,2 und 0,4 lagen und somit nach der Bewertung von Landis und Koch als „mittelmäßige“ Übereinstimmungen zu werten sind (Landis et al. 1977). In vielen Fällen lagen die einzelnen  $\kappa$ -Werte bei 0 oder sogar darunter, was bedeutet, dass die Übereinstimmung dem Zufall entsprach oder sogar gegenläufig war, was die Ergebnisse von Gitlin et al. (2004) bestärkt.

Ein direkter Vergleich mit der Literatur ist leider aus den in Abschnitt 2.5.2 genannten Gründen nicht möglich. Eine weitere Schwierigkeit in der Vergleichbarkeit der Studien liegt darin, dass  $\kappa$  nur in manchen Studien gewichtet wurde. Bei der Beurteilung der Röntgenbilder wurden sowohl ungewichtete als auch gewichtete  $\kappa$  berechnet. Bei der Beurteilung des Parenchyms (CT) wurde nicht gewichtet, bei der Pleura (CT) wurde gewichtet. Allgemein lässt sich festhalten, dass in Studien, wo sowohl ungewichtetes als



auch gewichtetes  $\kappa$  berechnet wurde, das gewichtete  $\kappa$  immer einen deutlich höheren Wert erlangte. Zudem waren sowohl die Anzahl der beurteilten Bilder in den Studien als auch die Grundvoraussetzungen sehr unterschiedlich. Sie reicht von 14 Bildern mit positivem Vorbefund (Begin et al. 1993) bis zu 7095 Bildern von einem repräsentativen Querschnitt der finnischen Bevölkerung (Impivaara et al. 1998). Die Anzahl der Beurteiler schwankte zwischen zwei und sechs und auch der Vergleichsmodus war sehr unterschiedlich: teilweise wurden Paare gebildet, deren Befunde direkt verglichen wurden, teilweise galt der Median der anderen Beurteiler als Gold-Standard, einmal ist von generalisiertem  $\kappa$  die Rede (Muir et al 1992), einige Male ist der Vergleichsmodus nicht genau nachzuvollziehen.

Trotz der schlechten Vergleichbarkeit der einzelnen  $\kappa$ -Werte lassen sich Tendenzen feststellen. Obwohl die Übereinstimmung meist mittelmäßig bis gut ist, sind erhebliche Schwankungen bei der Feststellung von Abnormalitäten zu erkennen (Welch et al. 1998). Impivaara et al. kommen zu dem Schluss, dass trotz detaillierter Richtlinien die Klassifikationen immer subjektiv sind (Impivaara et al. 1998). In einigen Studien wurde auch die Intraindividuelle Varianz (die Übereinstimmung der Befunde durch ein und denselben Beurteiler zu unterschiedlichen Zeitpunkten) untersucht, auf die hier jedoch nicht weiter eingegangen werden soll.

Tabelle 5.3 stellt die durchschnittlichen  $\kappa$ -Werte der einzelnen Beurteilerpaare dar. Paar 2+4 stach mit einem durchschnittlichen  $\kappa$  von 0,68 deutlich hervor. Nach Landis und Koch ist diese Übereinstimmung als beachtlich anzusehen (s. Tab. 2.21). Den nächstbesten Wert erlangte Paar 4+5 mit einem  $\kappa$  von 0,51, womit es im Bereich einer guten Leistung lag. Zwei Paare erlangten Werte von 0,17 und 0,20, welche nach Landis und Koch als schlechte Übereinstimmungen gelten.

Als mögliche Gründe für die eher schlechten  $\kappa$ -Werte könnte aufgeführt werden, dass die Kategorienzahl höher war als in anderen Studien, dass Veränderungen schwerer zu erkennen gewesen sein könnte als angenommen. Zudem haben nicht eindeutige Definitionen der Veränderungen insbesondere für parenchymale Veränderung zu schlechten Übereinstimmungen geführt. Desweiteren hatten die sieben Ärzte eine unterschiedliche medizinische Vorbildung - drei waren Arbeitsmediziner, vier waren Radiologen- und hatten zum Beispiel aufgrund ihres Alters sehr unterschiedliche praktische Erfahrungen mit der Interpretation von Pneumokoniosebildern. Manche von ihnen hatten vorher bereits eng zusammengearbeitet, andere nicht, so dass sie sich gegenseitig vor Studienbeginn unterschiedlich stark beeinflusst hatten. Paar 2 + 4 mit seiner überdurchschnittlich guten Übereinstimmung hatte zum Beispiel in der Vergangenheit eng zusammengearbeitet.

### 5.3.3. Diskussion Hypothese 1

Es wurde davon ausgegangen, dass bei der Interpretation von Röntgenbildern die Interreader-Varianz größer sein würde als bei der Interpretation von CT-Aufnahmen. CT-Aufnahmen sind leichter zu lesen und müssten deshalb zu besseren Übereinstimmungen führen, obwohl bei der Interpretation von Röntgenbildern bezüglich Pneumokoniosen viel mehr Übung besteht.

Diese Hypothese wurde in Abschnitt 5.2.1 bestätigt. Bei der Computertomographie

war die Übereinstimmung besser als beim Röntgen. Bezüglich des Medians zeigte die Computertomographie zu sechs Fragestellungen, die Röntgentechnik lediglich zu drei Fragestellungen bessere Ergebnisse. Bezüglich des Mittelwertes ergab sich ein Verhältnis CT:Röntgen von 5:3 (sowie bei einer Fragestellung gleichwertige Übereinstimmungen).

In der Literatur gab es nur eine Studie, die sich mit der Interreadervarianz sowohl bei der Interpretation von CT- als auch von Röntgenbildern befasste und gleichzeitig eine ähnliche Kategorisierung wie die Autorin dieser Arbeit vornahm (Begin et al. 1993). Sie kategorisierten nach den vier ILO-Hauptkategorien und verwendeten demnach eine Kategorie weniger als in der vorliegenden Studie. Demzufolge wären bei ihrer Studie etwas höhere  $\kappa$ -Werte zu erwarten. Bei der Beurteilung des Röntgenparenchyms erhielten sie bei der Beurteilung der Streuung kleiner Fleckschatten (rund und unregelmäßig) bei 61 Probanden und fünf Beurteilern einen  $\kappa$ -Median von 0,20. In der vorliegenden Studie wurde tatsächlich ein niedriger durchschnittlicher  $\kappa$ -Median bezüglich der Streuung von 0,13 erhalten.

Begin et al. erreichten bei der Beurteilung von Parenchymveränderungen auf CT-Bildern mit einem in Anlehnung an den ILO-1980-Bogen entwickelten Bogen bei den selben Fragestellungen einen  $\kappa$ -Wert von 0,28. In der vorliegenden Studie wurde bei den entsprechenden Fragestellungen auf den CT-Bögen ein durchschnittlicher  $\kappa$ -Median von 0,22 erreicht. Die absoluten Werte waren wie erwartet etwas schlechter. Dies bestätigt die Ergebnisse von Muir et al. (Muir et al. 1992), dass mehr Kategorien zu schlechteren Übereinstimmungen führen. Das Verhältnis der  $\kappa$ -Werte Röntgen/CT war bei Begin et al. etwa 0,7, in der vorliegenden Studie 0,6. Somit konnte die Ergebnisse von Begin et al., dass die CT dem Röntgen bezüglich der Beurteilung des Parenchyms überlegen ist, deutlich bestätigt werden.

Bei einer dichotomen Beurteilung erhielten sie Werte von 0,44 (Röntgen) und 0,63 (CT), was einem Verhältnis von 0,7 entspricht. In der vorliegenden Studie wurden bei einer dichotomen Beurteilung Werte von 0,10 (Röntgen) und 0,14 (CT) erlangt, was einem Verhältnis von 0,7 entspricht. Hierbei sind die absoluten Werte in der vorliegenden Studie bedeutend schlechter, das Verhältnis Röntgen/CT jedoch gleich, so dass auch hier die Ergebnisse von Begin et al. bestätigt werden können.

Eine andere Studie (Copley et al. 2001) befasste sich mit der Interreadervarianz bei der Beurteilung der Pleura sowohl beim Röntgen als auch bei der CT. Die Beurteilungsbögen waren jedoch nicht aufeinander abgestimmt, so dass man die  $\kappa$ -Werte der Röntgen- und CT-Verfahren nicht vergleichen kann. Zudem wurden andere Kategorien gewählt als in der vorliegenden Studie, so dass auch kein selektiver Vergleich der  $\kappa$ -Werte bei der Interpretation von Röntgenbildern möglich ist.

Im Abschnitt 5.2.1 wurde erarbeitet, dass die Computertomographie zu besseren Übereinstimmungen kam bei der Beurteilung des Vorhandenseins und der Streuung kleiner runder Fleckschatten, des Vorhandenseins kleiner unregelmäßiger Fleckschatten, der Ausdehnung und Dicke umschriebener Pleuraveränderungen und Pleuraverkalkungen. Das Röntgenverfahren erlangte lediglich höhere Übereinstimmungen bei der Beurteilung der Streuung kleiner unregelmäßiger Fleckschatten und der Ausdehnung und Dicke diffuser Pleuraverdickungen. An dieser Stelle zeigt sich das größere Potenzial des CT-Verfahrens. Der CT-Bogen wurde eigens für diese Studie entwickelt und CT-Aufnahmen waren bis

dato nur in Ausnahmefällen zum Erheben eines Pneumokoniosestatus üblich. Dementsprechend war die Übung in der Interpretation dieser Aufnahmen klein, ganz im Gegensatz zum etablierten Röntgenverfahren. Dennoch wurden in der Mehrheit bessere Übereinstimmung bei der Interpretation von CT-Bildern erzielt. Für die Zukunft ist zu erwarten, dass die Interreadervarianz bei der Computertomographie im Vergleich zu der des Röntgenverfahrens noch deutlicher sinken wird.

Es wurde angenommen, dass die Übereinstimmungen bei der CT-Beurteilung deshalb höher seien, weil die CT-Bilder einfacher zu lesen seien. Gälte dies pauschal, müssten bei allen Fragestellungen bessere Übereinstimmungen erzielt worden sein. Dies ist nicht der Fall. Dafür gibt es verschiedene Erklärungsmöglichkeiten. Zum einen ist es möglich, dass sich verschiedene Veränderungen in den unterschiedlichen Darstellungstechniken unterschiedlich gut erkennen lassen. Zum anderen ist es möglich, dass manche Definitionen des CT-Bogens nicht präzise genug waren, so dass zu viele eigene Definitionen bezüglich dieser Veränderungen entstanden, woraus naturgemäß schlechtere Übereinstimmungen resultierten. Es wäre auch zu überlegen, ob manche Veränderungen lediglich röntgentypisch sind, so dass man sie auf CT-Bildern kaum finden kann. In einem solchen Fall sollten entsprechende Veränderungen gar nicht auf dem CT-Bogen aufgeführt werden.

Muir et al. (1992) konnten feststellen, dass die Übereinstimmung höher wird, wenn kein Integrationsprozess im Kopf des Beurteilers (wie beim ILO-Bogen gefordert) stattfindet, sondern jeder Lungenteil einzeln beurteilt wird. So wurde beim CT vorgegangen und anschließend ein künstlicher Integrationsprozess durchgeführt. Nach den Ergebnissen von Muir et al. könnten die besseren  $\kappa$ -Werte bei der CT zumindest zum Teil auch dadurch erklärt werden.

Abschließend muss noch einmal betont werden, dass durch die unterschiedlichen Aufnahmetechniken ganz andere Ansichten möglich und dadurch andere Definitionen der Veränderungen nötig sind. Dementsprechend ist ein Vergleich der Werte der beiden Verfahren nur unter Vorbehalt möglich und möglicherweise nicht ganz so aussagekräftig, wie es im ersten Moment scheinen könnte.

#### 5.3.4. Diskussion Hypothese 2

Die zweite Hypothese besagt, dass die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen auf Röntgenaufnahmen größer sein würde als bei pleuralen Veränderungen, da sich pneumokoniotische Veränderungen besser im Pleurabereich erkennen lassen als im Parenchym. Sie wurde im Abs. 5.2.2 bestätigt. Der Pleuramedian war um den Faktor 3,2, der Pleuramittelwert um den Faktor 1,5 höher als der Parenchymmedian bzw. -mittelwert.

In der Literatur hat lediglich die Arbeitsgruppe von Welch et al. Parenchym und Pleura beurteilt (Welch et al. 1998). Bei der Beurteilung des Parenchyms lagen die  $\kappa$ -Werte bei einer dichotomen Beurteilung ( $<0/1$  gegen  $>0/1$ ) bei 0,04 bis 0,73, bei einer Beurteilung nach den vier Hauptkategorien bei -0,04 bis 0,58. Bei der Beurteilung der Pleura wurde bei einer dichotomen Beurteilung (Positivität, wenn Pneumokoniose (3A) und spezifische Befunde in den Sektionen 3B-D gefunden wurden)  $\kappa$ -Werte von 0,23 bis 0,65 erreicht. Hier ließen sich auf die  $\kappa$ -Werte bezogen keine ausgeprägten Unterschiede

bei der Beurteilung von Pleura und Parenchym nachweisen. Allerdings wurden die beiden Bereiche auch nicht durch ähnliche Kategorisierung aufeinander abgestimmt und lassen deshalb keinen genauen Vergleich zu.

Da man davon ausgehen kann, dass sich Verkalkungen durch einen deutlich anderen Grauwert auf Röntgenbildern besonders abheben, ist es interessant, die Pleurafragestellungen zusätzlich zu differenzieren in Pleuraveränderungen mit/ ohne Verkalkungen. Die Ergebnisse sind in Tab. 5.6 und Abb. 5.4 dargestellt.

Es kann festgehalten werden, dass die Übereinstimmung selbst bei den pleuralen Fragestellungen ohne Kalk deutlich höher (Median: Faktor 2,7, MW: Faktor 1,3) als bei den parenchymalen Fragestellungen. Demnach ist die bessere Übereinstimmung bei der Beurteilung der Pleura nicht alleine auf die leichte Beurteilung der Verkalkung zurückzuführen.

### 5.3.5. Diskussion Hypothese 3

Die dritte Hypothese besagt, dass die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen auf CT-Aufnahmen größer sein würde als bei pleuralen Veränderungen, da sich auch hier pneumokoniotische Veränderungen besser im Pleurabereich erkennen lassen als im Parenchym. Sie ließ sich im Abs. 5.2.3 bestätigen. Der Pleuramedian war um den Faktor 1,4, der Pleuramittelwert um den Faktor 1,1 höher als der Parenchymmedian bzw. -mittelwert. In der Literatur gibt es keine Studie, die vergleichend die Beurteilung der Pleura und des Parenchyms beurteilt.

Bei dieser Hypothese ist eine differenziertere Betrachtung in doppelter Hinsicht sinnvoll. Zum einen sollte auch hier der Einfluss der Fragestellungen nach Verkalkungen betrachtet werden, zum anderen ist für einen Vergleich mit dem Röntgenverfahren eine Reduktion der Parenchymfragestellungen auf die Fragestellungen nötig, die auf beiden Fragebögen vorkommen. Es wurde deshalb unterschieden zwischen Fragen, die auf beiden Bögen vorkommen (kleine runde und kleine unregelmäßige Fleckschatten, jeweils vorhanden/nicht vorhanden und Streuung) und allen Parenchymfragen, die auf dem CT-Bogen vorkommen (zusätzlich lineare Fleckschatten, Emphysem, Honeycombing, jeweils vorhanden/nicht vorhanden und Streuung) (s. Tab. 5.8 und Abb. 5.6).

Das Ergebnis ist nicht so eindeutig wie bei den Röntgenaufnahmen. Bei einer Betrachtung der Mediane lässt sich lediglich ein geringfügiger Unterschied zwischen den Pleurafragestellungen ohne Verkalkung und den gesamten Parenchymfragestellungen (inklusive lineare Fleckschatten, Emphysem, Honeycombing, jeweils vorhanden/ nicht vorhanden und Streuung) erkennen (Faktor 1,1 Pleura/Parenchym). Dagegen sind bei einer Mittelwertbetrachtung die  $\kappa$ -Werte der Beurteilungen des Parenchyms (komplett) höher als bei der Pleura ohne Verkalkung (Faktor 0,9 Pleura/Parenchym).

Dies zeigt, dass bei einer globalen Betrachtung aller Parenchymfragestellungen die Pleurae Beurteilung nur dann relevant besser ist, wenn auch Verkalkungen beurteilt werden. Auch im CT sind Verkalkungen durch einen anderen Grauton leichter zu erkennen als andere Veränderungen, so dass zu erwarten war, dass hierbei die Übereinstimmung besonders gut sei. Parenchymale Veränderungen lassen sich im CT aufgrund fehlender Überlagerungen leichter wahrnehmen als auf Röntgenaufnahmen. Dies wird ein maß-

geblicher Grund dafür sein, dass bei der Computertomographie die  $\kappa$ -Werte der pleuralen und parenchymalen Fragestellungen näher beieinanderliegen. Dementsprechend empfiehlt sich die Benutzung der Computertomographie besonders zur Feststellung parenchymaler Veränderungen.

Um einen korrekteren Vergleich mit den Röntgenergebnissen zu ermöglichen, dürften lediglich die Fragen betrachtet werden, die auf beiden Bögen vorkommen (s.o.). Hierbei wird ein etwas deutlicherer Unterschied zwischen parenchymalen und pleuralen Fragestellungen (ohne Kalk) ersichtlich: Median Faktor 1,5, MW Faktor 1,0 (Pleura/Parenchym). Dies lässt darauf schließen, dass die Beurteilung kleiner runder und kleiner unregelmäßiger Fleckschatten auf den CT-Bildern deutlich schwieriger ist als die Beurteilung von linearen Fleckschatten, Honeycombing und Emphysemen. Dies unterstützt den in der Diskussion der ersten Hypothese geäußerten Gedanken, dass die Verfahren jeweils bei bestimmten Fragen überlegen sein könnten. Demnach hat jedes Verfahren weiterhin eine gut begründete Berechtigung und die Anwendung beider Verfahren führt zu der größtmöglichen Ausschöpfung der Potentiale.

Bei der Beurteilung der Pleura beim Röntgen und CT war nicht nur die durchschnittliche Übereinstimmung meist besser, die Mediane und Mittelwerte lagen auch deutlich näher beieinander als bei der Beurteilung des Parenchyms. Dies zeigt, dass alle Paare bei der Beurteilung der Pleura geschlossen in sehr ähnlicher Qualität beurteilt haben.

Bei einer genauen Betrachtung lässt sich erkennen, dass bei der Computertomographie die Pleura-Mediane (ohne Kalk) meist deutlich höher lagen als die Parenchym-Mediane. Eine Ausnahme bildete jedoch das zweite Jahr, wo das Parenchym einen besseren Wert erlangte als die Pleurafragestellungen. In diesem Jahr gab es eine sehr große Übereinstimmung bei der Beurteilung der Streuung kleiner runder Fleckschatten, wodurch die gesamte Parenchym-Übereinstimmung dieses Jahres, und letztlich auch der Gesamt-Betrachtung, angehoben wurde. Es ist nicht sicher zu beurteilen, ob dieser eine stark abweichende Wert eine tatsächliche begründete bessere Übereinstimmung im 2. Jahr wiedergibt oder ob er ein Zufallsprodukt ist, der die Ergebnisse verfälscht.

### 5.3.6. Diskussion Hypothese 4 und 5

Die vierte und fünfte Hypothese postulieren, dass im zeitlichen Verlauf die Interreader-varianz bei der Beurteilung von Röntgen- und CT-Bildern sinkt, da hier ein Lernprozess stattfindet. Sie konnten im Abs. 5.2.4 und 5.2.5 bestätigt werden. Nach Wissen der Autorin ist dies die erste Studie weltweit, die das Lernverhalten bei der Interpretation von asbestinduzierten Pneumokoniosen untersuchte, dementsprechend ist kein Vergleich mit der Literatur möglich.

Es kann festgehalten werden, dass bei beiden Verfahren Lerneffekte nachzuvollziehen waren. Dies war zu erwarten, da sowohl von Jahr zu Jahr mehr Übung in der Interpretation von Pneumokoniosebildern bestand, als auch jährliche Konsensuskonferenzen stattfanden, um Definitionsfragen zu klären. Das Lernverhalten entspricht jedoch nicht den Erwartungen bezüglich der Ausprägung des Lernverhaltens bei beiden Verfahren. Es wäre zu erwarten, dass das Lernverhalten bei der Interpretation von CT-Bildern insgesamt ausgeprägter sei, da es sich hier um ein Pilot-Projekt handelte. Bei einem neu

eingeführten Verfahren ist anfangs das höchste Lernverhalten zu erwarten. Im Gegensatz dazu ist die ILO-Kategorisierung lange etabliert. Hier waren deshalb viel geringere Lerneffekte zu erwarten. Sie waren jedoch in einer vergleichbaren Größe. Vielleicht konnten hierbei die Konsensuskonferenzen zu einer Angleichung der bis dahin freieren Definitionen führen.

Bei der Interpretation von CT-Bildern war das Lernverhalten teilweise so ausgeprägt wie erwartet, teilweise wurden die Übereinstimmungen jedoch über die Zeit auch deutlich schlechter. Ein besonders ausgeprägtes Lernverhalten war bei der Computertomographie bei den Fragestellungen festzustellen, die nur auf dem CT-Bogen vorkommen (4,0 %). Bei den Fragestellungen, die auf beiden Bögen vorkommen, war das Lernverhalten beim Röntgenverfahren deutlicher (Rö: 1,4 %, CT: 0,3 %). Das eigentliche Lernpotenzial bei der CT-Interpretation zeigt sich bei den mit Honeycombing zusammenhängenden Fragestellungen. Die hierbei insgesamt hohen Zuwachsraten waren in einer zufriedenstellenden Größenordnung (14,6 bzw. 20 %). Ein deutliches Lernverhalten zeigt sich bei der CT neben dem Honeycombing auch bei kleinen runden Fleckschatten und Emphysemen (3,3 bis 10,4 %). Dagegen zeigt sich eine Verschlechterungstendenz bei linearen und kleinen unregelmäßigen Fleckschatten (-6,5 bis -10,6 %). Bei der Beurteilung der Pleuraverdickungen zeigte sich durchweg eine Verschlechterungstendenz (-0,2 bis -3,1 %), während bei der Beurteilung von Verkalkungen ein Lerneffekt nachzuweisen war (4,9 %). Die beiden zusammengehörigen Fragestellungen (z.B. Emphysem: Streuung; vorhanden/nicht vorhanden) wiesen immer die gleiche Lerntendenz auf. Dies könnte auf ein grundsätzliches Definitionsproblem der Veränderungen hinweisen (z.B. wann ist eine Pleuraverdickung umschrieben und wann diffus?). Die durchgeführten Konsensus-Konferenzen haben bei der CT offensichtlich nicht zu einer befriedigenden Klärung der offengebliebenen Fragen beigetragen. Vielleicht haben sich einzelne Beurteiler hier noch stärker eigene Definitionen gebildet als bei der ILO-Klassifikation und die Diskussionen haben immer wieder neue Ansichten zutage gebracht, wonach manche ihr Klassifikationsverhalten geändert haben und andere nicht. Es ist nicht anzunehmen, dass die Begründung lediglich in der Einfachheit/Schwere, eine Veränderung zu erkennen, zu suchen ist, da hierbei dennoch ein Lerneffekt stattfinden sollte.

## 5.4. Zusammenfassung

Die vorliegende Arbeit basiert auf Daten einer Längsschnittstudie und bearbeitet die Übereinstimmungsraten zweier Beurteiler bei der Befundung von pneumokoniotisch veränderten Röntgenbildern. Sie wurde im Zeitraum von 1993-1997 in sechs Zentren durchgeführt. Es wurde ein Hochrisikokollektiv zusammengestellt, aus dem 636 Personen bereit waren, jährlich einen Vorsorgetermin mit Röntgen- und CT-Diagnostik wahrzunehmen. Die Untersuchungen geschahen unter standardisierten Bedingungen, die Befunde wurden in standardisierte Bögen eingetragen (Röntgen: ILO-1980-Bogen, CT: selbstentworfenen Bogen, in Anlehnung an den ILO-1980-Bogen). Die Bilder wurden jeweils von zwei Beurteilern befundet, wobei der Zweitbeurteiler nicht über die Befunde des Erstbeurteilers informiert war. Insgesamt gab es sieben Beurteiler, die zehn Beurteilerpaare

bildeten. Sie hatten sich durch einen großen Erfahrungsreichtum in der Befundung von Pneumokoniosebildern ausgezeichnet. Die Variablen der Bögen wurden zu Kategorien zusammengefasst, anschließend wurden zur Beurteilung der Übereinstimmung  $\kappa$ -Werte errechnet.

Ziel dieser Arbeit war es, folgende fünf Hypothesen zu prüfen:

- 1) Bei der Interpretation von Röntgenbildern wird die Interreadervarianz größer sein als bei der Interpretation von CT-Aufnahmen. CT-Aufnahmen sind leichter zu lesen und führen deshalb zu übereinstimmenderen Aussagen, obwohl bei der Interpretation von Röntgenbildern bezüglich Pneumokoniosen viel mehr Übung besteht.
- 2) Bei Röntgen-Aufnahmen wird die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen größer sein als bei pleuralen Veränderungen. Allgemein lassen sich pneumokoniotische Veränderungen besser im Pleurabereich erkennen als im Parenchym.
- 3) Bei CT-Aufnahmen wird die Interreadervarianz bei der Beurteilung von parenchymalen Veränderungen größer sein als bei pleuralen Veränderungen mit der gleichen Begründung wie 2).
- 4) Im zeitlichen Verlauf wird die Interreadervarianz bei der Beurteilung von Röntgenbildern sinken, da ein Lernprozess stattfinden wird.
- 5) Im zeitlichen Verlauf wird die Interreadervarianz bei der Beurteilung von CT-Bildern sinken, da ein Lernprozess stattfinden wird.

Alle Hypothesen ließen sich bestätigen.

Bei insgesamt häufig nicht befriedigender Übereinstimmung stellt sich die Forderung der ILO nach Zweitbeurteilern als sehr begründet dar. Insgesamt ist festzustellen, dass die Computertomographie einen höheren Stellenwert bei der Vorsorgeuntersuchung ehemals asbeststaubexponierter Personen einnehmen sollte. Auch eine Überarbeitung der Bögen oder Kriterien zur Erstattung einer BK-Anzeige sowie eine kontrolliertere Ausbildung der Zweitbeurteiler könnten zu einer Verbesserung der Übereinstimmungen führen. Weiterhin sollte diskutiert werden, ob gegebenenfalls Ringversuche zur Qualitätsverbesserung und Schulung von Beurteilern sinnvoll sein könnten.

## 5.5. Ausblick

Die vorliegende Arbeit konnte zeigen, dass zwei Beurteiler bei der Interpretation von Röntgen- und CT-Bildern ehemals asbeststaubexponierter Personen oft nicht befriedigend übereinstimmen. Die Übereinstimmung war bei der CT höher als beim Röntgen



und jeweils bei Pleurabeurteilungen besser als bei Parenchymbeurteilungen. Ein Lerneffekt war bei beiden Verfahren zu beobachten, wobei das Lernverhalten beim Röntgen höher und bei der CT niedriger war als erwartet.

In zukünftigen Studien könnte man versuchen nachzuvollziehen, warum die Übereinstimmung bei der Beurteilung der Streuung kleiner unregelmäßiger Fleckschatten und der Ausdehnung und Dicke diffuser Pleuraverdickungen beim Röntgen besser war als bei der CT. Desweiteren könnte untersucht werden, ob der oben erwähnte stark abweichende Wert bei der Beurteilung der Streuung kleiner runder Fleckschatten im zweiten Jahr eine tatsächliche begründete bessere Übereinstimmung wiedergibt oder ob er ein Zufallsprodukt ist, der die Ergebnisse verfälscht. In einer weiteren Studie könnte der Frage auf den Grund gegangen werden, warum bei manchen Fragen ein ausgeprägtes Lernverhalten festzustellen war und bei anderen eine deutliche Verschlechterung. Dies könnte helfen, in Zukunft Bögen zu entwickeln, die verlässlicher beurteilt werden können.

Die ILO fordert, dass mindestens zwei Beurteiler ein Bild beurteilen. Diese Forderung zeigt sich dabei als sehr begründet. Wenn ein Beurteiler einen Befund als nicht pathologisch beurteilt, während der andere ihn als pathologisch ansieht, muss eine Diskussion stattfinden, die dem Patienten mit großer Wahrscheinlichkeit zugute kommen wird. So wird die Wahrscheinlichkeit höher, dass zum einen frühe Veränderungen eher erkannt werden, so dass vielleicht noch ein heilender Eingriff möglich sein wird, zum anderen werden Veränderungen gegebenenfalls eher als Berufskrankheit mit all ihren ökonomischen Folgen erkannt. Es kann also festgehalten werden, dass eine Zweitbeurteilung weiterhin stattfinden sollte, wobei über eine geregelte Ausbildung für Zweitbeurteiler (wie in den USA) nachgedacht werden sollte. Weiterhin sollte diskutiert werden, ob gegebenenfalls Ringversuche zur Qualitätsverbesserung und Schulung von Beurteilern sinnvoll sein könnten.

Die Computertomographie hat sich in dieser Studie bereits insofern bewährt, dass durch computertomographische Befunde mehr BK-Anzeigen gemacht wurden als alleine durch radiologische Befunde gemacht worden wären. Mit der vorliegenden Arbeit konnte zudem gezeigt werden, dass die Übereinstimmung bei der Computertomographie höher war als beim Röntgen. Dies zeigt deutlich, dass die Computertomographie zukünftig bei den Vorsorgeuntersuchungen einen viel höheren Stellenwert einnehmen sollte. Dabei sollten die Vorzüge der Computertomographie auch insofern genutzt werden, als eine weitere Bearbeitung der Daten durch den Computer unterstützt werden sollte. Zum Beispiel zeigt das Lernverhalten bei der Beurteilung von Pleuraverdickungen eine Verschlechterungstendenz und damit einen deutlichen Bedarf, sich mit diesem Bereich intensiver zu beschäftigen. Ein Aachener Elektrotechnikstudent entwickelte in seiner Diplomarbeit ein Computerprogramm, das den Beurteiler bei der Befundung von Pleuraverdickungen auf CT-Bildern unterstützt (Vogel S 2003). Die Fortentwicklung dieses Programmes würde zu exakteren Beurteilungen und zu einer besseren Übereinstimmung in der Befundung von Pleuraveränderungen führen und wäre damit sehr erstrebenswert.

Neben den weiterhin empfehlenswerten Zweitbeurteilungen wäre darüber nachzudenken, die Klassifikationsbögen nach den Ergebnissen dieser und zukünftiger Studien zu überarbeiten. Es konnte gezeigt werden, dass einige Beurteilungs-Differenzen durch uneindeutige Definitionen oder andere Defizite der Bögen bedingt sein könnten. Hier ist



---

eine Weiterentwicklung und vor allem Vereinfachung des CT-Bogens erforderlich. Zum Beispiel wäre zu überlegen, ob die Unterteilung in verschiedene Ausprägungen einer Veränderung (z.B. diffus/ umschrieben) weiterhin durchgehalten werden sollte. Die Ergebnisse dieser Studie weisen darauf hin, dass eine strikte Trennung in der Praxis oft nicht möglich ist. Sollte dies in zukünftigen Studien bestätigt werden, sollte daraus eine Zusammenfassung der nicht klar differenzierbaren Veränderungen oder eine Anpassung der Weisungen für die Erstattung einer BK-Anzeige resultieren.

# A. Röntgen-Bogen

**Satz III**  
 Vers.-Nr.: \_\_\_\_\_  
 Fam.- u. Vorname: \_\_\_\_\_

Datum der Untersuchung  
 Tag: \_\_\_\_\_ Monat: \_\_\_\_\_ Jahr: \_\_\_\_\_

**RÖNTGENBEFUND nach der ILO Klassifikation 1980/Bundesrepublik** (gem. Nr. 3.2.1/4.3 der Grundsätze G 1.1, G 1.2 bzw. G 1.3 sowie Nr. 5 G 1.2 bzw. G 1.3)

Bildgüte: \_\_\_\_\_ + \_\_\_\_\_ ± \_\_\_\_\_ ± \_\_\_\_\_ u \_\_\_\_\_

---

**Lunge**

**Kleine Schatten**  
 Runde Form  
 Größe: 

p	q	r
---	---	---

  
 Streuung: 

0-	1/0	2/1	3/2
0/0	1/1	2/2	3/3
0/1	1/2	2/3	3+

  
 Felder: 

RO	LO
RW	LM
RU	LU

**Kleine Schatten**  
 Unregelmäßige Form  
 Größe: 

s	t	u
---	---	---

  
 Streuung: 

0-	1/0	2/1	3/2
0/0	1/1	2/2	3/3
0/1	1/2	2/3	3+

  
 Felder: 

RO	LO
RW	LM
RU	LU

**Kleine Schatten**  
 Gemischte Formen  
 Größe: 

--	--	--

  
 Streuung: 

0-	1/0	2/1	3/2
0/0	1/1	2/2	3/3
0/1	1/2	2/3	3+

  
 Felder: 

RO	LO
RW	LM
RU	LU

**Große Schatten**  
 Größe: 

A	B	C
---	---	---

  
 Typ: 

wd	id
----	----

  
 Felder: 

RO	LO
RW	LM
RU	LU

---

**Pleura**

**Pleuraverdickung**  
 diffus  
 seitl. Brustwand  
 Größe: 

1	a
2	b
3	c

  
 Verbreitung/Dicke rechts: 

1	a
2	b
3	c

  
 Verbreitung/Dicke links: 

1	a
2	b
3	c

  
 Felder: 

RO	LO
RW	LM
RU	LU

**Adhärenz des kostophrenischen Winkels**  
 Größe: 

o.B.
------

  
 Seite: 

R
---

  
 Seite: 

L
---

**Pleuraverdickung**  
 unschneben (Plaques)  
 Größe: 

1	a
2	b
3	c

  
 Verbreitung/Dicke rechts: 

1	a
2	b
3	c

  
 Verbreitung/Dicke links: 

1	a
2	b
3	c

  
 Lokalisation: 

Zwerdteil	R	L
Brustwand	R	L

**Pleuraverkalkung**  
 unschneben (Plaques)  
 Größe: 

1	2	3
---	---	---

  
 Verbreitung/Dicke rechts: 

1
2
3

  
 Verbreitung/Dicke links: 

1
2
3

  
 Lokalisation: 

Zwerdteil	R	L
Brustwand	R	L
Sonstige	R	L

---

**ERGÄNZENDE BEFUNDE** (in begründeten Fällen gem. Nr. 3.2.1/4.3 der Grundsätze G 1.1, G 1.2 bzw. G 1.3 sowie Nr. 5 G 1.2 bzw. G 1.3)

\_\_\_\_\_

---

**BK-BEURTEILUNG**

Keine Hinweise auf anzeigenpflichtige Veränderungen  Begründung: \_\_\_\_\_

Anzeigespflicht: Begründeter Verdacht auf

- Silikose (BK-Nr. 4101)
- Siliko-Tuberkulose (BK-Nr. 4102)
- Asbestose Lunge/Pleura (BK-Nr. 4103)
- Asbestassoziierter Lungenkrebs (BK-Nr. 4104)
- Asbestassoziierter Kehlkopfkrebs (BK-Nr. 4104)
- Pleuranesotheliom (BK-Nr. 4105)

Sonstiges: \_\_\_\_\_  
 (Bitte BK-Anzeige erstellen und an den zuständigen Unt-Träger senden sowie den Versicherten unterrichten!)

---

**VORSCHLÄGE UND/ODER VERANLASSTE MASSNAHMEN**

\_\_\_\_\_

Stempel und Unterschrift des Arztes

Abbildung A.1.: Röntgen-Bogen

# B. CT-Bogen

Name / Versicherungsnummer / Aufkleber		Einstelldaten CT		Datum:		Bildgüte	
		Nr.:		Schichtzahl:		+ 1050	
		Schichtlage		Schichtdicke:		+/- 1051	
				Fensterlagen: HU +/- 1052 1.) HU U 1053 2.) HU 3.) HU		Lagerung	
				KV:    mA:    S:		BL    RL	

**Lunge:**

<b>KFS rundlich</b> p` q` r` 601 602 603	Streuung/Felder 0 604	RO 1 605 2 606 3 607	RM 1 608 2 609 3 610	RU 1 611 2 612 3 613	LO 1 614 2 615 3 616	LM 1 617 2 618 3 619	LU 1 620 2 621 3 622	Symbole # 1010 ax 1011 be 1012 br 1013 bu 1014 ca 1015 cl 1016 cn 1017 co 1018 cp 1019 cv 1020 di 1021 ef 1022 es 1023 fp 1024 fr 1025 gg 1026 hi 1027 id 1028 ih 1029 kl 1030 me 1031 od 1032 pi 1033 px 1034 ra 1035 rp 1036 tba 1037 tbu 1038
<b>KFS irregulär</b> s` t` u` 623 624 625	0 626	1 627 2 628 3 629	1 630 2 631 3 632	1 633 2 634 3 635	1 636 2 637 3 638	1 639 2 640 3 641	1 642 2 643 3 644	
<b>FS linear</b> v` w` x` 645 646 647	0 648	1 649 2 650 3 651	1 652 2 653 3 654	1 655 2 656 3 657	1 658 2 659 3 660	1 661 2 662 3 663	1 664 2 665 3 666	
<b>Häufigste Formen</b> 1. 2. 3. 667 668 669	0 670	1 671 2 672 3 673	1 674 2 675 3 676	1 677 2 678 3 679	1 680 2 681 3 682	1 683 2 684 3 685	1 686 2 687 3 688	
<b>Honeycombing</b> subpl Mantel Kern 689 690 691	0 692	1 693 2 694 3 695	1 696 2 697 3 698	1 699 2 700 3 701	1 702 2 703 3 704	1 705 2 706 3 707	1 708 2 709 3 710	
<b>Emphysem</b> subpl azi. cica. 711 712 713	0 714	1 715 2 716 3 717	1 718 2 719 3 720	1 721 2 722 3 723	1 724 2 725 3 726	1 727 2 728 3 729	1 730 2 731 3 732	
<b>Große Schatten</b> wd id ag 733 734 735	0 736	1 737 2 738 3 739	1 740 2 741 3 742	1 743 2 744 3 745	1 746 2 747 3 748	1 749 2 750 3 751	1 752 2 753 3 754	

**Pleura:**  
 W = Thoraxwand  
 M = Mediastinum  
 D = Diaphragma

**Ausdehnung:**  
 1 = <math>\leq \frac{1}{2}</math> Circumferenz/Seite  
 2 = <math>\frac{1}{2}</math> Circumferenz/Seite  
 3 = > <math>\frac{1}{2}</math> Circumferenz/Seite

**Dicke:**  
 a = < 5 mm  
 b = 5 - 10 mm  
 c = > 10 mm

K = Verkalkung  
 (keine Angaben zur  
 Dicke o. Ausdehnung)

<b>Pleuraverdickung, unbeschrieben</b> W M D 1101 1102 1103	0 1104	erhaben 1105 nicht erhaben 1106	Lunge + 1107 Lunge # 1108
<b>Pleuraverdickung, gleichförmig</b> W M D 1221 1222 1223	0 1224	Lunge + 1225 Lunge # 1226	

<table border="1" style="width:100%; border-collapse: collapse;"> <tr><th colspan="4">R</th><th colspan="4">L</th></tr> <tr><th>RO</th><th>RM</th><th>RU</th><th>RD</th><th>LO</th><th>LM</th><th>LU</th><th>LD</th></tr> <tr><td>1</td><td>1109</td><td>1116</td><td>1123</td><td>1130</td><td>1</td><td>1165</td><td>1172</td><td>1179</td><td>1186</td></tr> <tr><td>2</td><td>1110</td><td>1117</td><td>1124</td><td>1131</td><td>2</td><td>1166</td><td>1173</td><td>1180</td><td>1187</td></tr> <tr><td>3</td><td>1111</td><td>1118</td><td>1125</td><td>1132</td><td>3</td><td>1167</td><td>1174</td><td>1181</td><td>1188</td></tr> <tr><td>a</td><td>1112</td><td>1119</td><td>1126</td><td>1133</td><td>a</td><td>1168</td><td>1175</td><td>1182</td><td>1189</td></tr> <tr><td>b</td><td>1113</td><td>1120</td><td>1127</td><td>1134</td><td>b</td><td>1169</td><td>1176</td><td>1183</td><td>1190</td></tr> <tr><td>c</td><td>1114</td><td>1121</td><td>1128</td><td>1135</td><td>c</td><td>1170</td><td>1177</td><td>1184</td><td>1191</td></tr> <tr><td>K</td><td>1115</td><td>1122</td><td>1129</td><td>1136</td><td>K</td><td>1171</td><td>1178</td><td>1185</td><td>1192</td></tr> </table>	R				L				RO	RM	RU	RD	LO	LM	LU	LD	1	1109	1116	1123	1130	1	1165	1172	1179	1186	2	1110	1117	1124	1131	2	1166	1173	1180	1187	3	1111	1118	1125	1132	3	1167	1174	1181	1188	a	1112	1119	1126	1133	a	1168	1175	1182	1189	b	1113	1120	1127	1134	b	1169	1176	1183	1190	c	1114	1121	1128	1135	c	1170	1177	1184	1191	K	1115	1122	1129	1136	K	1171	1178	1185	1192	<table border="1" style="width:100%; border-collapse: collapse;"> <tr><th colspan="4">R</th><th colspan="4">L</th></tr> <tr><th>RO</th><th>RM</th><th>RU</th><th>RD</th><th>LO</th><th>LM</th><th>LU</th><th>LD</th></tr> <tr><td>1</td><td>1137</td><td>1144</td><td>1151</td><td>1158</td><td>1</td><td>1193</td><td>1200</td><td>1207</td><td>1214</td></tr> <tr><td>2</td><td>1138</td><td>1145</td><td>1152</td><td>1159</td><td>2</td><td>1194</td><td>1201</td><td>1208</td><td>1215</td></tr> <tr><td>3</td><td>1139</td><td>1146</td><td>1153</td><td>1160</td><td>3</td><td>1195</td><td>1202</td><td>1209</td><td>1216</td></tr> <tr><td>a</td><td>1140</td><td>1147</td><td>1154</td><td>1161</td><td>a</td><td>1196</td><td>1203</td><td>1210</td><td>1217</td></tr> <tr><td>b</td><td>1141</td><td>1148</td><td>1155</td><td>1162</td><td>b</td><td>1197</td><td>1204</td><td>1211</td><td>1218</td></tr> <tr><td>c</td><td>1142</td><td>1149</td><td>1156</td><td>1163</td><td>c</td><td>1198</td><td>1205</td><td>1212</td><td>1219</td></tr> <tr><td>K</td><td>1143</td><td>1150</td><td>1157</td><td>1164</td><td>K</td><td>1199</td><td>1206</td><td>1213</td><td>1220</td></tr> </table>	R				L				RO	RM	RU	RD	LO	LM	LU	LD	1	1137	1144	1151	1158	1	1193	1200	1207	1214	2	1138	1145	1152	1159	2	1194	1201	1208	1215	3	1139	1146	1153	1160	3	1195	1202	1209	1216	a	1140	1147	1154	1161	a	1196	1203	1210	1217	b	1141	1148	1155	1162	b	1197	1204	1211	1218	c	1142	1149	1156	1163	c	1198	1205	1212	1219	K	1143	1150	1157	1164	K	1199	1206	1213	1220
R				L																																																																																																																																																																									
RO	RM	RU	RD	LO	LM	LU	LD																																																																																																																																																																						
1	1109	1116	1123	1130	1	1165	1172	1179	1186																																																																																																																																																																				
2	1110	1117	1124	1131	2	1166	1173	1180	1187																																																																																																																																																																				
3	1111	1118	1125	1132	3	1167	1174	1181	1188																																																																																																																																																																				
a	1112	1119	1126	1133	a	1168	1175	1182	1189																																																																																																																																																																				
b	1113	1120	1127	1134	b	1169	1176	1183	1190																																																																																																																																																																				
c	1114	1121	1128	1135	c	1170	1177	1184	1191																																																																																																																																																																				
K	1115	1122	1129	1136	K	1171	1178	1185	1192																																																																																																																																																																				
R				L																																																																																																																																																																									
RO	RM	RU	RD	LO	LM	LU	LD																																																																																																																																																																						
1	1137	1144	1151	1158	1	1193	1200	1207	1214																																																																																																																																																																				
2	1138	1145	1152	1159	2	1194	1201	1208	1215																																																																																																																																																																				
3	1139	1146	1153	1160	3	1195	1202	1209	1216																																																																																																																																																																				
a	1140	1147	1154	1161	a	1196	1203	1210	1217																																																																																																																																																																				
b	1141	1148	1155	1162	b	1197	1204	1211	1218																																																																																																																																																																				
c	1142	1149	1156	1163	c	1198	1205	1212	1219																																																																																																																																																																				
K	1143	1150	1157	1164	K	1199	1206	1213	1220																																																																																																																																																																				
<table border="1" style="width:100%; border-collapse: collapse;"> <tr><th colspan="4">R</th><th colspan="4">L</th></tr> <tr><th>RO</th><th>RM</th><th>RU</th><th>RD</th><th>LO</th><th>LM</th><th>LU</th><th>LD</th></tr> <tr><td>1</td><td>1227</td><td>1234</td><td>1241</td><td>1248</td><td>1</td><td>1283</td><td>1290</td><td>1297</td><td>1304</td></tr> <tr><td>2</td><td>1228</td><td>1235</td><td>1242</td><td>1249</td><td>2</td><td>1284</td><td>1291</td><td>1298</td><td>1305</td></tr> <tr><td>3</td><td>1229</td><td>1236</td><td>1243</td><td>1250</td><td>3</td><td>1285</td><td>1292</td><td>1299</td><td>1306</td></tr> <tr><td>a</td><td>1230</td><td>1237</td><td>1244</td><td>1251</td><td>a</td><td>1286</td><td>1293</td><td>1300</td><td>1307</td></tr> <tr><td>b</td><td>1231</td><td>1238</td><td>1245</td><td>1252</td><td>b</td><td>1287</td><td>1294</td><td>1301</td><td>1308</td></tr> <tr><td>c</td><td>1232</td><td>1239</td><td>1246</td><td>1253</td><td>c</td><td>1288</td><td>1295</td><td>1302</td><td>1309</td></tr> <tr><td>K</td><td>1233</td><td>1240</td><td>1247</td><td>1254</td><td>K</td><td>1289</td><td>1296</td><td>1303</td><td>1310</td></tr> </table>	R				L				RO	RM	RU	RD	LO	LM	LU	LD	1	1227	1234	1241	1248	1	1283	1290	1297	1304	2	1228	1235	1242	1249	2	1284	1291	1298	1305	3	1229	1236	1243	1250	3	1285	1292	1299	1306	a	1230	1237	1244	1251	a	1286	1293	1300	1307	b	1231	1238	1245	1252	b	1287	1294	1301	1308	c	1232	1239	1246	1253	c	1288	1295	1302	1309	K	1233	1240	1247	1254	K	1289	1296	1303	1310	<table border="1" style="width:100%; border-collapse: collapse;"> <tr><th colspan="4">R</th><th colspan="4">L</th></tr> <tr><th>RO</th><th>RM</th><th>RU</th><th>RD</th><th>LO</th><th>LM</th><th>LU</th><th>LD</th></tr> <tr><td>1</td><td>1255</td><td>1262</td><td>1269</td><td>1276</td><td>1</td><td>1311</td><td>1318</td><td>1325</td><td>1332</td></tr> <tr><td>2</td><td>1256</td><td>1263</td><td>1270</td><td>1277</td><td>2</td><td>1312</td><td>1319</td><td>1326</td><td>1333</td></tr> <tr><td>3</td><td>1257</td><td>1264</td><td>1271</td><td>1278</td><td>3</td><td>1313</td><td>1320</td><td>1327</td><td>1334</td></tr> <tr><td>a</td><td>1258</td><td>1265</td><td>1272</td><td>1279</td><td>a</td><td>1314</td><td>1321</td><td>1328</td><td>1335</td></tr> <tr><td>b</td><td>1259</td><td>1266</td><td>1273</td><td>1280</td><td>b</td><td>1315</td><td>1322</td><td>1329</td><td>1336</td></tr> <tr><td>c</td><td>1260</td><td>1267</td><td>1274</td><td>1281</td><td>c</td><td>1316</td><td>1323</td><td>1330</td><td>1337</td></tr> <tr><td>K</td><td>1261</td><td>1268</td><td>1275</td><td>1282</td><td>K</td><td>1317</td><td>1324</td><td>1331</td><td>1338</td></tr> </table>	R				L				RO	RM	RU	RD	LO	LM	LU	LD	1	1255	1262	1269	1276	1	1311	1318	1325	1332	2	1256	1263	1270	1277	2	1312	1319	1326	1333	3	1257	1264	1271	1278	3	1313	1320	1327	1334	a	1258	1265	1272	1279	a	1314	1321	1328	1335	b	1259	1266	1273	1280	b	1315	1322	1329	1336	c	1260	1267	1274	1281	c	1316	1323	1330	1337	K	1261	1268	1275	1282	K	1317	1324	1331	1338
R				L																																																																																																																																																																									
RO	RM	RU	RD	LO	LM	LU	LD																																																																																																																																																																						
1	1227	1234	1241	1248	1	1283	1290	1297	1304																																																																																																																																																																				
2	1228	1235	1242	1249	2	1284	1291	1298	1305																																																																																																																																																																				
3	1229	1236	1243	1250	3	1285	1292	1299	1306																																																																																																																																																																				
a	1230	1237	1244	1251	a	1286	1293	1300	1307																																																																																																																																																																				
b	1231	1238	1245	1252	b	1287	1294	1301	1308																																																																																																																																																																				
c	1232	1239	1246	1253	c	1288	1295	1302	1309																																																																																																																																																																				
K	1233	1240	1247	1254	K	1289	1296	1303	1310																																																																																																																																																																				
R				L																																																																																																																																																																									
RO	RM	RU	RD	LO	LM	LU	LD																																																																																																																																																																						
1	1255	1262	1269	1276	1	1311	1318	1325	1332																																																																																																																																																																				
2	1256	1263	1270	1277	2	1312	1319	1326	1333																																																																																																																																																																				
3	1257	1264	1271	1278	3	1313	1320	1327	1334																																																																																																																																																																				
a	1258	1265	1272	1279	a	1314	1321	1328	1335																																																																																																																																																																				
b	1259	1266	1273	1280	b	1315	1322	1329	1336																																																																																																																																																																				
c	1260	1267	1274	1281	c	1316	1323	1330	1337																																																																																																																																																																				
K	1261	1268	1275	1282	K	1317	1324	1331	1338																																																																																																																																																																				

<b>Ergüßvershwart.</b> O 1339 R 1340 L 1341	<b>Resterguß</b> O 1342 R 1343 L 1344	<b>Adhärenz CWP</b> O 1345 R 1346 L 1347
--	--	---

<b>BK-unabhängige Erkrankung:</b> Berufskrankheit BK-Nr. Silikose/Parenchym 4101 1060 Silikose/Pleura 4101 1061 Siliko-Tbc 4102 1062 Carcinom 1063	<b>Asbestose/Parenchym</b> 4103 1064 <b>Asbestose/Pleura</b> 4103 1065 <b>Carcinom/Parenchym</b> 4104 1066 <b>Carcinom/Pleura</b> 4104 1067 <b>Mesotheliom</b> 4105 1068	Unterschrift: Datum:
---	--	-------------------------

Abbildung B.1.: CT-Bogen

# Abkürzungsverzeichnis

Im folgenden werden häufig verwendete Abkürzungen aufgeführt:

BK	Berufskrankheit
CXR	Röntgen
CT	Computertomographie
d	dichotom
Em S	Emphysem: Streuung
Em v/n	Emphysem: vorhanden/ nicht vorhanden
HK	Hauptkriterien
Hon S	Honeycombing: Streuung
Hon v/n	Honeycombing: vorhanden/ nicht vorhanden
HRCT	High Resolution Computertomography
ILO	International Labour Organization
IQ	Interquartil
Kalk v/n	Kalk: vorhanden/ nicht vorhanden
$\kappa$	Kappa: Übereinstimmungsmaß zweier Beurteiler
$\kappa_w$	gewichtetes Kappa
KFR S	kleine runde Fleckschatten: Streuung
KFR v/n	kleine runde Fleckschatten: vorhanden/ nicht vorhanden
KFu S	kleine unregelmäßige Fleckschatten: Streuung
KFu v/n	kleine unregelmäßige Fleckschatten: vorhanden/ nicht vorhanden
KZ	Kategorienanzahl
linFS S	lineare Fleckschatten: Streuung
linFS v/n	lineare Fleckschatten: vorhanden/ nicht vorhanden
Med	Median
MW	Mittelwert
PdA	diffuse Pleuraverdickung: Ausdehnung
PdD	diffuse Pleuraverdickung: Dicke
PuA	umschriebene Pleuraverdickung: Ausdehnung
PuD	umschriebene Pleuraverdickung: Dicke
Rö	Röntgen
s	Standardabweichung
v/n	vorhanden/ nicht vorhanden
ZAs	Zentralstelle asbeststaubgefährdeter Arbeitnehmer

# Tabellenverzeichnis

2.1. Liste der Berufskrankheiten . . . . .	13
2.2. Rö: Pneumokoniosezeichen . . . . .	15
2.3. BK-Anzeige: Rö-Parenchym . . . . .	16
2.4. BK-Anzeige: Rö-Pleura . . . . .	16
2.5. ILO-Bogen: Symbole . . . . .	18
2.7. ILO-Bogen: Parenchym . . . . .	19
2.8. ILO-Bogen: Pleura . . . . .	20
2.9. CT: Pneumokoniosezeichen . . . . .	21
2.10. Positiver prädiktiver Wert: Rö, CT und HRCT . . . . .	21
2.11. CT-Bogen: Parenchym . . . . .	24
2.13. CT-Bogen: Pleura . . . . .	25
2.14. Prototyp Vierfeldertafel . . . . .	25
2.15. $\kappa$ -Beispiel 1 . . . . .	25
2.16. $\kappa$ -Beispiel 2 . . . . .	26
2.17. Randsummen $\kappa$ -Beispiel 1 . . . . .	26
2.18. Erwartete Übereinstimmung 1 . . . . .	26
2.19. Erwartete Übereinstimmung 2 . . . . .	27
2.20. Erwartete Übereinstimmung Beispiel 2 . . . . .	28
2.21. $\kappa$ -Bewertung nach Landis und Koch . . . . .	29
2.22. Studien zum Thema Interreadervarianz bei der Beurteilung a) des Parenchym chym auf Röntgenbildern . . . . .	30
2.24. b) der Pleura auf Röntgenbildern . . . . .	31
2.25. c) des Parenchym auf CT-Bildern . . . . .	32
2.26. d) der Pleura auf CT-Bildern . . . . .	33
4.1. Die Untersuchungszentren der Studie . . . . .	41
4.2. CT: BK-Anzeige . . . . .	42
4.3. CT-Geräte der Studie . . . . .	42
4.4. CT-Einstellparameter . . . . .	42
4.5. Anzahl der untersuchten Probanden pro Jahr . . . . .	43
4.6. Die Beurteiler der Studie . . . . .	44
4.7. Beurteilerpaare: Anzahl beurteilter Personen . . . . .	44
4.8. CT: Parenchymbefunde . . . . .	45
4.9. CT: Pleurabefunde . . . . .	45
5.1. $\kappa$ -Werte: Rö . . . . .	51

---

5.2. $\kappa$ -Werte: CT . . . . .	54
5.3. Durchschnittliches $\kappa$ der einzelnen Paare; ermittelt als Mittelwert sämtlicher $\kappa$ -Werte dieses Paares . . . . .	57
5.4. Verfahrensvergleich bzgl. Frage und Jahr . . . . .	59
5.5. Rö: $\kappa$ -Parenchym und Pleura . . . . .	61
5.6. Rö: $\kappa$ -Parenchym und Pleura +/- Kalk . . . . .	61
5.7. CT: $\kappa$ -Parenchym und Pleura . . . . .	62
5.8. CT: $\kappa$ -Parenchym und Pleura +/- Kalk . . . . .	63
5.9. Beispieltabelle Lernverhalten . . . . .	64
5.10. Rö: Lerneffekt . . . . .	66
5.11. CT: Lerneffekt Teil 1 . . . . .	66
5.12. CT: Lerneffekt Teil 2 . . . . .	68

# Abbildungsverzeichnis

5.1. Median: überlegenes Verfahren . . . . .	58
5.2. MW: überlegenes Verfahren . . . . .	58
5.3. Rö: $\kappa$ Parenchym und Pleura . . . . .	61
5.4. Rö: $\kappa$ Parenchym und Pleura . . . . .	62
5.5. CT: $\kappa$ -MW Parenchym und Pleura . . . . .	63
5.6. CT: $\kappa$ -MW Parenchym und Pleura +/- Kalk . . . . .	64
5.7. Rö: zeitl. Medianverläufe . . . . .	65
5.8. Rö: zeitl. Mittelwertverläufe . . . . .	65
5.9. CT: zeitl. Medianverläufe der mit Rö gemeinsamen Fragestellungen . . . . .	67
5.10. CT: zeitl. Mittelwertverläufe der mit Rö gemeinsamen Fragestellungen . . . . .	67
5.11. CT: zeitl. Medianverläufe der exklusiven CT-Bogen-Fragestellungen . . . . .	67
5.12. CT: zeitl. MW-Verläufe der exklusiven CT-Bogen-Fragestellungen . . . . .	68
A.1. Röntgen-Bogen . . . . .	82
B.1. CT-Bogen . . . . .	83

# Literaturverzeichnis

- [1] Aberle DR, Gamsu G, Ray CS: High-Resolution CT of Benign Asbestos-Related Diseases: Clinical and Radiographic Correlation; American Journal of Roentgenology 1988; 151: 883-891
- [2] Aberle DR, Gamsu G, Ray CS, Feuerstein IM: Asbestos-Related Pleural and Parenchymal Fibrosis: Detection with High-Resolution CT; Radiology 1988; 166: 729-734
- [3] Akira M, Yamamoto S, Yokoyama K, Kita N, Moriga K, Higashihara T, Kozuka T: Asbestosis: High-Resolution CT - Pathologic Correlation; Radiology 1990; 176: 389-394
- [4] Akira M, Yokoyama K, Yamamoto S, Higashihara T, Moriga K, Kita N, Morimoto S, Ikezoe J, Kozuka T: Early Asbestosis: Evaluation with High-Resolution CT; Radiology 1991; 178: 409-416
- [5] Albin M, Horstmann V, Jakobsson K, Welinder H: Survival in cohorts of asbestos cement workers and controls; Occupational and Environmental Medicine 1996; 53: 87-93
- [6] Banaei A, Auvert B, Goldberg M, Gueguen A, Luce D, Goldberg S: Future trends in mortality of French men from mesothelioma; Occupational and Environmental Medicine 2000; 57: 488-494
- [7] Baur X, Marczynski B, Rozynek P, Voss B: Bronchopulmonale Präkanzerosen und Tumoren - Risikogruppen aus arbeitsmedizinischer Sicht; Pneumologie 1994; 48: 825-834
- [8] Begin R, Ostiguy G, Filion R, Colman N, Bertrand P: Computed tomography in the early detection of asbestosis; British Journal of Industrial Medicine 1993; 50: 689-698
- [9] Bergin CJ, Castellino RA, Blank N, Moses L: Specificity of High-Resolution CT Findings in Pulmonary Asbestosis: Do Patients Scanned for Other Indications Have Similar Findings?; American Journal of Roentgenology 1994; 163: 551-555
- [10] Berry G, Newhouse ML, Wagner JC: Mortality from all cancers of asbestos factory workers in east London 1933-80; Occupational and Environmental Medicine 2000; 57: 782-785



- [11] Bessis L, Callard P, Gotheil C, Biaggi A, Grenier P: High-Resolution CT of Parenchymal Lung Disease: Precise Correlation with Histologic Findings; *RadioGraphics* 1992; 12: 45-58
- [12] BK-DOK '99, Dokumentation des Berufskrankheiten-Geschehens in Deutschland; HVBG, Mai 2001
- [13] Bortz, Lienert: Kurzgefasste Statistik für die klinische Forschung; Springer-Verlag Berlin Heidelberg 1998
- [14] Brodtkin CA, McCullough J, Stover B, Balmes J, Hammar S, Omenn GS, Checkoway H, Barnhart S: Lobe of Origin and Histologic Type of Lung Cancer Associated With Asbestos Exposure in the Carotene and Retinol Efficacy Trial (CARET); *American Journal of Industrial Medicine* 1997; 32: 582-591
- [15] Butz M: Beruflich verursachte Krebserkrankungen - Eine Darstellung der im Zeitraum 1978 bis 2000 anerkannten Berufskrankheiten; HVBG, April 2002
- [16] Coenen W, Schenk H: Ermittlung von Risikogruppen bei Asbestexponierten. *Die BG* 1990; 718-726
- [17] Copley SJ, Wells AU, Rubens MB: Functional Consequences of Pleural Disease Evaluated with Chest Radiography and CT; *Radiology* 2001; 220: 237-243
- [18] de Klerk NH, Musk AW, Eccles JL, Hansen J, Hobbs MS: Exposure to crocidolite and the incidence of different histological types of lung cancer; *Occupational and Environmental Medicine* 1996; 53: 157-159
- [19] De Raeve H, Verschakelen JA, Gevenois PA: Observer variation in computed tomography of pleural lesions in subjects exposed to indoor asbestos; *Eur Resp J* 2001; 17: 916-921
- [20] Drechsel-Schlund C: Asbestverursachte Berufskrankheiten in Deutschland - Entstehung und Prognose, HVBG, August 2003
- [21] Finkelstein MM: Radiographic Asbestosis Is Not a Prerequisite for Asbestos-Associated Lung Cancer in Ontario Asbestos-Cement Workers; *American Journal of Industrial Medicine* 1997; 32: 341-348
- [22] Förster H: Asbestos and Other Fibrous Silicates. Health Effects; *Advanced Mineralogy* 1998; 312-322
- [23] Friedman AC, Fiel SB, Fisher MS, Radecki PD, Lev-Toaff AS, Caroline DF: Asbestos-Related Pleural Disease and Asbestosis: A comparison of CT and Chest Radiography; *American Journal of Roentgenology* 1988; 150: 269-275
- [24] Frumkin H, Pransky G, Cosmatos I: Radiologic Detection of Pleural Thickening; *Am Rev Respir Dis* 1990; 142:1325-1330

- [25] Gitlin JN, Cook LL, Linton OW, Garrett-Mayer E: Comparison of "B"Readers' Interpretations of Chest Radiographs for Asbestos Related Changes, *Academic Radiology* 2004, 11: 843-856
- [26] Hausmann K, Calavrezos A, Voitowitz H-J: Asbestbedingte Non-Hodgkin-Lymphome und Plasmozytome. Schlüsselbeobachtungen, nachfolgende Beobachtungsreihen mit Krankheitsclustern und epidemiologische Ergebnisse; *Zbl Arbeitsmed* 2000; 50: 358-371
- [27] Havemann K, Wolf M: Lungentumoren; Ostendorf PC und Seeber S: Hämatologie. *Onkologie* 1997; 606-643
- [28] Hering KG, Jacobsen M, Bosch-Galetke E, Elliehausen HJ, Hieckel HG, Hofmann-Preiss K, Jacques W, Jeremie U, Kotschy-Lang N, Kraus T, Menze B, Raab W, Raithel HJ, Schneider WD, Strassburger K, Tuengerthal S, Voitowitz HJ, ILO: Further development of the International Pneumoconiosis Classification -from ILO 1980 to ILO 2000 and to ILO 2000/German Federal Republic version; *Pneumologie* 2003; 57:576-584
- [29] Hering KG, Tuengerthal S, Kraus T: Standardisierte CT/HRCT-Klassifikation der Bundesrepublik Deutschland für arbeitsbedingte Thoraxerkrankungen; *Radiologie* 2004, 44:500-511
- [30] Hillerdal G: Pleural Plaques and Risk for Bronchial Carcinoma and Mesothelioma - A Prospective Study; *Chest* 1994; 105: 144-150
- [31] Internationales Arbeitsamt Genf, Richtlinien für die Anwendung der internationalen Klassifikation des IAA von Pneumokoniosen-Röntgenfilmen, überarbeitete Ausgabe 1980; 1981
- [32] Impivaara O, Zitting AJ, Kuusela T, Alanen E, Karjalainen A: Observer Variation in Classifying Chest Radiographs for Small Lung Opacities and Pleural Abnormalities in a Population Sample; *American Journal of Industrial Medicine* 1998; 34: 261-265
- [33] Katabami M, Dosaka-Akita H, Honma K, Saitoh Y, Kimura K, Uchida Y, Mikami H, Ohsaki Y, Kawakami Y, Kichuki K: Pneumoconiosis-related Lung Cancers - Preferential Occurrence from Diffuse Interstitial Fibrosis-type Pneumoconiosis; *Am J Crit Care Med* 2000; 162: 295-300
- [34] Kiesel J, Zutt M, Voitowitz RH, Voitowitz HJ: Follow-up asbestassoziierter Lungen- und Pleuraerkrankungen aus arbeitsmedizinischer Sicht; Dokumentationsband über die Verhandlungen der Deutschen Gesellschaft für Arbeitsmedizin und Umweltmedizin e.V. 1995; 243-247
- [35] Kjærgaard J, Andersson M: Incidence rates of malignant mesothelioma in Denmark and predicted future number of cases among men; *Scand J Work Environ Health* 2000; 26(2): 112-117

- [36] Konietzko N: Asbestinduzierte pleuropulmonale Erkrankungen; Handbuch der Arbeitsmedizin, Ecomed-Verlag, Landsberg, April 1998, 1-3
- [37] Koskinen K, Zitting A, Tossavainen A, Rinne JP, Roto P, Kivekas J, Reijula K, Huuskonen MS: Radiographic abnormalities among Finnish construction, shipyard and asbestos industry workers; *Scand J Work Environ Health* 1998; 24(2): 109-17
- [38] Kraus T, Raithel HJ, Hering KG: Evaluation and classification of high-resolution computed tomographic findings in patients with pneumoconiosis, *International Archives of Occupational and Environmental Health* 1996, 68: 249-254
- [39] Kraus T, Raithel HJ: Frühdiagnostik asbeststaubverursachter Erkrankungen; Hauptverband der Gewerblichen Berufsgenossenschaften 1998
- [40] Landis J, Koch GG: The measurement of Observer agreement for categorical data; *Biometrics* 1977; 33:159-174
- [41] Lange S: Radiologische Diagnostik der Lungenerkrankungen; Thieme, Stuttgart 1986
- [42] Levin SM, Kann PE, Lax MB: Medical Examination for Asbestos-Related Disease; *American Journal of Industrial Medicine* 2000; 37: 6-22
- [43] Liste der Berufskrankheiten; § 9 Abs.1 Siebtes Buch Sozialgesetzbuch (SGB VII)
- [44] Müller K-M, Krismann M: Asbestassoziierte Erkrankungen; *Deutsches Ärzteblatt* 93 1996 (27); 9, C387-391
- [45] Muir DCF, Bernholz CD, Morgan WKC: Classification of chest radiographs for pneumoconiosis: a comparison of two methods of reading; *British Journal of Industrial Medicine* 1992; 49: 869-871
- [46] Musch DC, Higgins ITT, Landis JR: Some Factors influencing interobserver variation in classifying simple pneumoconiosis; *British Journal of Industrial Medicine* 1985, 42: 346-349
- [47] Patz EF, Rusch VW, Heelan R: The Proposed New International TNM Staging System for Malignant Pleural Mesothelioma: Application to Imaging; *American Journal of Roentgenology* 1996; 166: 323-327
- [48] Popp W: Neue Erkenntnisse zur Synkanzerogenese bei exogen verursachten Tumoren; *Zbl. Hyg.* 198, 407-428 (1996)
- [49] Pschyrembel: *Klinisches Wörterbuch*, 1998: 258. Auflage; Walter de Gruyter, Berlin, New York 1998
- [50] Puntoni R, Merlo F, Borsa L, Reggiardo G, Garrone E, Ceppi M: A Historical Cohort Mortality Study Among Shipyard Workers in Genoa, Italy; *American Journal of Industrial Medicine* 2001; 40: 393-370

- 
- [51] Ribak J, Selikoff IJ: Survival of asbestos insulation workers with mesothelioma; British Journal of Industrial Medicine 1992; 49: 732-735
- [52] Rudd RM: New developments in asbestos-related pleural disease; Thorax 1996; 51: 210-216
- [53] Tiitola M, Kivisaari L, Zitting A: Computed tomography of asbestos-related pleural abnormalities; Int Arch Occup Environ Health 2002; 75: 224-228
- [54] Vogel S: Semiautomatische Segmentierung, quantitative Vermessung und Verlaufsdokumentation von Pleuramesotheliomen in Spiral-CT-Bildsequenzen; Lehrstuhl für Messtechnik und Bildverarbeitung an der RWTH Aachen, Juli 2003
- [55] Welch LS, Hunting KL, Balmes J, Bresnitz EA, Guidotti TL, Lockey JE, Myo-Lurin T: Variability in the Classification of Radiographs Using the 1980 International Labor Organization Classification for Pneumoconioses; Chest 1998; 114: 1740-1748
- [56] Wiethage T, Müller K-M: Rauchen, Radon, Asbest und Quarz; Atemw.-Lungenkrkh. 1999, Jahrgang 25, 6: 309-314
- [57] Weitowitz HJ, Beierl L, Rathgeb M, Schmidt K, Rodelsperger K, Greven U, Weitowitz RH, Lange HJ, Ulm K: Asbestos-Related Diseases in the Federal Republic of Germany; American Journal of Industrial Medicine 1981; 2: 71-78
- [58] Zentrale Erfassungsstelle asbeststaubgefährdeter Arbeitnehmer, Übersicht über den Datenbestand am 31.12.2003, bei der ZAs angefordert

# Lebenslauf Tanja Carl

<b>Persönliche Daten</b>	Name	Tanja Carl
	Geboren	05.12.1977 in Bonn-Bad Godesberg
	Familienstand	ledig
<b>Schulbildung</b>	1984-1997	Grundschule und Gymnasium, Bonn
	09/94-02/95	Luther College, Regina, Kanada
	Abschluss	Abitur
<b>Studium</b>	Humanmedizin	
	1997-2003	RWTH Aachen
	2003-2004	TU Dresden
	02-09/04	PJ in Glarus, Schweiz
	09/99	Physikum
	03/01	1. Abschnitt der Ärztlichen Prüfung
	09/03	2. Abschnitt der Ärztlichen Prüfung
	11/04	3. Abschnitt der Ärztlichen Prüfung