

Aus dem Institut für Medizinische Biometrie und Statistik
der Universität zu Lübeck
Direktor: Univ.- Prof. Dr. ret. nat. habil. Andreas Ziegler

Sprechen Ärzte, Apotheker und Juristen die gleiche Sprache?

**Eine empirische Untersuchung zum Gebrauch von
umgangssprachlichen Wahrscheinlichkeitsbegriffen.**

Inauguraldissertation
zur
Erlangung der Doktorwürde
der Universität zu Lübeck
-Aus der Sektion Medizin-

vorgelegt von
Steffi Mehlbeer, geb. Gockel
Schwerin

Lübeck 2012

1. Berichterstatter: Univ.- Prof. Dr. rer. nat. habil. Andreas Ziegler

2. Berichterstatter: Prof. Dr. rer. nat. Michael Bader

Tag der mündlichen Prüfung: 11.12.2012

Zum Druck genehmigt: Lübeck, den 11.12.2012

-Promotionskommission der Sektion Medizin-

INHALTSVERZEICHNIS

1	EINLEITUNG	1
1.1	ALLGEMEINE EINLEITUNG	1
1.2	HISTORISCHER ÜBERBLICK.....	2
1.3	FRAGESTELLUNG	4
2	MATERIAL UND METHODIK	6
2.1	AUSWAHL DER PROBANDEN	6
2.2	EMPIRISCHE ERHEBUNG.....	7
2.3	ERSTELLUNG DES ANSCHREIBENS UND DES FRAGEBOGENS.....	7
2.4	PRETEST.....	10
2.5	ABLAUF DER STUDIE.....	10
2.6	METHODIK ZUR BEWERTUNG DES RÜCKLAUFS	11
2.7	METHODIK ZUM ERSTELLEN UND VERGLEICH DER RANGREIHEN.....	12
2.8	METHODEN ZUM VERGLEICH DER KONTEXTFREIEN UND - ABHÄNGIGEN INTERPRETATION	18
2.9	VERGLEICH MIT DER OFFIZIELLEN KLASSIFIKATION FÜR ARZNEIMITTELNEBENWIRKUNGEN	23
3	ERGEBNISSE	25
3.1	RÜCKLAUF UND AUSSCHÖPFUNGSQUOTE	25
3.2	SOZIODEMOGRAPHISCHE DATEN	28
3.3	RANGREIHENFOLGEN UND -VERGLEICH	29
3.4	VERGLEICH VON WAHRSCHEINLICHKEITSBEGRIFFEN „KONTEXTFREI“ VERSUS „IM KONTEXT NEBENWIRKUNGEN“	36
3.5	WERDEN WAHRSCHEINLICHKEITSBEGRIFFE IM KONTEXT „RECHTSSTREIT“ ANDERS INTERPRETIERT?	52
3.6	VERGLEICH MIT DER OFFIZIELLEN KLASSIFIKATION FÜR ARZNEIMITTELNEBENWIRKUNGEN	55
4	DISKUSSION	64
4.1	WIE GUT UND RELEVANT IST DIESE STUDIE?.....	64
4.2	GIBT ES BERUFSGRUPPENSPEZIFISCHE RANGREIHEN DER WAHRSCHEINLICHKEITSBEGRIFFE UND UNTERSCHIEDEN SICH DIESE VONEINANDER?	65
4.3	BESTEHT EINE KONTEXTABHÄNGIGKEIT DER WAHRSCHEINLICHKEITSBEGRIFFE UND GIBT ES EINEN UNTERSCHIED ZWISCHEN DEN BERUFSGRUPPEN?	67
4.4	KENNEN ÄRZTE, APOTHEKER UND JURISTEN DIE RICHTLINIE FÜR „HÄUFIGKEITSANGABEN VON NEBENWIRKUNGEN“?	70
5	ZUSAMMENFASSUNG	73
6	LITERATURVERZEICHNIS	75
7	ANHANG	78
7.1	SPSS- SYNTAX ZUM BOOTSTRAP.....	78
7.2	FRAGEBOGEN, ANSCHREIBEN UND ERINNERUNGSSCHREIBEN	79
7.3	RÜCKLAUFDAUER	85
7.4	ALTERSVERTEILUNG NACH BERUFSGRUPPEN UNTERTEILT	87

7.5	GRAPHISCHE DARSTELLUNG DER NUMERISCHEN INTERPRETATION DER 20 WAHRSCHEINLICHKEITSBEGRIFFE; ALLE TEILNEHMER ZUSAMMENGEFASST	87
7.6	VERGLEICH DER DREI NEUEN RANGREIHEN MIT DER IDEALEN RANGREIHE DER LITERATUR	91
7.7	GRAPHISCHE GEGENÜBERSTELLUNG DER KONTEXTFREIEN UND DER KONTEXTABHÄNGIGEN INTERPRETATION VON „MÖGLICHERWEISE“, „OFT“ UND „WAHRSCHEINLICH“	93
8	DANKSAGUNG.....	100
9	LEBENS LAUF.....	101

1 Einleitung

1.1 Allgemeine Einleitung

Fragen Sie nach, wenn Ihr Arzt Ihnen sagt, dass Ihr neues Medikament nur „selten“ schwere Nebenwirkungen verursachen kann? Fragen Sie nach, wie viel genau „selten“ bedeutet?

In den letzten Jahren und Jahrzehnten wurde zunehmend realisiert, dass Patienten informiert sein wollen und müssen, um fundierte Entscheidungen bezüglich ihrer weiteren Behandlungsmöglichkeiten zu treffen. Schon lange entscheidet nicht mehr der Arzt allein über die weitere Behandlung. Bei nahezu unbegrenzten technischen Möglichkeiten, mit schneller Informationsbeschaffung via Internet, kann der Patient noch in der Arztpraxis mit dem Mobiltelefon alle wichtigen Informationen aus dem Internet herunterladen. Dennoch bleibt das persönliche Gespräch der wichtigste Schritt zum informierten Patienten. Doch verstehen wir einander?

Wenn der Apotheker in einem Rechtsstreit über „häufige“ Nebenwirkungen berichtet, fragen Sie dann nach, was genau „häufig“ heißt? Oder wenn der Anwalt argumentiert, dies seien Risiken, die nun mal „gelegentlich“ auftreten würden. Fragen Sie dann nach?

Fragen Sie nach!

1.2 Historischer Überblick

Seit Beginn der 60iger Jahre beschäftigten sich zahlreiche Studien mit der Frage, wie Wahrscheinlichkeiten und Risiken am effektivsten dargestellt werden können.

Simpson [1] war einer der ersten Autoren, der sich über viele Jahre mit diesem Thema beschäftigte. In seinen Studien stellte er fest, dass die numerische Interpretation von einigen Wahrscheinlichkeitsbegriffen wie „always“, „very often“ oder „seldom“ über 20 Jahre in ihren Mittelwerten stabil waren. Allerdings fiel schon hier die extreme Spannweite der Interpretationen auf.

In den folgenden Jahren wurden verschiedenste Personengruppen bezüglich der Darstellung, des Verständnisses und der Interpretation von Wahrscheinlichkeiten untersucht. So befragte *Simpson* [1] zunächst Studenten. Im Verlauf verschob sich das Interesse aber immer mehr in Richtung Medizin. Man stellte fest, dass der Weg zum informierten Patienten nur über ein gelungenes Arzt-Patienten-Gespräch führt, in dem sich die Gesprächspartner uneingeschränkt verstehen. So wurden Ärzte, Patienten, Krankenschwestern, Eltern und auch Schwangere als Probanden ausgewählt [1-7].

Zunächst stellte sich die Frage, in welcher Form die Wahrscheinlichkeiten darzustellen seien und welche Vorlieben es gäbe?

Mazur und Hickam [8] und auch *Ohnishi et al* [9] stellten in ihren Studien fest, dass viele Patienten eine verbale Darstellung von Wahrscheinlichkeiten wie zum Beispiel „häufig“, „nie“ oder „selten“ bevorzugen. Einige Patienten hätten sogar Schwierigkeiten, Risiken mit Hilfe von Prozentangaben richtig einzuschätzen [8, 9]. Im Gegensatz hierzu zogen in anderen Studien Patienten oder Eltern von Patienten eine numerische Wahrscheinlichkeitsdarstellung vor [10, 11].

Bezüglich der Mediziner wurden ähnliche ambivalente Ergebnisse erzielt. Die Arbeitsgruppe um *Merz* [4] stellte beispielsweise fest, dass die Mediziner ihrer Studie (1991) verbale Wahrscheinlichkeitsbegriffe bevorzugten. So sei es primär natürlicher in Worten zu sprechen, als Zahlen zu verwenden. Es bestehe der Wunsch, dem Patienten keine Genauigkeiten vorzutäuschen und oftmals fehlten auch detaillierte numerische Wahrscheinlichkeiten oder würden je nach Lehrmeinung unterschiedlich dargestellt [4].

Eine einfache Lösung war somit nicht zu finden und weitere Studien waren nötig. Wie schon in *Simpsons* Studie [1], so ließen sich auch in zahlreichen weiteren Studien die Spannweite der numerischen Interpretation von verbalen Wahrscheinlichkeitsbegriffen belegen. Die Abstände zwischen Minimal- und Maximalwert waren dabei sehr groß und damit sehr ungenau [1-4, 6, 9, 11-14]. Missverständnisse im Arzt-Patienten-Gespräch seien vorprogrammiert, wenn verbale Wahrscheinlichkeitsbegriffe für die Darstellung von Risiken und Nebenwirkungen gebraucht werden [15, 16]. Einige Autoren wie z.B. *Nakao und Axelrod* [7], *Budescu und Wallsten* [16] und *Hay et al* [17] gingen sogar so weit, Empfehlungen auszusprechen, Wahrscheinlichkeiten nur noch numerisch darzustellen. Doch dass dies alleine nicht möglich sein würde, war auch vielen Autoren

klar. Somit stellte sich die Frage, welche Faktoren die numerische Interpretation von Wahrscheinlichkeiten beeinflussen.

Zunächst konnte gezeigt werden, dass das Alter, die ethnische Zugehörigkeit, die Vorerkrankungen und die Erfahrungen der Probanden einen Einfluss auf die Ergebnisse hatten [18, 19]. Des Weiteren postulierten *Pepper und Prytulak* [20] bereits 1974, der Kontext würde die Interpretation der Wahrscheinlichkeitsbegriffe beeinflussen. Dies bewiesen sie anschaulich anhand von allgemeinen Ereignissen („selten“ wie ein Erdbeben versus „häufig“ wie Schüsse in einem Western).

Weitere Studien bewiesen, dass im Vergleich zur kontextfreien Interpretation verbale Wahrscheinlichkeitsbegriffe im medizinischen Kontext mit geringeren Prozentwerten belegt werden. Insbesondere die Schwere einer Erkrankung hatte einen starken Einfluss auf die Interpretation [9, 18]. Dies bestärkte diejenigen Autoren, welche keinen Platz für umgangssprachliche Wahrscheinlichkeitsbegriffe in der Medizin sehen [7, 17]. Wiederum wurde der Ruf nach einer einheitlichen Kodierung laut [15, 21, 22].

1990 erstellten *Mosteller und Youtz* [21] eine Liste von Wahrscheinlichkeitsbegriffen, die als Grundlage einer internationalen Kodierung dienen sollte. Allerdings enthielt diese 52 Ausdrücke. Im Weiteren wurde nach möglichen Rangreihen von Wahrscheinlichkeitsausdrücken gesucht [4, 16, 23, 24]. Es konnte gezeigt werden, dass intraindividuell der Gebrauch von verbalen Wahrscheinlichkeiten sehr stabil ist. Für Ärzte konnte eine Rangreihe von 20 Begriffen erstellt werden [24]. Diese Konstanz im Gebrauch der Wahrscheinlichkeitsbegriffe ließ darauf hoffen, eine Richtlinie für den Gebrauch verbaler Wahrscheinlichkeiten erstellen zu können.

Eine standardisierte Kodierung von Wahrscheinlichkeitsbegriffen wurde vom European Commission (EC) Pharmaceutical Committee erarbeitet und in the Summary of Product Characteristics (SPC) 1999 eingearbeitet [25]. Das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) hat auf dieser Basis 2002 die Richtlinie „Häufigkeitsangaben von Nebenwirkungen“ herausgegeben [26]. Hierin wurden die fünf Begriffe „sehr häufig“, „häufig“, „gelegentlich“, „selten“ und „sehr selten“ in Zusammenhang mit Nebenwirkungen klar definiert (siehe Tabelle 3).

Mittlerweile wurde diese Richtlinie schon mehrfach überarbeitet, um die Grenzbereiche uneindeutig voneinander abzugrenzen. Letztmalig erfolgte die Anpassung der Angaben in Deutschland 2007.

Eine Arbeitsgruppe um *Berry und Knapp* untersuchte in mehreren Experimenten/Studien diese SPC-Richtlinie [27-30]. Sie konnten beweisen, dass Patienten diese fünf Begriffe eindeutig überinterpretieren [27, 29, 31]. Die wirklichen Risiken würden falsch eingeschätzt und Medikamente würden nicht genommen, weil Patienten die verbalen Wahrscheinlichkeiten überinterpretieren [29]. *Moore et al* [32] stellten fest, dass insbesondere seltene Ereignisse/Risiken überinterpretiert und alltägliche Risiken/Ereignisse unterschätzt werden.

Ein weiterer interessanter Lösungsansatz war und ist die graphische Darstellung von Wahrscheinlichkeiten [2, 33-36]. Als Beispiel sei hier *Paling* [34] dargestellt. Er entwickelte eigene Skalen und Bilder mit deren Hilfe Wahrscheinlichkeiten veranschaulicht werden können. Außerdem gab er nach jahrelanger Forschung gezielte Empfehlungen ab, wie Ärzte Risiken klarer darstellen könnten. Neben der Empfehlung möglichst visuelle Skalen zu benutzen, stellt auch er fest, dass Patienten die Begriffe der EU- Richtlinie nicht richtig interpretieren. Weitere Studien sind somit nötig, um die Richtlinie und damit die Kommunikation weiter zu optimieren. Oder muss die SPC-Richtlinie wohlmöglich wieder fallengelassen werden?

1.3 Fragestellung

Nicht nur in der Medizin treten Probleme in der Kommunikation auf. Auch in Partnerschaft, Ehe und Familie können Missverständnisse entstehen, wenn verbale Wahrscheinlichkeiten falsch interpretiert werden. Ziehen diese vielleicht kleinere Diskussionen nach sich, so können Missverständnisse in der Politik oder der Rechtsprechung weitreichende Folgen haben. Ein Patient, der sich falsch beraten fühlt, wird eventuell den Arzt wechseln. Tritt eine Komplikation auf, von deren Häufigkeit er nichts wusste bzw. sie anders einschätzte, wird er gegebenenfalls sogar eine Klage einreichen.

Die SPC-Richtlinie für „Häufigkeitsangaben von Nebenwirkungen“ (siehe Tabelle 3) wird mittlerweile regelhaft in jedem Beipackzettel mit abgebildet und erläutert. Sie ist eine standardisierte Kodierung der fünf Wahrscheinlichkeitsbegriffe. Von Patienten werden diese fünf Begriffe allerdings eindeutig überinterpretiert. Wie ist dies bei Ärzten und Apothekern? Kennen sie diese Richtlinie? Diese Fragen sollen in dieser Studie bearbeitet werden. Außerdem wird geprüft, ob auch andere Berufsgruppen, die seltener mit medizinischen Fragestellungen konfrontiert werden, die Richtlinie kennen. Kennen zum Beispiel Juristen diese Leitlinie?

Zunächst wird in dieser Arbeit der Ansatz der Rangreihen nochmal aufgegriffen. Es wird geprüft, ob es auch für Apotheker und Juristen Rangreihen der Wahrscheinlichkeitsbegriffe gibt. Anschließend soll ermittelt werden, inwieweit der Kontext „Nebenwirkungen“ die Interpretation von verbalen Wahrscheinlichkeitsausdrücken beeinflusst. Dies ist nötig, da bereits gezeigt wurde, dass auch der Kontext selbst Einfluss auf die Interpretation hat [9]. Welchen Einfluss hat also der Kontext „Nebenwirkung“?

Zum Schluss soll geklärt werden, ob die verschiedenen Berufsgruppen die Richtlinie für „Häufigkeitsangaben von Nebenwirkungen“ kennen und richtig verwenden. Gibt es ähnliche Schwierigkeiten wie bei den Patienten in der Interpretation der fünf Wahrscheinlichkeitsbegriffe?

Die zentralen Fragen dieser Arbeit lauten somit:

1. Gibt es berufsgruppenspezifische Rangreihen der Wahrscheinlichkeitsbegriffe und unterscheiden sich diese voneinander?
2. Besteht eine Kontextabhängigkeit in Zusammenhang mit Nebenwirkungen und gibt es einen Unterschied zwischen den Berufsgruppen?
3. Kennen und benutzen diese Berufsgruppen die SPC-Richtlinie für „Häufigkeitsangaben bei Nebenwirkungen“?

Nachfolgend werden die Probanden und der Fragebogen vorgestellt und der Ablauf der Studie beschrieben. Der Fragebogen wurde in vier Blöcke bzw. Abschnitte eingeteilt. Die Methodik wird in Anlehnung an die Abschnitte des Fragebogens abschnittsweise dargestellt. Die Darstellung der Ergebnisse, die Diskussion und die Zusammenfassung folgen anschließend.

2 Material und Methodik

Im folgenden Abschnitt soll die Methodik dieser empirischen Studie beschrieben werden. Es werden die Probanden und der Fragebogen vorgestellt und der Ablauf der Studie beschrieben. Anschließend wird die Methodik zur Berechnung der einzelnen Fragestellungen erläutert.

2.1 Auswahl der Probanden

Ein bekannter Werbeslogan, verankert im Heilmittelwerbegesetz §4 Abschnitt 3, lautet: Zu Risiken und Nebenwirkungen fragen Sie Ihren Arzt oder Apotheker. Schaut man nur zwei Stunden am Abend fern, so bekommt man diese Aufforderung mindestens einmal zu hören. Dies wurde zum Anlass genommen diese beiden Berufsgruppen als Probanden auszuwählen. Es sollte geprüft werden, wie Wahrscheinlichkeiten von Ärzten und Apotheker dargestellt werden.

Hierbei wurde sich in dieser Studie auf das Bundesland Schleswig- Holstein beschränkt. Da das Institut für Medizinische Biometrie und Statistik der Universität zu Lübeck diese Studie betreute, lag die Wahl des Landes nahe. Eine Vergleichbarkeit mit den norddeutschen Bundesländern war zu unterstellen. Ob ein Unterschied im Vergleich zwischen norddeutschen und süddeutschen Ärzten bestehe, war eine Fragestellung einer parallel laufenden Studie.

Die Adressen der Ärzte waren auf der Homepage der Ärztekammer Schleswig- Holstein als PDF-Datei herunterzuladen. Diese wurden in eine Excel-Datei überführt. Mittels Zufallsprinzip wurden die Ärzte ausgewählt und angeschrieben. Ähnlich wurde mit den Apothekern verfahren. Die schleswig-holsteinische Apothekerkammer überließ der Autorin freundlicherweise eine CD-ROM mit Adressen in Form einer Excel-Datei.

Eine weitere Überlegung war, eine Berufsgruppe auszuwählen, die nicht unmittelbaren Patientenkontakt habe. Trotzdem sollten sich die Teilnehmer mit Wahrscheinlichkeitsbegriffen beschäftigen. Dies trifft zum Beispiel für Juristen zu, die sich mit Arzthaftungsfragen und Behandlungsfehler auseinandersetzen. Somit wurden die Juristen als dritte Berufsgruppe ausgewählt.

Die Adressen der Rechtsanwälte und Richter wurden über die Schleswig-Holsteinische Rechtsanwaltskammer und über das Handbuch der Juristen ermittelt. Wiederum wurden die Adressen in eine Excel-Datei überführt.

Aus allen Adressen wurden die Probanden zufällig per Computer ausgewählt. Insgesamt wurden 300 Ärzte, 200 Apotheker und 200 Juristen aus Schleswig-Holstein angeschrieben.

2.2 Empirische Erhebung

Im Weiteren stellte sich die Frage, welches Erhebungsverfahren für diese Studie das Richtige sei.

Allgemein standen die Primär- und die Sekundärforschung als Methoden der Meinungsforschung zur Auswahl. Primärforschung bedeutet, dass Daten speziell für die Studie erhoben werden. Bei der Sekundärforschung wird mit Daten aus bereits bestehenden Erhebungen gearbeitet [37]. Da keine Daten über Wahrscheinlichkeitsbegriffe dieser Berufsgruppen vorlagen, wurde sich in dieser Arbeit für die Primärforschung entschieden.

Grundsätzlich stehen die Befragung und die Beobachtung als Methoden der Meinungsforschung zur Verfügung. In dieser Studie bot sich die schriftliche Befragung an.

Vorteile sind die Vergleichbarkeit der Daten und damit eine standardisierte Auswertung. Das Antwortverhalten ist als neutral anzusehen, da die Qualität des Interviewers in den Hintergrund tritt. Somit können weder positive noch negative Einflüsse auf den Befragten einwirken [37]. Des Weiteren ist der geringere Zeit- und Kostenfaktor gegenüber dem persönlichen oder telefonischen Interview vorteilhaft.

Nachteil einer schriftlichen Befragung ist oftmals eine geringe Rücklaufquote. Vielfach werden nur Quoten von 10-20 % erreicht, 20-40 % gelten als zufrieden stellend. Werte ab 40 % werden als hervorragend betrachtet [37].

Insgesamt ist festzustellen, dass die Qualität der Untersuchungsergebnisse und die Rücklaufquote entscheidend vom Fragebogen und dessen Gestaltung abhängen [37, 38].

2.3 Erstellung des Anschreibens und des Fragebogens

2.3.1 Anschreiben

In einem freundlichen, aber auch zum Nachdenken anregenden Anschreiben wurden die Teilnehmer gebeten, den Fragebogen auszufüllen. Es wurden der Aufbau und die Ziele der Studie, sowie das Institut und die Kontaktperson vorgestellt. Außerdem wurde auf die Vertraulichkeit der Daten hingewiesen [38].

Das Anschreiben wurde auf dem offiziellen Briefpapier der Universität zu Lübeck gedruckt. Dabei wurde auf das aktuelle Datum und die namentliche Anrede im Anschreiben geachtet. Dies entspricht den Empfehlungen von Dillman [39].

Das Anschreiben wurde eigenhändig mit blauer Tinte unterschrieben und ein adressierter und frankierter Rückumschlag beigelegt. Dies sollte das Zurücksenden vereinfachen und somit die Rücklaufquote erhöhen. Ein Exemplar des Anschreibens ist im Anhang hinterlegt.

2.3.2 Fragebogen

Der Fragebogen wurde unter Berücksichtigung des international als Standard geltenden Verfahrens von Dillmann erstellt [38, 39]. Es wurden vier thematische Blöcke gebildet. Mithilfe von Überleitungssätzen wurden die jeweiligen Abschnitte eingeleitet, um die Probanden auf die neuen Themen vorzubereiten.

Da die Aufmerksamkeitskurve bei einer schriftlichen Befragung deutlich zum Ende abflacht, wurden die wichtigen Themenblöcke zu Beginn erhoben. Die soziodemographischen Daten wurden deshalb an das Ende des Fragebogens gestellt [38]. Es wurde nach Geschlecht, Alter, Fachrichtung und Einrichtung, in der die Befragten arbeiten, gefragt. Außerdem sollten Angaben über den Ort des Aufwachsens, des Studierens, den Ort des Examensabschlusses und des derzeitigen Arbeitens gemacht werden.

Inhaltlich war der Fragebogen wie nachfolgend beschrieben eingeteilt: Der Abschnitt A enthielt eine Liste mit verbalen Wahrscheinlichkeitsausdrücken. Aufgabe war es, diese Begriffe numerisch zu interpretieren. So sollten die Probanden beispielsweise die Begriffe „selten“, „oft“ und „nie“ mit einer frei wählbaren Prozentzahl interpretieren.

Im Abschnitt B wurde ein Fallbeispiel konstruiert. Die Probanden sollten sich in einen Arzt hineinversetzen, der die drei Wahrscheinlichkeitsbegriffe „häufig“, „gelegentlich“ und „selten“ im Kontext mit Nebenwirkungen als Prozentangaben darstellen sollte. Diese drei Begriffe sind in Zusammenhang mit Nebenwirkungen klar von der Pharmaindustrie definiert.

Umgekehrt wurde im Abschnitt C verfahren. In aufsteigender Reihenfolge von niedrig zu hoch wurden numerischen Wahrscheinlichkeiten wie 1:1000 oder 1:100 angeboten. Diese sollten mit frei wählbaren Begriffen („selten“, „nie“, usw.) definiert werden. Auch dieses Fallbeispiel bezog sich auf Nebenwirkungen bzw. Komplikationen.

Das Fallbeispiel des Abschnitts D wurde offener gestaltet, d.h. unabhängig von ‚Nebenwirkungen‘. Es wurde ein Rechtsstreit konstruiert. Hier sollte den Begriffen „möglicherweise“, „oft“, und „wahrscheinlich“ abermals eine Prozentzahl zugeordnet werden.

Die Wahrscheinlichkeitsbegriffe, die in diesem Fragebogen verwendet wurden, entstammen einer Studie im Rahmen einer Dissertation [24]. Dort wurde eine optimale Reihe aus 20 deutschen Wahrscheinlichkeitsbegriffen erstellt. Es konnte bewiesen werden, dass die Reihenfolge dieser Begriffe unter Ärzten als konstant angesehen werden kann. Somit existiert eine Rangreihe der Wahrscheinlichkeitsausdrücke bei den Ärzten.

In einem Vergleich dieser mit einer idealen Reihe aus der englischsprachigen Literatur, zeigte sich ein hohes Maß an Zusammenhang, so dass von einer Vergleichbarkeit zwischen den Rangreihen zu sprechen ist [24].

In dieser Studie wurde die ‚optimale‘ deutschsprachige Reihe der Suß-Dissertation [24] als ideale Reihe der deutschen Wahrscheinlichkeitsbegriffe (siehe Tabelle 1) angesehen und damit als Grundlage des Abschnitts A verwendet.

Tabelle 1: Die ideale Reihe der Wahrscheinlichkeitsbegriffe ist der deutschsprachigen Literatur entnommen. Sie stellt eine Reihenfolge dar. Die Rangreihenfolge wurde von Ärzten erstellt, wobei die Begriffe aufsteigend gelistet sind.

Rangplatz	ideale Reihe
1.	niemals
2.	selten
3.	nie
4.	gelegentlich
5.	unsicher
6.	denkbar
7.	fraglich
8.	möglicherweise
9.	unklar
10.	oft
11.	wahrscheinlich
12.	häufig
13.	überwiegend
14.	meist
15.	normalerweise
16.	typischerweise
17.	fraglos
18.	stets
19.	sicher
20.	immer

Der Fragebogen wurde insgesamt 1000 Mal verschickt. Er wurde so erstellt, dass er zur Bearbeitung von zwei zueinander parallel laufenden Studien genutzt werden konnte. Dies geschah aus Kostengründen. Für diese Dissertation kamen im Rahmen der Fragestellung 700 der 1000 verschickten Exemplare zur Auswertung infrage. Die Fragebögen wurden in zwei Durchläufen mit jeweils 350 Exemplaren verschickt. Sie waren mit einem fortlaufenden Nummerncode versehen, um bei Nichtantworten gezielt Erinnerungsschreiben verschicken zu können. Dem Erinnerungsschreiben wurde wiederum ein Fragebogen beigelegt.

Bei sehr gutem Rücklauf mussten keine Veränderungen vorgenommen werden, so dass in beiden Durchläufen der gleiche Fragebogen verwendet wurde. Ein Exemplar ist im Anhang hinterlegt.

2.3.3 Erinnerungsschreiben

Um den Rücklauf zu erhöhen, wurde ein Erinnerungsschreiben erstellt. Die Probanden wurden hierin nochmals auf die Wichtigkeit ihrer Teilnahme hingewiesen. Auch dieses Anschreiben wurde eigenhändig mit blauer Tinte unterschrieben und mit einem Fragebogen und frankiertem Rückumschlag verschickt.

Im ersten Durchlauf wurde das Erinnerungsschreiben nach drei und im zweiten Durchlauf nach fünf Wochen verschickt. Dieser Zeitraum wurde gewählt, um den Probanden Zeit für die Rücksendung zu lassen und damit unnötige Doppelsendungen und Kosten zu vermeiden.

2.4 Pretest

Es blieb nun zu prüfen, ob der Fragebogen verständlich war oder ob es Probleme beim Beantworten gab. Hierfür wurde zu Beginn der Studie ein Pretest durchgeführt. Dies ist bei großen empirischen Studien Voraussetzung für eine optimale Gestaltung des Hauptfeldes [40]. Der Pretest dient zur Bestimmung der durchschnittlichen Befragungszeit, prüft die Verständlichkeit der Fragen und gibt letztlich Hinweise auf die zu erwartende Rücklaufquote des Hauptfeldes [41].

Insgesamt 1000 Fragebögen sollten im Rahmen der zwei zueinander parallel laufenden Studien verschickt werden. Zuvor wurden 20 Exemplare im Freundes- und Familienkreis der Studienleiter verteilt. Unter den Testpersonen befanden sich Ärzte, Medizin- und Jurastudenten sowie Familienangehörige ohne medizinische Ausbildung. Dieser vereinfachte Pretest diente vorrangig dazu Verständnisprobleme aufzudecken. Der organisatorische Ablauf wurde nicht extra getestet. Allerdings wurde im Vorfeld entschieden zunächst nur die Hälfte der Probanden anzuschreiben und ggf. bei größeren Problemen Änderungen im nächsten Durchlauf vorzunehmen. Aufgrund der Ergebnisse des ersten Durchlaufs und des guten Rücklaufs mussten aber keine Korrekturen durchgeführt werden.

2.5 Ablauf der Studie

Die Auswahl der Adressen erfolgte per Zufallsprinzip, elektronisch gesteuert. Im Rahmen eines Serienbriefs wurden somit die Ärzte, Apotheker und Juristen angeschrieben. Der Brief bestand wie bereits beschrieben aus einem Anschreiben, einem Fragebogen und einem frankierten Rückumschlag.

Im August 2004 wurde der erste Durchlauf verschickt. Zur erneuten Testung des Rücklaufs wurden zunächst nur 350 Exemplare versandt, um nötigenfalls bei schlechtem Antwortverhalten den Fragebogen zu überarbeiten. Dies war im Verlauf aber nicht nötig.

Das Erinnerungsschreiben wurde drei Wochen nach Aussendung des ersten Durchlaufs verschickt. Es wurden gezielt die Personen angeschrieben, die bis zu diesem Zeitpunkt nicht geantwortet hatten. Dies war aufgrund einer Kodierung der einzelnen Fragebögen möglich. Diese Kodierung wurde zur Berechnung bis auf die Berufsgruppencodes aufgehoben, sodass die Bögen anonym ausgewertet werden konnten.

Ein zweiter Durchlauf wurde sechs Wochen nach Beginn des Ersten mit weiteren 350 Exemplaren durchgeführt. Das zugehörige Erinnerungsschreiben wurde hier fünf Wochen nach dem Anschreiben versandt. Dieses verlängerte Intervall zwischen Verschicken des Fragebogens und des Erinnerungsschreibens beruht auf der Vermutung, der Ferienbeginn könne zu einem geringeren Rücklauf führen [39].

Die zurückgesandten Fragebögen wurden mit dem Ankunftsdatum versehen und alle ausgefüllten Exemplare anschließend digitalisiert. Zur Überprüfung der Richtigkeit dieser neu erstellten Tabelle, wurden alle Daten durch eine unabhängige Person des Instituts für medizinische Biometrie und Statistik erneut eingegeben und miteinander verglichen.

Des Weiteren wurde eine Plausibilitätskontrolle durchgeführt. Alle Fragebögen, die Verständnisprobleme erkennen ließen, wurden aussortiert. Außerdem wurden alle Fragebögen ausgeschlossen, die unvollständig ausgefüllt waren.

Zur Auswertung der Daten wurde das Programm PASW Statistics (SPSS für Windows) in der Version 12.8/18 verwendet. Bei der Digitalisierung der Fragebögen wurden fehlende Werte mit einem Wert (missing value) versehen. Das SPSS- Programm rechnet nur mit „complete cases“, sodass alle unvollständigen Fragebögen nicht in die Berechnung eingingen.

Zur Berechnung des Bootstraps in SPSS wurde eine spezielle Syntax verwendet, die im Anhang aufgeführt ist.

Des Weiteren wurde zur Berechnung des Kruskal- Wallis-Test und des Hodges-Lehmann- Punktschätzer für den Median der Mediandifferenzen das Programm SAS/StatXact 8 PROCs verwendet.

2.6 Methodik zur Bewertung des Rücklaufs

Um die Qualität einer Umfrage zu beurteilen wurde die Ausschöpfungsquote nach PORST herangezogen [40]. Sie errechnet sich folgendermaßen:

Tabelle 2: Berechnung der Ausschöpfungsquote

Bruttostichprobe	- stichprobenneutrale Ausfälle	= bereinigte Stichprobe
Bereinigte Stichprobe	- nicht-stichprobenneutrale Ausfälle	= Ausschöpfung
Auschöpfung/ bereinigte Stichprobe	* 100	= Ausschöpfungsquote

Es sei die Bruttostichprobe die Anzahl aller verschickten Fragebögen. Stichprobenneutrale Ausfälle ergeben sich, wenn der Brief zurückgesendet wird oder der Empfänger unbekannt verzogen ist, d.h. wenn Fehler in der Adressenliste vorliegen. Nichtstichprobenneutrale Ausfälle entstehen, wenn der Empfänger die Teilnahme verweigert [40].

Die Ausschöpfungsquote ist der Quotient aus Ausschöpfung, d.h. aus allen verwertbaren Fragebögen und der bereinigten Stichprobe. In einem weiteren Schritt wird dies mit 100 multipliziert [40].

Die Frage nach einer Mindestausschöpfungsquote wird in der Literatur kontrovers betrachtet. Werte ab 50% werden als tolerierbare Ausschöpfungsquote angesehen, während Raten ab 80% als unrealistisch eingeordnet werden [40].

2.7 Methodik zum Erstellen und Vergleich der Rangreihen

Im ersten Abschnitt des Fragebogens sollten 20 Wahrscheinlichkeitsbegriffe von den Probanden mit Prozentwerten belegt werden. Diese 20 Begriffe sind Bestandteile der idealen Rangreihe der deutschsprachigen Literatur. Im Folgenden werden die Fragestellung und die Methodik zur Bearbeitung der Fragestellung beschrieben.

2.7.1 Fragestellung und Methodenübersicht

Die Ergebnisse der 20 Wahrscheinlichkeitsbegriffe aller Studienteilnehmer werden zunächst rein deskriptiv dargestellt.

Dann wird untersucht, wie verbale Wahrscheinlichkeitsausdrücke gebraucht werden. Dabei sind nicht die absoluten Werte von Interesse, sondern die Rangreihe der Begriffe. Für jede Person wird eine Reihe der 20 Begriffe gebildet. Die einzelnen Personen werden den drei Berufen Ärzte, Apotheker und Juristen zugeordnet.

Die wichtigen Fragen dieses Abschnitts lauten:

1. Existiert für jede Berufsgruppe eine Rangreihenfolge der Wahrscheinlichkeitsausdrücke?
2. Stimmen die ermittelten Rangreihenfolgen mit der ‚idealen‘ Rangreihenfolge aus der Literatur überein?

Die erste Frage wird mithilfe des Kendall'schen Konkordanzkoeffizienten W bearbeitet. Anschließend wird mit der χ^2 -Verteilung die Signifikanz von Kendall W geprüft und mit einem Bootstrap-Verfahren das Konfidenzintervall für Kendall W bestimmt.

Die zweite Frage soll mithilfe der Rangkorrelationskoeffizienten Kendall τ (tau) und Spearman's ρ (Rho) bearbeitet werden. Zur statistischen Überprüfung wird ein Permutationstest verwendet.

Nachfolgend werden die einzelnen Methoden vorgestellt.

2.7.2 Methodik zum Erstellen der Rangreihen

2.7.2.1 Konkordanzkoeffizient Kendall W

Der Konkordanzkoeffizient Kendall W ist ein Maß für den Zusammenhang zwischen mehreren (k) Rangreihen von (N) Datenträgern [42].

Die folgende Formel dient der Berechnung Kendall W's:

$$W = \frac{12S}{k^2 \cdot (N^3 - N)}$$

W = Konkordanzkoeffizient

S = Quadratsumme der Rangsummen

N = Anzahl der in einer Rangreihe angeordneten Einheiten

k = Anzahl der Rangreihen

In dieser Studie sind N die 20 verwendeten Wahrscheinlichkeitsausdrücke und k die Anzahl der Rangreihen der teilnehmenden Personen.

Der Konkordanzkoeffizient wird folglich verwendet, um zu prüfen, ob und wie sehr sich die einzelnen Rangreihen ähneln. Kendall W kann dabei Werte zwischen 0 und 1 annehmen, wobei $W = 1$ eine vollständige Übereinstimmung der einzelnen Rangreihen bedeutet. Beträgt $W = 0$ ist eine perfekte Diskrepanz anzunehmen, d.h. zwischen den gebildeten Rangreihen gäbe es keinen Zusammenhang [43]. Ähneln sich die Rangreihen der einzelnen Personen, so sind hohe Werte von W zu erwarten, d.h. je größer der Wert von W ist, desto größer ist die Ähnlichkeit der Rangreihen [44].

Hieraus ergeben sich die zu überprüfenden Hypothesen:

$H_0: W = 0$ Es besteht kein Zusammenhang zwischen den Rangreihen.

$H_1: W > 0$ Es besteht ein Zusammenhang zwischen den Rangreihen.

Mithilfe des Programms SPSS/ Windows wird aus den gegebenen Daten für jede einzelne Person eine Rangfolge der 20 Begriffe erstellt und der Konkordanzkoeffizient nach Kendall für jede Berufsgruppe berechnet. Gleichzeitig bildet SPSS für jede Berufsgruppe eine neue berufsspezifische Rangreihe der 20 Begriffe, so dass insgesamt drei neue „echte“ Rangreihen entstehen [42].

2.7.2.2 Signifikanz von Kendall W

Die Signifikanz gibt an, mit welcher Wahrscheinlichkeit die Ergebnisse einer Stichprobe (Teilmenge) repräsentativ für die Grundgesamtheit sind, d.h. mit welcher Wahrscheinlichkeit auf die Grundgesamtheit geschlossen werden kann, ohne größere Fehler zu begehen [37].

Die Signifikanz von Kendall W für kleine Stichproben mit Werten von $N < 7$ und für $k < 20$ ist tabelliert. Für Stichproben mit größerem Umfang ($N > 7$) benutzt man die χ^2 -Verteilung. Hierbei ist W asymptotisch χ^2 verteilt [42] bei $Fg = N - 1$.

Die Berechnung erfolgt mit der Formel:

$$X^2 = k(N - 1) \times W$$

X^2 = Chi – Quadrat

k = Anzahl der Rangreihen

N = Anzahl der Begriffe

Wenn nun der errechnete Wert von X^2 für ein bestimmtes Signifikanzniveau und einen bestimmten Freiheitsgrad von $Fg = N - 1$ gleich dem Wert in einer spezifischen Tafel oder größer als dieser ist, so kann die Nullhypothese auf diesem Signifikanzniveau zurückgewiesen werden [42, 44].

Hohe oder signifikante Werte von W bedeuten, dass die Probanden bei der Bildung der Rangreihen dieselben Standards verwenden [42].

2.7.2.3 Bootstrap zur Bestimmung des Konfidenzintervalls

Wieweit die Werte aus dieser Stichprobe von den „wahren“ Werten der Grundgesamtheit entfernt liegen, kann nicht sicher abgeschätzt werden. Somit soll ein Konfidenzintervall berechnet werden. Für Kendall W existiert jedoch kein approximatives Konfidenzintervall, sodass nur mithilfe eines Bootstrap- oder Jackknife-Verfahrens ein Vertrauensintervall bestimmt werden kann.

Hier wird der Bootstrap verwendet. Er dient nur zur Berechnung des Konfidenzintervalls, d.h. es kann keine Überprüfung der Signifikanz von Kendall W damit durchgeführt werden.

Zur Durchführung des Bootstrap-Verfahrens benötigt man im SPSS- Programm eine Syntax, die im Anhang aufgeführt ist. Aus dem ursprünglichen Datensatz wird zufällig und mit Zurücklegen die gleiche Anzahl an Datenfällen, wie im ursprünglichen Datensatz, gezogen. Aufgrund des Zufallsprinzips können einige Fälle mehrfach gezogen werden, während andere Fälle gar nicht ausgewählt werden (siehe SPSS-Hilfe-Programm).

Diese Prozedur wird $n = 1000$ mal durchgeführt. Im Anschluss werden aus allen 1000 neu gebildeten Kendall W der Mittelwert, die Varianz des Mittelwerts und das Konfidenzintervall mit dem 2,5 und 97,5 Perzentil als untere und obere Begrenzung verwendet. Die Irrtumswahrscheinlichkeit wurde mit $\alpha = 0,05$ definiert.

Dieses Verfahren wird für alle Berufsgruppen separat durchgeführt. Anschließend werden die einzelnen Ergebnisse aus den Rohdaten und den Bootstrap-Verfahren miteinander verglichen.

2.7.3 Methodik zum Vergleich der einzelnen Rangreihenfolgen mit der idealen Rangreihenfolge der Literatur

2.7.3.1 Rangkorrelationskoeffizient Spearmans ρ :

Der Rangkorrelationskoeffizienten ρ nach Spearman ist ein Maß für die Enge des monotonen Zusammenhangs zwischen zwei Rangwertreihen.

Dieses Korrelationsmaß setzt ein Ordinalskalenniveau voraus. Man berechnet den Koeffizienten, indem man die Differenzen der Rangplätze der beiden Reihen quadriert und deren Summe bildet. Mittels folgender Formel wird anschließend der Spearman'schen Rangkorrelationskoeffizient errechnet [42, 43]:

$$d_i = x_i - y_i$$

$$\rho = 1 - \frac{6 \cdot \sum_{i=1}^N d_i^2}{N^3 - N}$$

d_i = Differenzwert

x_i = Rang in der ersten Reihe

y_i = Rang in der zweiten Reihe

ρ = Spearmans Rangkorrelationskoeffizient

N = Anzahl der Begriffe

Mithilfe des Rangkorrelationskoeffizienten Spearman's ρ wird die ideale Reihe aus der Literatur jeweils mit den drei neuen Rangreihe verglichen. Die Werte von ρ liegen zwischen -1 und 1, wenn der Zusammenhang perfekt positiv monoton bzw. perfekt negativ monoton ist [43].

Für diese Berechnung lauten die Hypothesen:

H_0 : Es besteht kein monotoner Zusammenhang zwischen der idealen Reihe und der jeweiligen Rangreihe der Berufsgruppen. $\rho = 0$

H_1 : Es besteht ein Zusammenhang zwischen den Reihen.

2.7.3.2 Signifikanz von Spearmans ρ

Bei kleinen Stichproben mit $N = 4$ bis 30 können kritische Werte des Rangkorrelationskoeffizienten aus einer spezifischen Tafel abgelesen werden. Für größere Stichproben $N \geq 10$ ist ρ verteilt wie die Prüfgröße t nach Student [42] mit $Fg = N - 2$:

$$t = \rho \cdot \sqrt{\frac{N - 2}{1 - \rho^2}}$$

ρ = Spearmans Rangkorrelationskoeffizient

N = Anzahl der Begriffe

t = Prüfgröße

2.7.3.3 Rangkorrelationskoeffizient Kendall τ

Der Rangkorrelationskoeffizient Kendall τ dient ebenfalls als Maß für den monotonen Zusammenhang zwischen zwei Rangreihen.

Er beruht auf der „Fehlordnung“ der Ränge innerhalb einer Vergleichsreihe. Eine Reihe wird als die natürlich gegebene angesehen. In der Vergleichsreihe werden die einzelnen Rangwerte paarweise miteinander verglichen. Hierbei entstehen Proversionen mit einem (+) gekennzeichnet, wenn sich die Rangwerte in aufsteigender Ordnung darstellen. Es entstehen Inversionen (-), wenn die Rangwerte absteigend angeordnet sind. Die Summe aus Proversionen und Inversionen gibt dann den Zusammenhang zwischen den beiden Reihen an [43]:

$$S_{max} = \frac{1}{2} \cdot N(N - 1)$$

$$S = P - I$$

$$\tau = \frac{S}{\frac{1}{2} \cdot N(N - 1)}$$

N = Anzahl der Objekte

P = Proversionen

I = Inversionen

S = Kendall Summe

S_{max} = höchstmögliche positive Kendall Summe

τ = Kendall Korrelationskoeffizient

Kendall τ kann Werte zwischen -1 und +1 annehmen. Ist die Vergleichsreihe analog der natürlichen Reihe (in diesem Fall idealen Reihe) aufsteigend geordnet, so besteht eine perfekt positive Rangkorrelation (+1). Bei absteigender Anordnung wäre die Korrelation perfekt negativ (-1). Einen Wert von 0 würde man als eine fehlende Korrelation interpretieren. Das bedeutet gleichzeitig, dass bei Überwiegen der Proversionen ein positiver Zusammenhang besteht, während bei Überwiegen der Inversionen ein negativer Zusammenhang anzunehmen ist [43, 45].

2.7.3.4 Signifikanz von Kendall τ :

Ist N = Anzahl der Objekte mit $N > 10$, so kann man τ als normalverteilt betrachten mit dem Mittelwert $\mu_\tau = 0$

und der Standardabweichung

$$\sigma_\tau = \sqrt{\frac{2(2N + 5)}{9N(N - 1)}}$$

Und damit ist

$$z = \frac{\tau - \mu_\tau}{\sigma_\tau} = \frac{\tau}{\sqrt{\frac{2(2N + 5)}{9N(N - 1)}}}$$

Die Signifikanz des z-Werts wird über eine spezifische Tafel bestimmt. Ist der p-Wert gleich oder kleiner als α ist H_0 zugunsten von H_1 zurückzuweisen [42].

2.7.3.5 Vergleich Spearmans ρ versus Kendall τ

Beide Rangkorrelationskoeffizienten bestimmen den monotonen Zusammenhang zweier Rangreihen. Sie sind jedoch numerisch nicht miteinander zu vergleichen, da sie auf anderen Rechenregeln basieren [43].

Beide Testverfahren weisen die gleiche Effizienz auf, d.h. sie weisen die Nullhypothese auf demselben Signifikanzniveau zurück.

Jedoch bestehen Besonderheiten. Die τ - Berechnung basiert ausschließlich auf rein ordinalen Informationen und stellt damit weniger Anforderungen an das Datenmaterial, d.h. es werden lediglich die Anzahl der „Größer-Relationen“ (Proversionen) und „Kleiner-Relationen“ (Inversionen) verwendet. Bei Spearmans Rangkorrelationskoeffizienten muss aber gerechtfertigt sein, dass zwischen aufeinander folgenden Rängen äquidistante Positionen auf dem Merkmal abgebildet werden. Damit kann es zu Überschätzungen von ρ kommen, wenn die Voraussetzungen nicht erfüllt sind [43]. Ein Vorteil ρ gegenüber Kendall τ ist, dass zwischen Kendall W und Spearmans ρ eine lineare Beziehung besteht, den es zwischen Kendall W und τ nicht gibt.

2.7.3.6 Permutationstest

Der Permutationstest ist ein nichtparametrisches Testverfahren. Es werden zwei Reihen miteinander verglichen, wobei nur eine Reihe zufällig beobachtet werden muss, während die andere Rangreihe vorgegeben sein kann [46].

In der Literatur ist eine „ideale“ Rangreihe vorgegeben. Um zu überprüfen, wie die einzelnen Rangreihen von der „idealen“ Rangreihe abweichen, wird hier der Permutationstest zur statistischen Überprüfung verwandt.

Die zu untersuchenden Hypothesen lauten:

H_0 : Die 20 Begriffe der echten Rangreihe sind zufällig angeordnet, d.h. es besteht keine Ähnlichkeit zwischen der echten und der idealen Rangreihe.

H_1 : Die Objekte der echten Rangreihe sind nicht zufällig angeordnet, d.h. es besteht Ähnlichkeit zwischen den beiden Reihen.

Zur Berechnung der Spearman'schen Richtlinie [46] dient folgende Formel:

$$\pi = \sum_{i=1}^n |R(X_i) - R(Y_i)|$$

$R(X_i)$ = Rang in der Reihe X des Wortes i.

$R(Y_i)$ = Rang in der Reihe Y des Wortes i.

Für den Mittelwert (AV) und die Standardabweichung (SD) der Spearman'schen Richtlinie gilt [46]:

$$AV(\pi) = \frac{1}{3}(n^2 - 1)$$
$$SD(\pi) = \frac{1}{45}(n + 1)(2n^2 + 7)$$

In dieser Arbeit ist $n = 20$, da 20 Wahrscheinlichkeitsbegriffe betrachtet werden.

Unter H_0 : „Die n Objekte sind zufällig angeordnet“ ist

$$T = \frac{\pi - AV}{\sqrt{SD}}$$

asymptotisch standardnormalverteilt [46].

Abweichungen von T sind in beide Richtungen möglich. H_0 wird bei einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$ abgelehnt falls $|T| > 1,96$ ist.

2.8 Methoden zum Vergleich der kontextfreien und kontextabhängigen Interpretation

Im Abschnitt B und D des Fragebogens sollten sechs Begriffe im Kontext interpretiert werden. Die Ergebnisse sollten dann mit der kontextfreien Interpretation verglichen werden. Nachfolgend werden die Fragestellungen und die Methodik zur Bearbeitung dieser dargestellt.

2.8.1 Fragestellung und Methodenübersicht

Zur Überprüfung, ob es eine Kontextabhängigkeit numerischer Interpretationen gibt, wurden im Fragebogen insgesamt 6 Begriffe in einem Kontext abgefragt. Die Begriffe „häufig“, „gelegentlich“ und „selten“ wurden im Kontext Nebenwirkungen interpretiert. Wie beschrieben, existiert eine Richtlinie des Bundesinstituts für Arzneimittel und Medizinprodukte.

Hier stellen sich die Fragen:

1. Gibt es eine Kontextabhängigkeit in der Interpretation?
2. Werden die Werte gemäß der Richtlinie richtig numerisch interpretiert?

Im Abschnitt D des Fragebogens wurde ein Rechtsstreit konstruiert. Hier wurden die Begriffe „möglicherweise“, „oft“ und „wahrscheinlich“ zur Interpretation angeboten. Diese drei Begriffe wurden willkürlich ausgewählt. Eine offizielle numerische Klassifikation dieser verbalen Wahrscheinlichkeiten gibt es nicht. Trotzdem wurde dieser Abschnitt eingebaut, um möglichst von der offiziellen Richtlinie abzulenken.

Die Fragestellungen des Abschnitts D lauten:

1. Stimmen die numerischen Interpretationen der drei Wahrscheinlichkeitsbegriffe im Kontext „Rechtsstreit“ mit den kontextfreien numerischen Interpretationen dieser Begriffe im ersten Abschnitt überein?
2. Gibt es Unterschiede zwischen den einzelnen Berufsgruppen?

Die Methodik des Abschnitts D entspricht der des Abschnitts B. Berechnet werden die Mediane, Minimum, Maximum, sowie das 25%- und das 75%-Quantil. Diese Berechnungen erfolgen für alle Studienteilnehmer zusammenfassend und anschließend berufsgruppenspezifisch.

Die Prozentwerte der Begriffe aus Abschnitt A werden dann den Werten aus Abschnitt B bzw. D graphisch gegenübergestellt. Dies erfolgt nach Berufsgruppen getrennt.

Im Anschluss werden die numerischen Wahrscheinlichkeiten der Probanden mit den klar definierten Wahrscheinlichkeiten für Nebenwirkungen (BfArM) verglichen. Die korrekten und inkorrekten Interpretationen werden aufgelistet.

Anschließend werden die Interpretationen der Begriffe aus Abschnitt A und B bzw. D mit dem Wilcoxon-Test verglichen. Der Kruskal-Wallis-Test untersucht, ob sich die drei Berufsgruppen in der Verteilung der Prozentwerte zwischen den Abschnitten unterscheiden.

Wenn es Unterschiede gibt, muss geklärt werden, wo die Unterschiede liegen. Dafür benutzt man den Mann-Whitney-Test, der jeweils zwei Berufsgruppen miteinander vergleicht.

2.8.2 Methodik zum Vergleich der kontextfreien und kontext-abhängigen Interpretation

2.8.2.1 Wilcoxon-Test für abhängige Paare:

Der Wilcoxon-Test ist ein nichtparametrisches Testverfahren. Zwei abhängige Stichproben werden hinsichtlich ihrer zentralen Tendenz verglichen. Er ist ein Test, der Daten auf Ordinalskalenniveau bearbeitet [43].

Man berechnet zunächst die Differenzbeträge d_i für jedes abhängige Paar aus den beiden Stichproben. Anschließend werden alle d_i -Werte ungeachtet ihres Vorzeichens in eine Rangreihe gebracht. Der kleinste d_i -Wert erhält dabei den Rang 1. Des Weiteren werden dann die Summen der negativen bzw. der positiven Rangplätze gebildet (T und T'). Hierbei werden Paardifferenzen, die aufgrund gleicher Messwerte Null ergeben, ignoriert und gehen somit nicht in die Berechnung ein [42].

Der Wilcoxon-Test diente in dieser Studie dem Vergleich zweier numerischer Interpretationen eines Wortes. Dabei wurden beispielsweise die Werte von „häufig“ aus dem kontextfreien ersten Abschnitt mit den Werten von „häufig“ aus dem Kontext „Nebenwirkungen“ des zweiten Abschnitts verglichen.

Mit diesem Test-Verfahren wurden die folgenden Hypothesen geprüft:

H_0 : Es besteht kein Unterschied zwischen den jeweiligen Paaren, d.h. die Summe der Rangplätze mit positiven Vorzeichen ist gleich der Summe der Rangplätze mit negativem Vorzeichen.

H_1 : Die jeweiligen Paare unterscheiden sich in ihrer numerischen Interpretation.

2.8.2.2 Signifikanzprüfung:

Zur Signifikanzprüfung wird die kleinere der beiden Rangsummen (T) verwendet.

Für kleine Stichproben sind die unteren T-Werte für verschiedene Signifikanzniveaus gelistet. Bei T- Werten gleich oder kleiner der Werte der Tafel kann die Nullhypothese auf dem jeweiligen Signifikanzniveau zurückgewiesen werden.

Bei großen Stichproben mit $N > 25$, wie in dieser Studie, geht die Verteilung der T-Werte in eine Normalverteilung über. Somit kann mit dem Mittelwert und der Standardabweichung eine Transformation des T-Werts in einen z-Wert erfolgen. In der spezifischen Tafel für Normalverteilungen ist dann die Überschreitungswahrscheinlichkeit für den errechneten z-Wert abzulesen und die H_0 zu verwerfen oder anzunehmen [42].

Die Formeln lauten:

$$\mu_T = \frac{N(N+1)}{4}$$

$$\sigma_T = \sqrt{\frac{N(N+1)(2N+1)}{24}}$$

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{T - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

μ_T = Erwartungswert

σ_T = Standardabweichung

N = Anzahl der Paardifferenzen

T = kleinere der beiden Rangsummen

Die Signifikanztestung erfolgte bei einem α - Niveau von 5% zweiseitig, da die Richtung des Unterschieds, d.h. sind die numerischen Interpretationen des Kontexts größer oder kleiner als die freien Interpretationen, nicht prognostiziert wurde.

2.8.2.3 Hodges-Lehmann- Punktschätzer:

Es wird der Punktschätzer nach Hodges-Lehmann bestimmt. Hierfür werden alle möglichen Paare von Studienteilnehmern aus jeweils zwei Berufsgruppen gebildet. Von diesen Paaren wird die Differenz berechnet und von diesen der Median über alle Paare ermittelt. Der Punktschätzer gibt somit den Wert der Mediantdifferenz an.

Da diese Untersuchung eine Stichprobe darstellt, kann der Punktschätzer für die Grundgesamtheit nur geschätzt werden, und es besteht die Möglichkeit eines Zufallsfehlers. Aus diesem Grunde wurde gleichzeitig ein 95%-Konfidenzintervall bestimmt. Somit kann mit 95%iger-Wahrscheinlichkeit gesagt werden, dass der Median innerhalb des Konfidenzintervalls liegt.

Das Prinzip des Hodges-Lehmann Punktschätzers wird von Everitt [47] folgendermaßen definiert:

“The principal form of this estimator is used to give an estimate of the difference between the values in two sets of data. If the two sets of data contain m and n data points respectively, $m \times n$ pairs of points (one from each set) can be formed and each pair gives a difference of values. The Hodges–Lehmann estimator for the difference is defined as the median of the $m \times n$ differences.”

Ist die Anzahl der Nulldifferenzen groß, so weist dies auf die Richtigkeit der Nullhypothese hin. Da beim Wilcoxon-Test die Nulldifferenzen jedoch unberücksichtigt bleiben, begünstigt dieser die H_1 . Somit wurde anschließend der Nulldifferenzen-Test nach Cureton durchgeführt [43, 48].

2.8.2.4 Kruskal-Wallis-Test

Der Kruskal-Wallis-Test ist ebenfalls ein nichtparametrisches Testverfahren, welches ein ordinalskaliertes Niveau fordert. Er wird verwendet, wenn mehr als zwei unabhängige Stichproben miteinander verglichen werden sollen. Er prüft, ob es einen Unterschied in ihrer Verteilung gibt.

Um herauszufinden, ob sich die drei Berufsgruppen hinsichtlich der Prozentwerte der Wahrscheinlichkeitsbegriffe aus Abschnitt A und B unterscheiden, wird der Kruskal-Wallis-Test für k unabhängige Stichproben durchgeführt.

Die zu untersuchenden zweiseitigen Hypothesen bei einem Signifikanzniveau von $\alpha = 0,05$ lauten:

$$H_0: F_1(x) = F_2(x) = F_3(x)$$

$$H_1: F_i(x) = F(x - \theta_i) (x \in \mathbb{R}) \text{ mit } \theta \neq 0 \text{ für mindestens ein } i$$

$F_1(x)$: Verteilung der Prozentangaben in der Berufsgruppe der Ärzte

$F_2(x)$: Verteilung der Prozentangaben in der Berufsgruppe der Apotheker

$F_3(x)$: Verteilung der Prozentangaben in der Berufsgruppe der Juristen

Die Nullhypothese H_0 besagt, dass es keinen Unterschied zwischen den Ärzten, Apothekern und Juristen für die Wahrscheinlichkeitsbegriffe gibt. Die Alternativhypothese H_1 sagt dagegen aus, dass sich mindestens zwei Berufsgruppen in der Verteilung unterscheiden.

Die genaue Methodik ist dem Bortz [43] zu entnehmen. Die Prozentwerte eines Begriffs (z.B. „häufig“) aller Teilnehmer erhalten Rangplätze von 1 bis N. Anschließend wird der Durchschnitt R_j der Spaltenrangsummen T_j berechnet. Man erhält die gesuchte Prüfgröße mit der Formel:

$$H = \frac{N-1}{N} \cdot \sum_{j=1}^k \frac{\left(R_j - \frac{N+1}{2}\right)^2}{\frac{N^2-1}{N_j}}$$

Der Test wird in dieser Studie verwandt, um die Prozentangaben der Berufsgruppen miteinander zu vergleichen. Außerdem sollen die Berufsgruppen hinsichtlich der Differenzen zwischen kontextfreier und kontextbezogener Interpretation verglichen werden.

2.8.2.5 Mann-Whitney-Test

Auch der Mann-Whitney-Test ist ein nichtparametrisches Testverfahren. Mit diesem Test kann die Hypothese überprüft werden, ob sich zwei Populationen (unabhängige Stichproben) in der Verteilung eines Merkmals unterscheiden.

Jeweils zwei Berufsgruppen (Ärzte/Apotheker; Ärzte/Juristen; Apotheker/Juristen) werden miteinander verglichen. So wird z.B. für die Prozentwerte von „häufig ohne Kontext“ von allen Ärzten und Apothekern eine gemeinsame Rangreihe transformiert. Der Unterschied in der Summe der Rangplätze (r_1 versus r_2) jeder Gruppe kennzeichnet möglicherweise den Unterschied in den Prozentwerten. Die Prüfgröße U beschreibt, wie häufig ein Rangplatz in der Stichprobe größer ist als in der anderen Stichprobe [43].

Hierbei ist:

$$u = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + n_2)}{2} - r_1$$

R_1 = Summe der Rangplätze in der Stichprobe 1

N_1 = Anzahl der Messwerte der Stichprobe 1

U = Wert der Teststatistik

Für große Stichprobenumfänge ($n_1 > 25$ oder $n_2 > 25$):

$$z = z^{korr} = \frac{u - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2}{n \cdot (n-1)} \cdot \left(\frac{n^3 - n}{12} - \sum_{k=1}^k \frac{t_k^3 - t_k}{12}\right)}}$$

k = Anzahl der Messwertgruppen mit den gleichen Werten

t_k = Anzahl der Bindungen in der k-ten Gruppe

Bei einem Signifikanzniveau von $\alpha = 0,05$ lauten die Hypothesen:

$$H_0: F_1(x) = F_2(x)$$

$$H_1: F_1(x) \neq F_2(x)$$

$F_1(x)$ = Verteilung der Prozentangaben in der ersten zu vergleichenden Gruppe

$F_2(x)$ = Verteilung der Prozentangaben in der zweiten zu vergl. Gruppe

Bei einem Signifikanzniveau von $\alpha = 0,05$ gilt die Nullhypothese als verworfen, wenn der p-Wert kleiner als das Signifikanzniveau ist.

Es werden die Vergleiche für die Begriffe „häufig“, „gelegentlich“ und „selten“ aus Abschnitt A (ohne Kontext) und B (Kontext) durchgeführt. Auch hier wird erneut der Hodges-Lehmann- Punktschätzer verwendet.

2.9 Vergleich mit der offiziellen Klassifikation für Arzneimittelnebenwirkungen

Im Abschnitt C des Fragebogens wurden Verhältniszahlen in einen Kontext gesetzt und sollten mit verbalen Wahrscheinlichkeitsausdrücken interpretiert werden. Nachfolgend wird das Vorgehen zur Auswertung dieses Abschnitts beschrieben.

2.9.1 Fragestellung

Das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) in Bonn erarbeitete 2002 eine Richtlinie für „Häufigkeitsangaben von Nebenwirkungen“ [26]. Die Begriffe „sehr häufig“, „häufig“, „gelegentlich“, „selten“ und „sehr selten“ wurden vom BfArM klar numerisch definiert. In Tabelle 3 erfolgt die Darstellung dieser Richtlinie.

Tabelle 3 Richtlinie für die „Häufigkeitsangaben für Nebenwirkungen“ des BfArM

SPC-Richtlinie	deutsch	%-Angaben
very common	sehr häufig	$\geq 10\%$
common	häufig	$\geq 1\% \text{ -- } < 10\%$
uncommon	gelegentlich	$\geq 0,1\% \text{ -- } < 1\%$
rare	selten	$\geq 0,01\% \text{ -- } < 0,1\%$
very rare	sehr selten	$< 0,01\%$

Im Abschnitt B des Fragebogens wurden den Probanden die drei Begriffe „häufig“, „gelegentlich“ und „selten“ zur numerischen Interpretation angeboten. Es wurden nur drei dieser Werte benutzt, um eine Assoziation mit der offiziellen Nebenwirkungstabelle zu vermeiden und somit die Fragestellung nicht zu offenkundig freizugeben.

Auch der Abschnitt C des Fragebogens war in einen Kontext eingebettet. Hier wurden jedoch Verhältniszahlen angeboten, die mit Begriffen belegt werden sollten. So sollte z.B. beschrieben werden, welchen Begriff man verwenden würde, um ein OP-Risiko von 1:1000 auszudrücken. Die Auswahl der Begriffe war frei wählbar.

Für die Wahrscheinlichkeiten „1:1000“, „1:100“, „1:10“, „1:5“ und „1:2“ werden die fünf häufigsten Begriffe tabellarisch dargestellt. Anschließend soll geprüft werden, wie viele korrekte und inkorrekte verbale Zuordnungen, gemäß der offiziellen Klassifikation für Arzneimittelnebenwirkungen (siehe Tabelle 3), zu den Wahrscheinlichkeiten „1:1000“, „1:100“, „1:10“ vorliegen.

Ziel ist es herauszufinden, ob die Studienteilnehmer die offizielle Klassifikation kennen und korrekt gebrauchen.

3 Ergebnisse

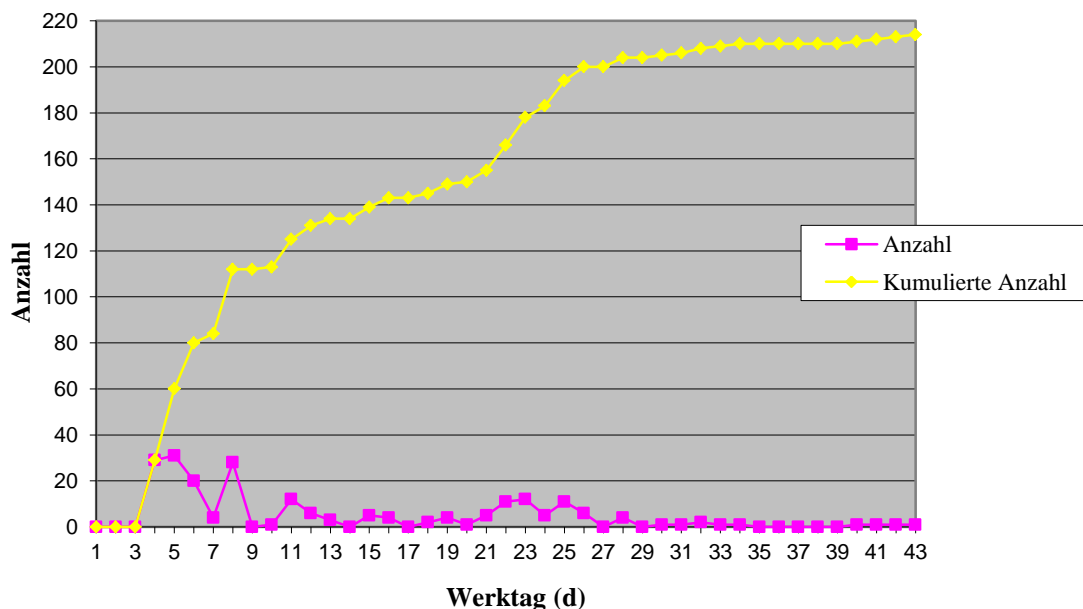
Im nächsten Abschnitt dieser Arbeit werden die Ergebnisse des Rücklaufs, der Ausschöpfung und die soziodemographischen Daten dargestellt. Im Anschluss werden die Ergebnisse der Rangreihenfolgen, des Kontextvergleichs und des Vergleichs mit der offiziellen Nebenwirkungstabelle dargelegt.

3.1 Rücklauf und Ausschöpfungsquote

3.1.1 Rücklauf und Rücksendedauer

Im ersten Durchlauf der Erhebung waren 350 der insgesamt 500 verschickten Fragebögen für diese Studie relevant. Insgesamt wurden 214 der 350 Bögen im ersten Durchlauf beantwortet und an das Institut für Medizinische Biometrie zurückgeschickt. Bereits am 8. Werktag waren 52,3% (112) der Fragebögen zurückgesandt worden. Ein Erinnerungsschreiben wurde am 17. Werktag verschickt. Zu diesem Zeitpunkt waren 66,8% der Antworten bereits eingetroffen. Im ersten Durchlauf vergingen somit vom Versandtag bis zur Ankunft des letzten Fragebogens 67 Werkstage. Nach diesem Tag trafen keine weiteren Antworten mehr ein. Eine Tabelle der Ankunftsdaten des ersten und zweiten Durchlaufs sind im Anhang (siehe Tabelle 51 und Tabelle 52) hinterlegt.

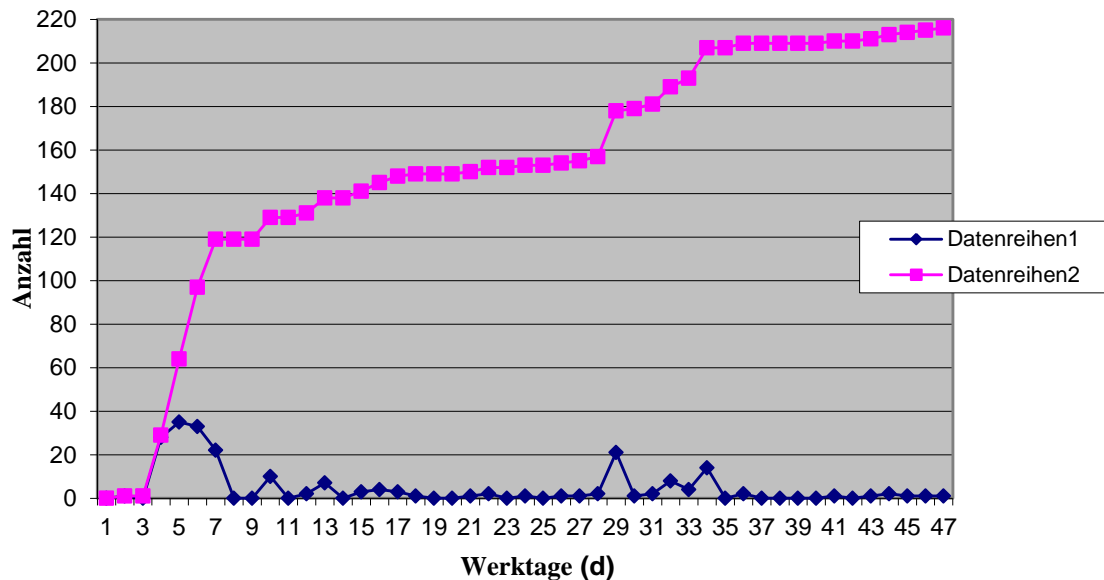
Ankunftstage des 1. Durchlaufs



Im zweiten Durchlauf wurden 216 von 350 Fragebögen zurückgeschickt. Damit verlief der zweite Durchlauf so gut wie der Erste. Hier waren am 7. Werktag bereits 55,1% der Briefe im Institut eingetroffen. Um das Erinnerungsschreiben nicht in die Ferienzeit fallen zu lassen, wurde dieses erst am 24. Werktag versandt. Bis zu diesem Zeitpunkt

waren 70,8% der Schreiben wieder eingetroffen. Bis zum Eintreffen des letzten Fragebogens vergingen im zweiten Lauf insgesamt 110 Werktage.

Ankunftstage des 2. Durchlaufs



3.1.2 Ausschöpfung

Wie bereits im Methodik-Abschnitt beschrieben, wird die Ausschöpfung als Maß für die Qualität des Fragebogens verwendet. Die Berechnung erfolgte standardisiert nach PORST [40].

Beispielsweise wurden im ersten Durchlauf 150 Ärzte angeschrieben, von denen ein Fragebogen nicht zugesandt werden konnte (siehe Tabelle 4). Es ergibt sich eine bereinigte Stichprobe von 149 möglichen Probanden. Bei 60 nicht-neutralen Ausfällen errechnet sich eine Ausschöpfung von 89. Hieraus kann man demzufolge eine Ausschöpfungsquote der Ärzte von 59,7% ermitteln. Für die Apotheker ergibt sich im ersten Durchlauf eine Quote von 52% und für die Juristen von 73,7%. Insgesamt liegt die Ausschöpfungsquote des ersten Durchlaufs bei 61,5%.

Tabelle 4: Ausschöpfung des ersten Durchlaufs

	Ärzte	Apotheker	Juristen
Bruttostichprobe	150	100	100
neutrale Ausfälle	1	0	1
bereinigte Stichprobe	149	100	99
nicht-neutrale Ausfälle	60	48	26
Auschöpfung	89	52	73
Auschöpfungsquote %	59,7	52,0	73,7

Neutrale Ausfälle: Briefe mit falscher Adresse oder verzogener Proband; nicht- neutrale Ausfälle sind Personen, die aktiv die Teilnahme verweigern

Im zweiten Durchlauf ergibt sich eine Gesamtausschöpfungsquote von 62,4 %. Auch in diesem Durchlauf haben prozentual wieder mehr Juristen geantwortet. Erneut haben weniger Apotheker geantwortet.

Tabelle 5: Ausschöpfung des zweiten Durchlaufs

	Ärzte	Apotheker	Juristen
Bruttostichprobe	150	100	100
neutrale Ausfälle	1	0	3
bereinigte Stichprobe	149	100	97
nicht-neutrale Ausfälle	56	45	29
Ausschöpfung	93	55	68
Ausschöpfungsquote %	62,4	55,0	70,1

Neutrale Ausfälle: Briefe mit falscher Adresse oder verzogener Proband; nicht- neutrale Ausfälle sind Personen, die aktiv die Teilnahme verweigern

Zusammenfassend wird somit aus dem gesamten Stichprobenumfang bei 700 verschickten Fragebögen eine Ausschöpfungsquote von 62 % erreicht.

Tabelle 6: Ausschöpfung des gesamten Untersuchungsfeldes

	Ärzte	Apotheker	Juristen	Gesamt
Bruttostichprobe	300	200	200	700
neutrale Ausfälle	2	0	4	6
bereinigte Stichprobe	298	200	196	694
nicht-neutrale Ausfälle	116	93	55	264
Ausschöpfung	182	107	141	430
Ausschöpfungsquote in %	61,1	53,5	71,9	62,0

Neutrale Ausfälle: Briefe mit falscher Adresse oder verzogener Proband; nicht- neutrale Ausfälle sind Personen, die aktiv die Teilnahme verweigern.

3.2 Soziodemographische Daten

Im letzten Abschnitt des Fragebogens wurden von den Teilnehmern soziodemographische Daten erfragt. Es sollten Angaben über das Geschlecht, das Alter, die Fachrichtung und die Einrichtung, in der die Befragten arbeiten, gemacht werden.

Von den 430 Teilnehmern haben 427 Angaben zu ihrer Berufsgruppe gemacht. 182 Probanden sind Ärzte. Dies entspricht einem prozentualen Anteil von 42,6% der gesamten Teilnehmer. 105 Befragte gaben an, Apotheker zu sein. Mit 140 Juristen stellen diese die zweigrößte Gruppe dar (siehe Tabelle 7).

Tabelle 7: absoluter und prozentualer Anteil der Befragten aufgeschlüsselt nach Berufsgruppen

	Häufigkeit	Prozent (%)
Ärzte	182	42,6
Apotheker	105	24,6
Juristen	140	32,8

3.2.1 Geschlechterverteilung der Probanden

Schaut man sich im Weiteren die Geschlechtsverteilung an, so stellt man fest, dass deutlich mehr Männer an der Studie teilgenommen haben. Insgesamt haben 424 Probanden Angaben zu ihrem Geschlecht gemacht. Hiervon sind 318 Studienteilnehmer männlich. Damit sind 75% der Probanden männlich und 25% weiblich (siehe Tabelle 8).

Auf die einzelnen Berufsgruppen verteilt ergibt sich bei den Ärzten ein Frauenanteil von 17,7%. Bei den Apothekern sind immerhin 37,1% der Probanden weiblich. Nicht berücksichtigt bleiben die sechs Fragebögen ohne Geschlechtsangabe.

Tabelle 8: Anzahl der Männer und Frauen der einzelnen Berufsgruppen, dargestellt in absoluten Werten und Prozentangaben.

Berufsgruppe	Geschlecht		Anzahl gesamt
	männlich	weiblich	
Ärzte	149 (82,3%)	32 (17,7%)	181 (100%)
Apotheker	66 (62,9%)	39 (37,1%)	105 (100%)
Juristen	103 (74,6%)	35 (25,4%)	138 (100%)
Gesamt	318 (75,0%)	106 (25,0%)	424 (100%)

3.2.2 Altersverteilung der Probanden

Auch bezüglich des Alters gab es einige unvollständige Fragebögen. Diese bleiben hier unberücksichtigt. Insgesamt haben 419 Studienteilnehmer Angaben zu ihrem Alter gemacht. Der jüngste Befragte ist 27 Jahre alt, der älteste 74 Jahre (siehe Tabelle 9). Das Durchschnittsalter liegt bei 49 Jahren. 25% aller Probanden sind jünger als 42 Jahren und 25% älter als 56 Jahre.

Zusammenfassend entsprechen die Altersverteilungen der einzelnen Berufsgruppen einander. Erfreulicherweise sind Probanden aller Altersschichten vertreten. Man kann somit von einer Gleichheit der Berufsgruppen in Hinblick auf das Alter ausgehen. Im Anhang ist eine Tabelle 53 mit den vollständigen Streuungsmaßen hinterlegt.

Tabelle 9: Alter der Teilnehmer der Studie nach Berufsgruppen unterteilt.

	N	Mittelwert	Minimum	Maximum
Ärzte	178	50	33	69
Apotheker	104	49	27	74
Juristen	137	47	30	73

N= Anzahl der Studienteilnehmer, Alter angegeben in Jahren,; die Dezimalstellen wurden nicht dargestellt. Vier Ärzte, ein Apotheker und drei Juristen haben keine Angaben zu ihrem Alter gemacht.

3.3 Rangreihenfolgen und -vergleich

Im ersten Abschnitt des Fragebogens wurden 20 Wahrscheinlichkeitsbegriffe angeboten, die von den Teilnehmern mit Prozentwerten belegt werden sollten. Zunächst werden diese, sowie die Lage- und Streuungsmaße (Median, Quantile usw.) dargestellt.

Anschließend wird überprüft, ob sich die individuellen Rangreihen in den einzelnen Berufsgruppen ähneln. Dann werden die Rangreihen der einzelnen Berufsgruppen mit der idealen Reihe aus der Literatur verglichen.

3.3.1 Numerische Interpretation der verbalen Wahrscheinlichkeiten

In der Literatur wird eine große Spannweite in der Interpretation von verbalen Wahrscheinlichkeiten beschrieben. Die meisten Studien wurden an Ärzten durchgeführt. Schaut man sich beispielsweise den Begriff „wahrscheinlich“ an, so errechnet sich für die Ärzte ein Median von 80%. Minimaler Wert ist immerhin 5% und der maximal vergebene Prozentwert ist 98% (siehe Tabelle 10). Bei den Apothekern liegt der Median bei 75%. Als niedrigster Wert wurde 2% und als höchster Wert 99% angegeben. Ähnlich verhält es sich auch bei den Juristen. Hier beträgt der Median wiederum 75%, aber der minimal gewählte Prozentwert lag bei 0,5% und das Maximum bei 95%.

Des Weiteren fällt auf, dass die Begriffe „unklar“, „fraglich“ und „fraglos“ von einigen Teilnehmern nicht mit Prozentangaben belegt wurden. Tabelle 13 verdeutlicht dies noch

etwas klarer. Außerdem werden in allen Berufsgruppen für „unklar“, „unsicher“, „fraglich“, „denkbar“ und auch „möglicherweise“ große Quantil-Abstände beobachtet. So ergibt sich für „unsicher“ bei den Apothekern ein 25%-Quantil von 10% und ein 75%-Quantil von 50%. Somit haben 50% der Apotheker „unsicher“ mit einem Wert zwischen 10 und 50% bewertet. Dies ist ein sehr breites Quantil. Auch die weiteren genannten Begriffe zeigen ähnlich breite Quantil-Abstände. Die graphische Darstellung der 20 Wahrscheinlichkeitsbegriffe erfolgt im Anhang (siehe 7.5).

Zusammenfassend zeigen diese Ergebnisse, dass auch Juristen und Apotheker eine große Spannweite in der numerischen Interpretation bieten. Einige der Begriffe wurden gar nicht erst mit Prozentwerten belegt bzw. die Spannbreite auch in den Quantilen ist groß.

Tabelle 10: gewählte Prozentwerte der 20 Wahrscheinlichkeitsbegriffe der Mediziner; mit Angabe der Lage und Streuungsmaße (in Prozent)

	n = Anzahl der Befragten	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
selten	179	5.00	0.05	100.00	2.50	10.00
häufig	178	75.00	5.00	100.00	60.00	80.00
niemals	178	0.00	0.00	100.00	0.00	0.00
unklar	149	35.00	0.00	100.00	15.00	50.00
sicher	175	99.00	0.00	100.00	95.00	100.00
normalerweise	178	90.00	10.00	100.00	80.00	90.00
unsicher	168	50.00	0.00	100.00	25.00	50.00
fraglos	162	95.00	0.00	100.00	80.00	100.00
oft	178	70.00	10.00	100.00	60.00	80.00
stets	178	99.00	0.00	100.00	90.00	100.00
möglicherweise	172	30.00	0.55	100.00	20.00	50.00
nie	178	0.00	0.00	100.00	0.00	0.10
gelegentlich	178	20.00	1.00	90.00	10.00	25.00
wahrscheinlich	179	80.00	5.00	98.00	70.00	90.00
überwiegend	177	75.00	20.00	100.00	70.00	80.00
immer	178	100.00	0.00	100.00	99.00	100.00
meist	178	90.00	20.00	99.00	75.00	90.00
fraglich	159	20.00	0.00	80.00	10.00	50.00
denkbar	173	20.00	0.00	90.00	10.00	40.00
typischerweise	179	80.00	0.00	100.00	75.00	90.00

Tabelle 11: gewählte Prozentwerte der 20 Wahrscheinlichkeitsbegriffe der Apotheker; mit Angabe der Lage und Streuungsmaße (in Prozent)

	n = Anzahl der Befragten	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
selten	105	5.00	0.00	95.00	1.00	10.00
häufig	105	75.00	5.00	95.00	60.00	80.00
niemals	104	0.00	0.00	100.00	0.00	0.00
unklar	99	20.00	0.00	95.00	10.00	50.00
sicher	104	99.00	2.00	100.00	90.00	100.00
normalerweise	104	80.00	1.00	100.00	70.00	90.00
unsicher	104	22.50	0.00	95.00	10.00	50.00
fraglos	98	95.00	0.00	100.00	80.00	100.00
oft	104	70.00	10.00	95.00	60.00	80.00
stets	105	99.00	0.00	100.00	90.00	100.00
möglicherweise	103	30.00	0.00	70.00	20.00	50.00
nie	104	0.00	0.00	100.00	0.00	0.98
gelegentlich	105	20.00	0.00	50.00	10.00	27.50
wahrscheinlich	105	75.00	2.00	99.00	60.00	82.50
überwiegend	105	75.00	0.00	99.00	60.00	80.00
immer	105	100.00	5.00	100.00	99.00	100.00
meist	105	80.00	1.00	99.00	75.00	90.00
fraglich	102	20.00	0.00	75.00	5.00	40.00
denkbar	103	20.00	0.00	75.00	5.00	50.00
typischerweise	103	75.00	0.00	100.00	60.00	80.00

Tabelle 12: gewählte Prozentwerte der 20 Wahrscheinlichkeitsbegriffe der Juristen; mit Angabe der Lage und Streuungsmaße (in Prozent)

	n = Anzahl der Befragten	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
selten	140	10.00	0.01	60.00	5.00	10.00
häufig	140	75.00	0.50	95.00	60.63	80.00
niemals	140	0.00	0.00	10.00	0.00	0.00
unklar	128	30.00	0.00	90.00	20.00	50.00
sicher	140	100.00	0.00	100.00	91.25	100.00
normalerweise	139	80.00	0.50	100.00	70.00	90.00
unsicher	138	50.00	0.00	80.00	25.00	50.00
fraglos	135	98.00	0.00	100.00	90.00	100.00
oft	140	75.00	0.50	95.00	60.00	80.00
stets	140	100.00	0.00	100.00	95.00	100.00
möglicherweise	139	40.00	0.50	90.00	30.00	50.00
nie	140	0.00	0.00	100.00	0.00	0.00
gelegentlich	140	20.00	0.50	75.00	20.00	30.00
wahrscheinlich	139	75.00	0.50	95.00	60.00	80.00
überwiegend	140	75.00	0.50	95.00	60.00	80.00

immer	140	100.00	0.50	100.00	99.93	100.00
meist	140	80.00	0.50	98.00	75.00	86.88
fraglich	136	40.00	0.10	90.00	25.00	50.00
denkbar	140	30.00	0.50	90.00	10.00	50.00
typischerweise	140	80.00	0.50	110.00	70.00	85.00

Tabelle 13: gewählte Prozentwerte der 20 Wahrscheinlichkeitsbegriffe für alle Probanden zusammengefasst; mit Angabe der Lage- und Streuungsmaßen (in Prozent angegeben)

	n = Anzahl der Geantworteten	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
selten	427	5	0.000001	100	3.00	10
häufig	426	75	0.50	100	60.00	80
niemals	425	0	0.00	100	0.00	0
unklar	379	30	0.00	100	12.00	50
sicher	422	99	0.00	100	95.00	100
normalerweise	424	80	0.50	100	75.00	90
unsicher	413	45	0.00	100	20.00	50
fraglos	398	95	0.00	100	89.25	100
oft	425	70	0.50	100	60.00	80
stets	426	100	0.00	100	95.00	100
möglicherweise	417	35	0.00	100	20.00	50
nie	425	0	0.00	100	0.00	0.05
gelegentlich	426	20	0.00	90	10.00	30
wahrscheinlich	426	75	0.50	99	60.00	85
überwiegend	425	75	0.00	100	65.00	80
immer	426	100	0.00	100	99.00	100
meist	426	80	0.50	99	75.00	90
fraglich	400	25	0.00	90	10.00	50
denkbar	419	20	0.00	90	10.00	49
typischerweise	425	80	0.00	110	70.00	90

3.3.2 Rangreihenfolgen

Zunächst wurde für jede einzelne Person eine Rangreihe der 20 Begriffe erstellt. Die Probanden wurden gemäß ihrer Berufe in drei Gruppen eingeteilt. Hieraus konnte für jede einzelne Berufsgruppe eine Rangreihe erstellt werden, so dass insgesamt drei neue Rangreihen entstanden (Tabelle 16). Es wurde anschließend der **Kendall Konkordanzkoeffizient W** für jede Berufsgruppe berechnet (siehe Tabelle 14). So wurden beispielsweise die Rangreihen aller Apotheker mithilfe von Kendall W verglichen. Bei einem Wert von 0,696 ähneln sich die einzelnen Rangreihen somit sehr.

Tabelle 14 Konkordanzkoeffizient Kendall W für jede Berufsgruppe mit χ^2 und dazugehörigem p-Wert

	Kendall W	χ^2	p-Wert
Ärzte	0,684	1807,35	$p \leq 0,0001$
Apotheker	0,696	1256,59	$p \leq 0,0001$
Juristen	0,709	1683,40	$p \leq 0,0001$

Kendall W ist hier ein Maß für den Zusammenhang der Rangreihen innerhalb einer Berufsgruppe (z.B. Ärzte), χ^2 dient zur Signifikanztestung.

Um nun zu erfahren, ob aus den Ergebnissen der Stichprobe auf die Grundgesamtheit geschlossen werden kann, mussten die jeweiligen Konkordanzkoeffizienten auf ihre Signifikanz getestet werden. Mithilfe einer spezifischen Tafel konnte das jeweilige Signifikanzniveau, auf dem die Nullhypothese zurückgewiesen werden kann, abgelesen werden.

Beispiel Ärzte: $\chi^2 = 139 * (20 - 1) * 0,684 = 1807,35$ mit $p \leq 0,0001$

Für alle drei Berufsgruppen liegen signifikante Kendall W's vor.

Wie beschrieben, wurde anschließend der **Bootstrap** verwandt, um ein Konfidenzintervall von $p = 95\%$ bei einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ zu erhalten. Als untere Grenze für das Konfidenzintervall wird somit das 2,5 Perzentil und für die obere Grenze das 97,5 Perzentil gewählt.

Tabelle 15 Vergleich Kendall W's mit dem gemittelten Kendall W_b 's des Bootstraps, sowie Angabe des Konfidenzintervalls (2,5 und 97,5 Perzentil) für die Ärzte, Apotheker und Juristen

	Kendall W	Mittelwert W_b	Varianz des Mittelwerts	Konfidenzintervall
Ärzte	0,684	0,688	0,001	(0,621 ; 0,755)
Apotheker	0,696	0,699	0,002	(0,615 ; 0,783)
Juristen	0,709	0,714	0,002	(0,634 ; 0,789)

Mittelwert W_b stellt den Mittelwert, der aus den 1000 Neuziehungen jeweils neu errechneten Kendall Ws, dar.

Der gemittelte Kendall W für die Ärzte nach dem Bootstrap- Verfahren liegt bei 0,688. Dieser Wert liegt deutlich im ermittelten Konfidenzintervall (0,621; 0,755). Auch für die Juristen und Apotheker liegen die, durch den Bootstrap ermittelten, neuen Kendall W- Werte im 95%- Konfidenzintervall.

Somit kann die Nullhypothese zugunsten der Alternativhypothese verworfen werden. Es besteht bei einem gemittelten Kendall W von 0,699 ein Zusammenhang zwischen den einzelnen Rangreihen der Apotheker. Dies trifft auch für die Juristen und die Mediziner zu. Die drei neuen echten Rangreihen der Berufsgruppen sind im Abschnitt 3.3.3 dargestellt.

3.3.3 Vergleich der drei neuen Rangreihen mit der idealen Rangreihe der Literatur

Die drei neuen Rangreihen der einzelnen Berufsgruppen sollten dann mit der idealen Rangreihe der Literatur verglichen werden. Die Tabelle 16 stellt die Rangreihen der Ärzte, Apotheker und Juristen der idealen Rangreihe gegenüber. Es werden die 20 Begriffe entsprechend ihres Rangplatzes aufgelistet. Im Anhang (siehe Tabelle 54) kann diese Tabelle mit den entsprechenden numerischen Rangplätzen eingesehen werden.

Tabelle 16: Vergleich der Rangreihen der Ärzte, Apotheker und Juristen mit der idealen Rangreihe der deutschsprachigen Literatur

Rangplatz	ideale Reihe	Ärzte	Apotheker	Juristen
1	niemals	niemals	niemals	niemals
2	selten	nie	nie	selten
3	nie	selten	selten	nie
4	gelegentlich	gelegentlich	denkbar	gelegentlich
5	unsicher	denkbar	fraglich	denkbar
6	denkbar	fraglich	gelegentlich	unklar
7	fraglich	unklar	unklar	unsicher
8	möglicherweise	möglicherweise	unsicher	fraglich
9	unklar	unsicher	möglicherweise	möglicherweise
10	oft	oft	typischerweise	oft
11	wahrscheinlich	häufig	überwiegend	häufig
12	häufig	überwiegend	häufig	überwiegend
13	überwiegend	wahrscheinlich	oft	wahrscheinlich
14	meist	typischerweise	wahrscheinlich	typischerweise
15	normalerweise	meist	normalerweise	normalerweise
16	typischerweise	normalerweise	meist	meist
17	fraglos	fraglos	fraglos	fraglos
18	stets	stets	stets	sicher
19	sicher	sicher	sicher	stets
20	immer	immer	immer	immer

Es wurde die ideale Reihe nun jeweils mit den Rangreihen der Berufsgruppen verglichen. Unter Verwendung des **Spearman'schen Rangkorrelationskoeffizienten** sollte die Stärke des monotonen Zusammenhangs zwischen den paarweise verglichenen Reihen überprüft werden (Tabelle 17).

Für die Apotheker erhält man z.B. ein ρ von 0,932. Mit Hilfe der Prüfgröße t nach Student konnte gezeigt werden, dass dieser Wert signifikant ist (siehe Tabelle 17). Somit ist die Nullhypothese, es bestehe kein Zusammenhang mit der idealen Reihe, zugunsten der Alternativhypothese bei einer Irrtumswahrscheinlichkeit von 0,05 zu verwerfen. Die

Werte für die Ärzte und Apotheker sind ebenfalls signifikant. Auch hier muss die Nullhypothese verworfen werden.

Tabelle 17: Spearman'scher Rangkorrelationskoeffizient ρ der einzelnen Berufsgruppe mit Angabe des t- Werts

	Spearman's ρ	t-Wert	Signifikanz von ρ
Ärzte	0,973	17,9	$p < 0,0005$
Apotheker	0,932	10,9	$p < 0,0005$
Juristen	0,976	19,0	$p < 0,0005$

Der Spearman'sche Rangkorrelationskoeffizient ρ gibt die Enge des monotonen Zusammenhang zwischen zwei Rangreihen an. Zur Signifikanztestung wird die Prüfgröße t nach Student verwendet.

Des Weiteren wurden die Rangreihen der Berufsgruppen mit dem **Rangkorrelationskoeffizienten Kendall τ** mit der idealen Reihe verglichen. Der Rangkorrelationskoeffizient Kendall τ kann Werte zwischen +1 (perfekte Korrelation) und -1 (negative Korrelation) annehmen.

Tabelle 18: Rangkorrelationskoeffizient Kendall τ der einzelnen Berufsgruppen mit Angabe der z- und p-Werte

	Kendall τ	z-Werte	p-Werte
Ärzte	0,895	5,517	$p < 0,001$
Apotheker	0,800	4,932	$p < 0,001$
Juristen	0,895	5,517	$p < 0,001$

Kendall τ ist ebenfalls ein Rangkorrelationskoeffizient der auf der Fehlordnung von Rangreihen beruht. Zur Signifikanztestung werden die z- und p- Werte ermittelt.

Zum Beispiel zeigen sich im Vergleich der idealen mit der echten Rangreihe der Ärzte +180 Proversionen und -10 Inversionen. Bei einer maximal möglichen Summe von 190 ergibt sich ein τ von 0,895. Bei einem z-Wert von 5,517 und einem $p < 0,001$ kann die Nullhypothese zugunsten der Alternativhypothese verworfen werden. Dies trifft auch für die beiden anderen Berufsgruppen zu. Somit ist eine Ähnlichkeit zwischen den jeweiligen Rangreihen und der idealen Reihe zu unterstellen.

Um nun zu prüfen, ob die neuen Rangreihen zufällig so angeordnet sind, wurde der **Permutationstest** durchgeführt. Wiederum wurden die Rangreihen der drei Berufsgruppen mit der idealen Rangreihe der Literatur verglichen. Die Permutation ist dabei die Summe der Differenzen der Rangplätze. Nimmt man beispielsweise den Begriff „selten“. In der idealen Reihe hat dieser den Rangplatz 2 und bei den Ärzten Platz 3. Daraus ergibt sich eine Differenz von 1. Für alle Begriffe wurden nun die Differenzen der Rangplätze gewählt und anschließend die Summe aller Differenzen gebildet. Die einzelnen Permutationen sind im Anhang (siehe Tabelle 55) dargestellt.

Tabelle 19 Permutationen der berufsgruppenspezifischen Rangreihen

	Permutation	SD	Mittelwert (AV)	T
Ärzte	18	19,41	133	-5,93
Apotheker	30	19,41	133	-5,31
Juristen	18	19,41	133	-5,93

Permutation: Summe der Differenzen der Rangplätze zwischen 2 Rangreihen. Die Standardabweichung (SD), der Mittelwert (AV) und der T- Wert dienen zur Signifikanztestung.

Bei den Ärzten ergibt sich so beispielsweise eine Permutation von 18. Hieraus kann ein T- Wert von -5,93 errechnet werden. Die Nullhypothese kann bei einer Irrtumswahrscheinlichkeit von 0,05 abgelehnt werden. Das bedeutet, dass die 20 Wahrscheinlichkeitsbegriffe nicht zufällig angeordnet sind. Diese Aussage trifft für alle Berufsgruppen zu.

3.4 Vergleich von Wahrscheinlichkeitsbegriffen „kontextfrei“ versus „im Kontext Nebenwirkungen“

In Abschnitt B sollten die drei Wahrscheinlichkeitsbegriffe „häufig“, „gelegentlich“ und „selten“, die im Zusammenhang mit Nebenwirkungen klar von der Pharmaindustrie definiert sind, numerisch interpretiert werden.

In diesem Abschnitt werden die Lage- und Streuungsmaße für die Sachverhalte „häufig“, „gelegentlich“ und „selten“ für alle Teilnehmer insgesamt und anschließend berufsgruppenspezifisch dargestellt.

Graphisch werden dann die drei Wahrscheinlichkeitsbegriffe aus Abschnitt A den Begriffen aus Abschnitt B gegenübergestellt. Dies erfolgt nach Berufsgruppen getrennt.

Danach werden die Begriffe „häufig“, „gelegentlich“ und „selten“ aus Abschnitt A mit den numerischen Interpretationen aus Abschnitt B anhand des Wilcoxon-Tests verglichen. Mit dem Kruskal-Wallis-Test wird anschließend untersucht, ob sich die drei Berufsgruppen in der Verteilung der Prozentwerte aus Abschnitt A und B unterscheiden.

Zum Schluss werden jeweils zwei Berufsgruppen mit dem Mann-Whitney-Test verglichen, um herauszufinden, welche Berufsgruppen sich voneinander unterscheiden.

3.4.1 Numerische Interpretation von „häufig“, „gelegentlich“ und „selten“ im Kontext Nebenwirkungen

Zunächst erfolgt die Auswertung aller Teilnehmer (siehe Tabelle 20). So haben insgesamt 425 Probanden den Begriff „häufig“ im Abschnitt B numerisch interpretiert. 50% der Probanden gaben „häufig“ einen Wert zwischen 40 und 75%. Der Median liegt bei 60%. „Selten“ wurde im Kontext „Nebenwirkungen“ von 50% der Probanden mit

einem Wert zwischen 1 und 10% belegt. Minimum war 0,01%, während das Maximum bei 95% lag.

Tabelle 20 gewählte Prozentwerte der Begriffe „häufig“, „gelegentlich“ und „selten“ im Kontext Nebenwirkungen. Alle Probanden sind zusammengefasst.

	n	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
häufig	425	60	1,00	99	40	75
gelegentlich	425	20	0,10	80	9	25
selten	423	5	0,01	95	1	10

n= Anzahl der Antworten (Probanden); Median, Minimum, Maximum und Quantil angeben in Prozentwerten

Nach Berufsgruppen unterteilt ergeben sich beispielsweise folgende Ergebnisse: Bei den Apothekern liegt der Minimalwert von „häufig“ bei 1%. Maximalwert ist 95%. 50% der Apotheker vergaben einen Wert zwischen 20 und 50%.

Tabelle 21 Prozentwerte mit Streuungsmaßen für „häufig“ interpretiert im Sachverhalt Nebenwirkungen.

	n	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
Ärzte	180	65	5	99	30	75
Apotheker	104	50	1	95	20	75
Juristen	140	70	3	95	50	80

n= die Anzahl der Teilnehmer. Median, Minimum, Maximum und Quantil angegeben in Prozent.

Bei der Interpretation von „gelegentlich“ haben die Ärzte und Apotheker ähnliche Medianwerte erreicht. Hier ähneln sich die 25%- und 75%- Quantile sehr. Die Juristen weisen hier deutlich höhere Werte auf (siehe Tabelle 22).

Tabelle 22 Prozentwerte mit Streuungsmaßen für „gelegentlich“ interpretiert im Sachverhalt Nebenwirkungen

	n	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
Ärzte	181	10	0,55	80	5	20
Apotheker	104	10	0,10	55	5	23,75
Juristen	139	20	1,00	75	15	30

n= die Anzahl der Teilnehmer. Median, Minimum, Maximum und Quantil angegeben in Prozent.

Schaut man sich die Interpretationen von „selten“ an, so fällt der niedrigere Median der Apotheker von 2,75% auf. Die Juristen haben „selten“ durchschnittlich einen höheren Prozentwert gegeben. Dies spiegelt sich auch in den Quantilen wieder (Tabelle 23).

Tabelle 23 Prozentwerte mit Streuungsmaßen für „selten“ interpretiert im Sachverhalt Nebenwirkungen.

	n	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
Ärzte	179	5	0.06	55	1	5
Apotheker	104	2.75	0.01	95	1	5
Juristen	139	5	0.01	55	5	10

n= die Anzahl der Teilnehmer. Median, Minimum, Maximum und Quantil angegeben in Prozent.

3.4.2 Vergleich der kontextfreien und kontextabhängigen Interpretation von „häufig“, „gelegentlich“ und „selten“

Für jede einzelne Person werden die numerischen Interpretationen der drei Wahrscheinlichkeitsbegriffe „häufig“, „gelegentlich“ und „selten“ paarweise aus Abschnitt A und B verglichen. So haben z.B. 66 Ärzte „gelegentlich“ im Abschnitt A und B mit dem gleichen Prozentwert belegt, während 73 Ärzte im Kontext einen niedrigeren und 38 Ärzte einen höheren Wert vergeben haben. Am häufigsten wurde der jeweilige Begriff im Kontext niedriger bewertet als ohne Kontext.

Tabelle 24: Vergleich der Begriffe im Kontext versus kontextfrei der Ärzte

Vergleich der Ärzte		Anzahl
"häufig"	Kontext häufig < häufig	89
	Kontext häufig > häufig	29
	Kontext häufig = häufig	58
	Gesamt:	176
"gelegentlich"	Kontext gelegentlich < gelegentlich	73
	Kontext gelegentlich > gelegentlich	38
	Kontext gelegentlich = gelegentlich	66
	Gesamt:	177
"selten"	Kontext selten < selten	66
	Kontext selten > selten	36
	Kontext selten = selten	75
	Gesamt:	177

Tabelle 25: Vergleich der Begriffe im Kontext versus kontextfrei **der Apotheker**

Vergleich der Apotheker		Anzahl
"häufig"	Kontext häufig < häufig	65
	Kontext häufig > häufig	13
	Kontext häufig = häufig	26
	Gesamt	104
"gelegentlich"	Kontext gelegentlich < gelegentlich	51
	Kontext gelegentlich > gelegentlich	21
	Kontext gelegentlich = gelegentlich	32
	Gesamt	104
"selten"	Kontext selten < selten	51
	Kontext selten > selten	13
	Kontext selten = selten	40
	Gesamt:	104

Tabelle 26: Vergleich der Begriffe im Kontext versus kontextfrei **der Juristen**

Vergleich der Juristen		Anzahl
"häufig"	Kontext häufig < häufig	61
	Kontext häufig > häufig	25
	Kontext häufig = häufig	55
	Gesamt	141
"gelegentlich"	Kontext gelegentlich < gelegentlich	47
	Kontext gelegentlich > gelegentlich	35
	Kontext gelegentlich = gelegentlich	58
	Gesamt	140
"selten"	Kontext selten < selten	44
	Kontext selten > selten	29
	Kontext selten = selten	67
	Gesamt	140

3.4.3 Graphische Gegenüberstellung der kontextfreien und der kontextabhängigen Interpretationen von „häufig“, „gelegentlich“ und „selten“

3.4.3.1 Gesamtes Studienkollektiv

Es werden die Histogramme der Interpretationen der Begriffe „häufig“, „gelegentlich“ und „selten“ aus Abschnitt A den kontextgebundenen Interpretationen aus Abschnitt B gegenübergestellt. Zunächst erfolgt die Darstellung zusammenfassend für alle Teilnehmer.

Abbildung „häufig“

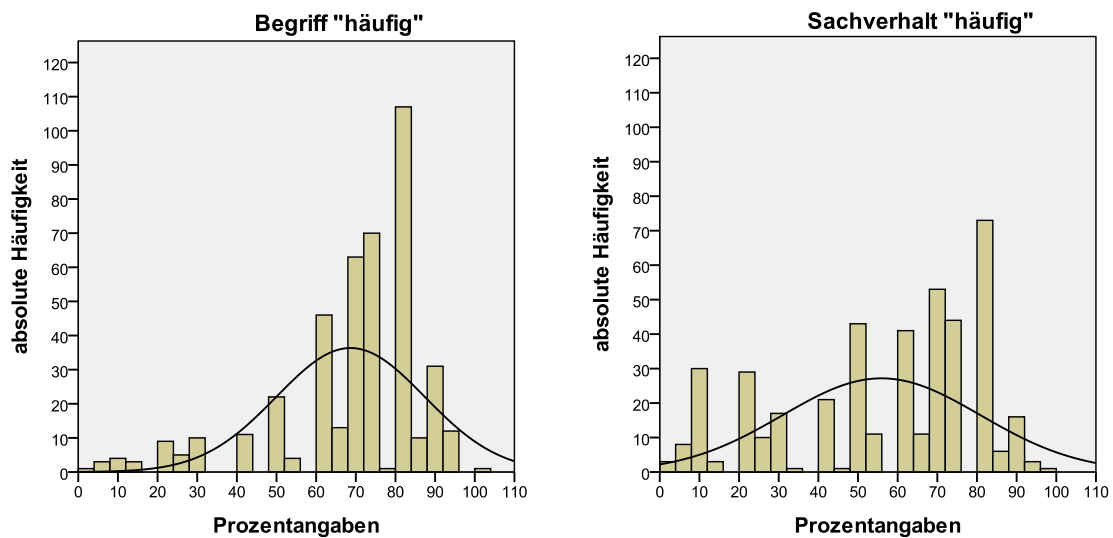


Abbildung 1: Anzahl der Nennungen von „häufig“. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung „gelegentlich“

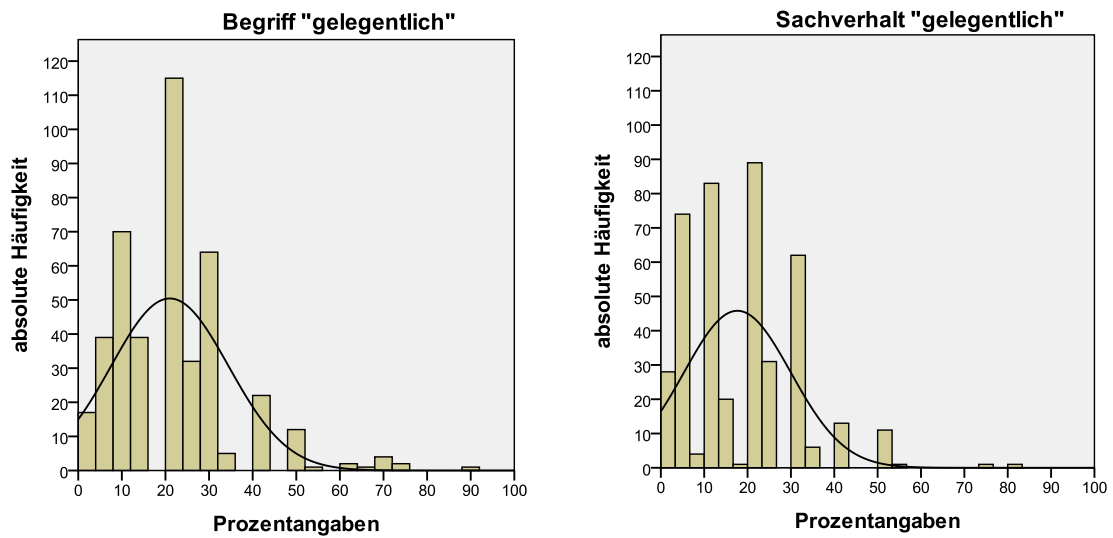


Abbildung 2: Anzahl der Nennungen von „gelegentlich“. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x- Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung „selten“

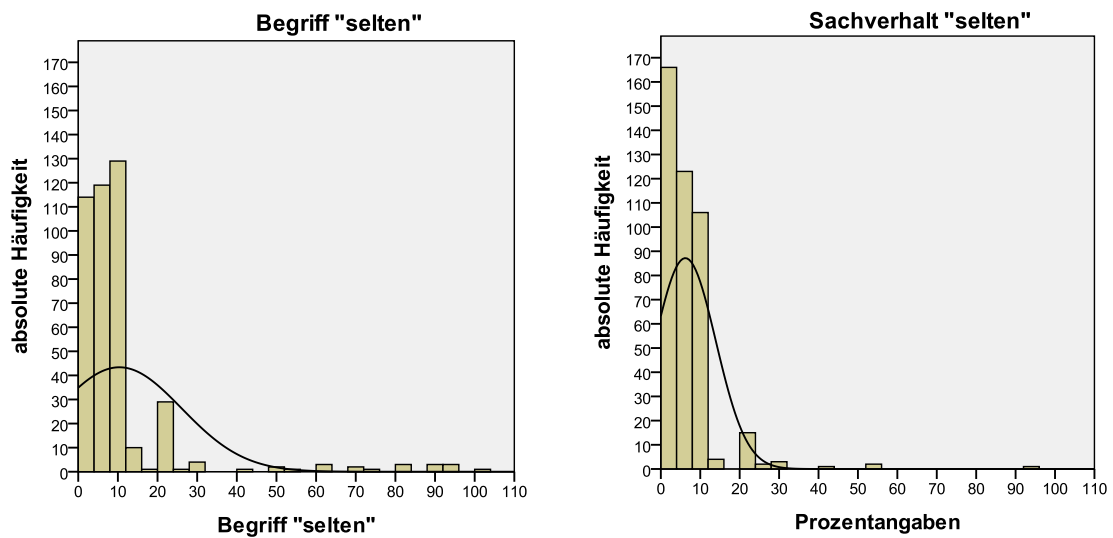


Abbildung 3:Anzahl der Nennungen von „selten“. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

3.4.3.2 Berufsgruppenspezifische Darstellung

Nachfolgend werden die Ergebnisse graphisch nach Berufsgruppen unterteilt dargestellt.

Anzahl der Nennungen der Prozentwerte von „häufig“

Abbildung: Ärzte

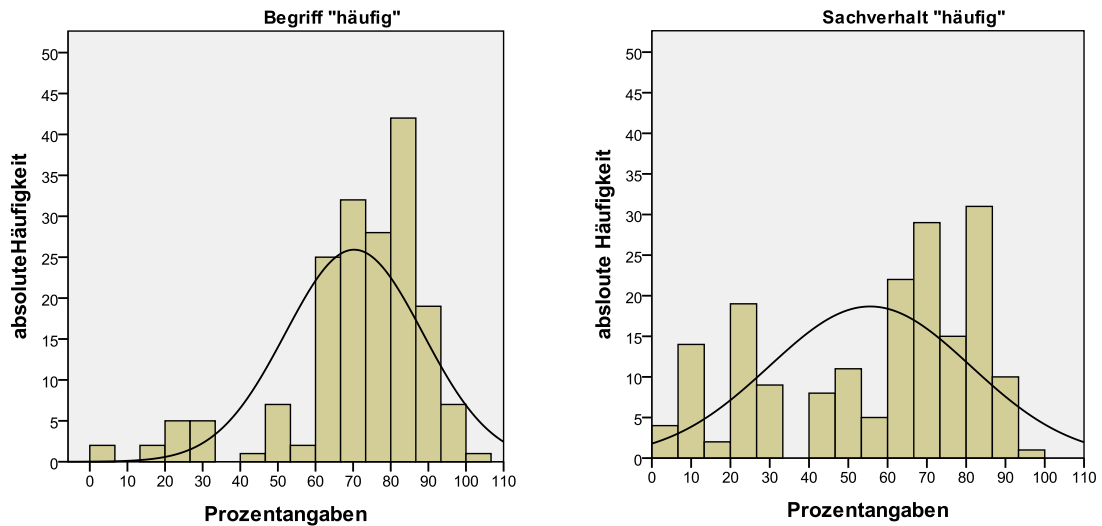


Abbildung 4: Anzahl der Nennungen des Begriffs „häufig“ in der Gruppe der **Ärzte**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x- Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung: Apotheker

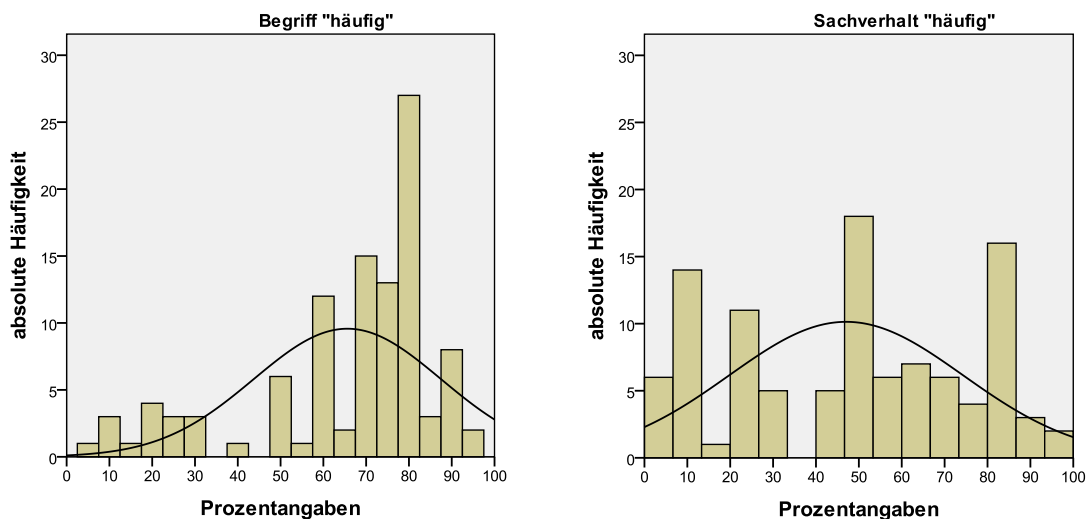


Abbildung 5: Anzahl der Nennungen von „häufig“ in der Gruppe der **Apotheker**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x- Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung: Juristen

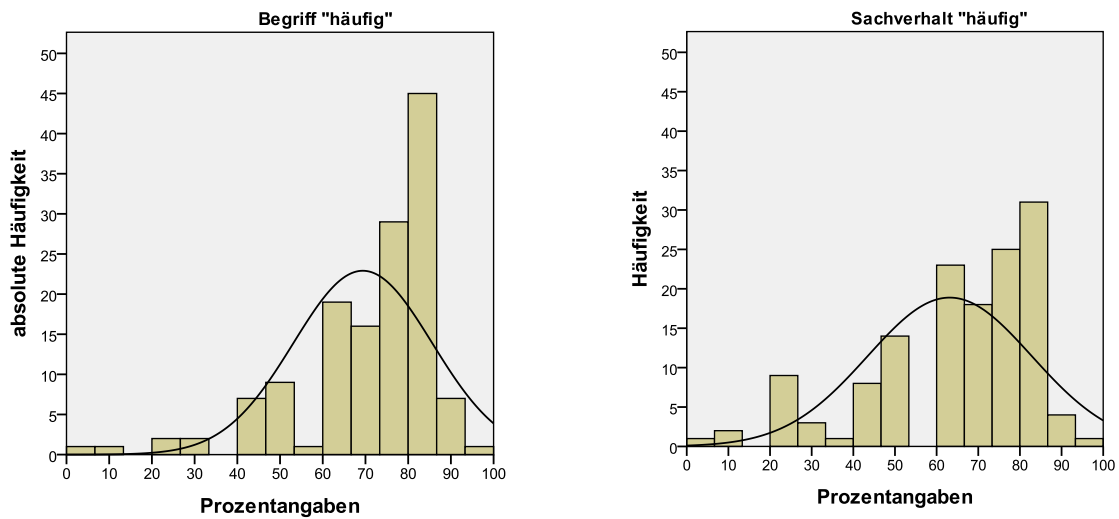


Abbildung 6: Anzahl der Nennungen von „häufig“ in der Gruppe der **Juristen**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x- Achse sind die numerischen Interpretationen in Prozent angegeben.

Anzahl der Nennungen der Prozentwerte von „gelegentlich“

Abbildung: Ärzte

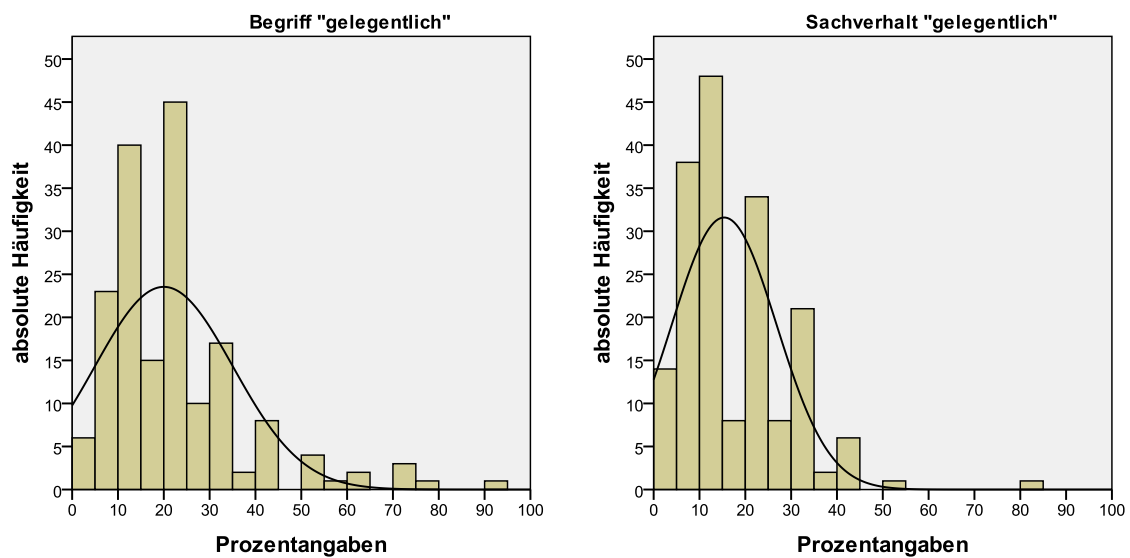


Abbildung 7: Anzahl der Nennungen von „gelegentlich“ in der Gruppe der **Ärzte**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x- Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung: Apotheker

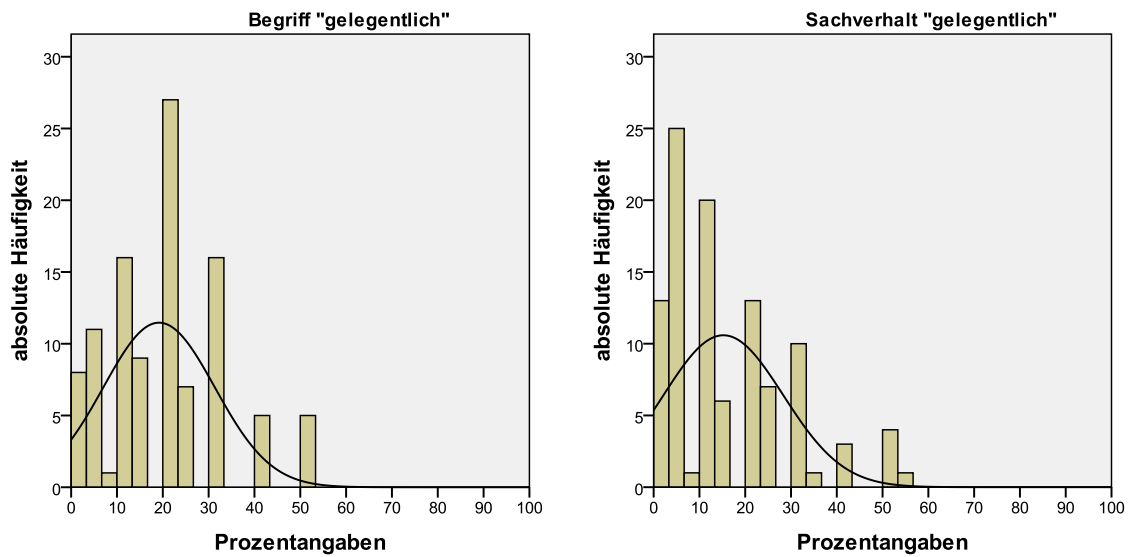


Abbildung 8: Anzahl der Nennungen von „gelegentlich“ in der Gruppe der **Apotheker**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung: Juristen

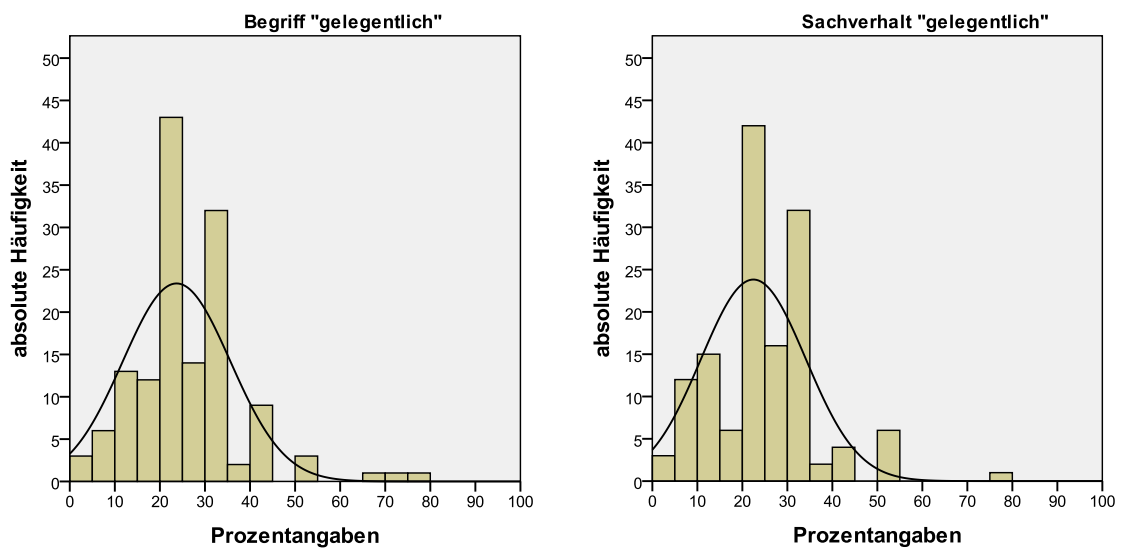


Abbildung 9: Anzahl der Nennungen von „gelegentlich“ in der Gruppe der **Juristen**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

Anzahl der Nennungen der Prozentwerte von „selten“

Abbildung: Ärzte

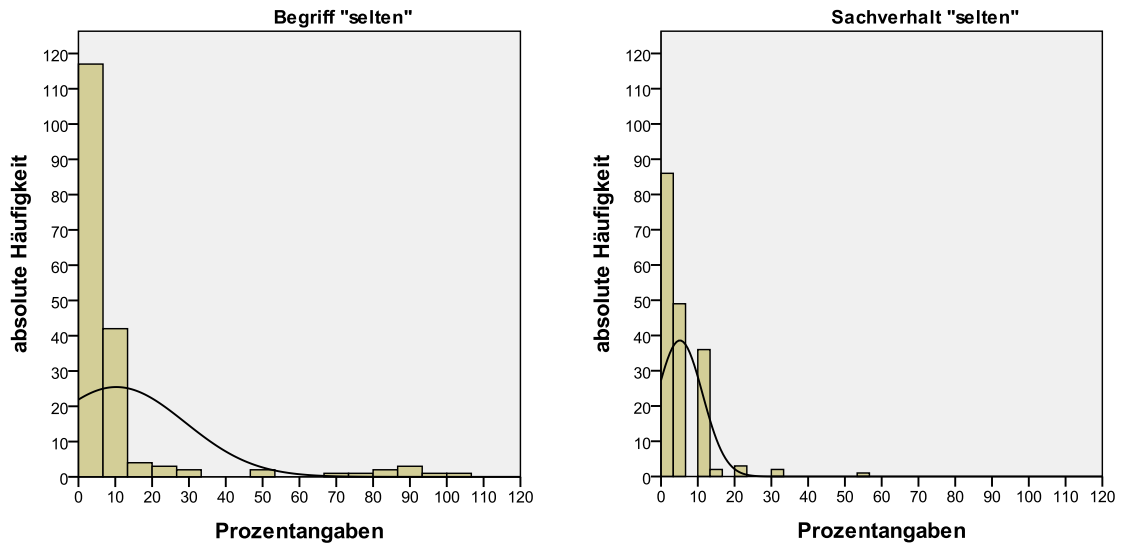


Abbildung 10: Anzahl der Nennungen von „selten“ in der Gruppe der **Ärzte**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung: Apotheker

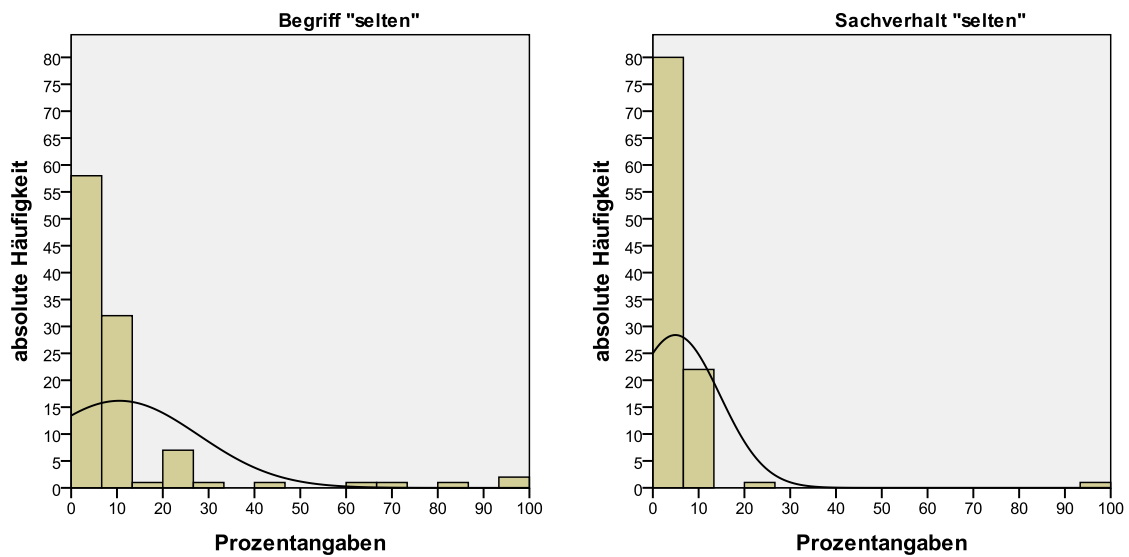


Abbildung 11: Anzahl der Nennungen von „selten“ in der Gruppe der **Apotheker**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung: Juristen

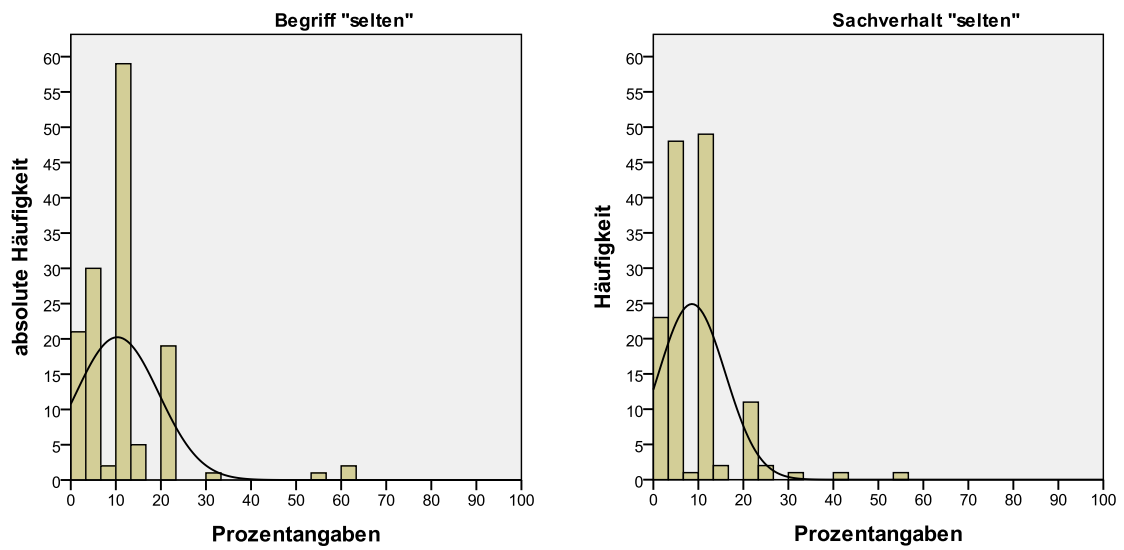


Abbildung 12: Anzahl der Nennungen von „selten“ in der Gruppe der **Juristen**. Links ohne Kontext; rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x- Achse sind die numerischen Interpretationen in Prozent angegeben.

3.4.4 Besteht ein kontextabhängiger Unterschied in der Interpretation von Wahrscheinlichkeitsbegriffen?

Zur Beantwortung dieser Frage wird zunächst der **Wilcoxon- Test** verwendet.

Beispielsweise wird für den Begriff „häufig“ aus dem Abschnitt A und B für jeden Studienteilnehmer die Differenz der beiden Prozentwerte ermittelt. Aus diesen Differenzen wird dann die Mediandifferenz gebildet. Des Weiteren wird der **Hodges-Lehmann-Punktschätzer** als Median der Differenzen und das dazugehörige 95%-Konfidenzintervall angegeben.

Tabelle 27: Vergleich der Variablen „häufig“, „gelegentlich“ und „selten“ im Kontext versus ohne Kontext interpretiert. Dargestellt sind die Mediane aus Abschnitt A und B, sowie die Mediandifferenzen mit Konfidenzintervall (KI) jeweils in Prozent und Angabe des p- Werts

	häufig			häufig	
	Kontext – ohne Kontext		p-Wert	ohne Kontext	Kontext
	Mediandifferenz	95% KI	p-Wert	Median	Median
Ärzte	-22.50	-27.50; -15.50	$5.5 * 10^{-12}$	75.00	65.00
Apotheker	-25.00	-32.50; -17.50	$2.3 * 10^{-10}$	75.00	50.00
Juristen	-10.00	-15.00; -5.00	$3.5 * 10^{-4}$	75.00	70.00
Gesamt	-20.00	-22.50; -15.00	$8.3 * 10^{-23}$	75.00	60.00

	gelegentlich			gelegentlich	
	Kontext – ohne Kontext		p-Wert	ohne Kontext	Kontext
	Mediandifferenz	95% KI	p-Wert	Median	Median
Ärzte	-6,00	-9.50; -2.50	$7.0 * 10^{-5}$	20.00	10.00
Apotheker	-7.00	-10.00; -2.50	$1.8 * 10^{-3}$	20.00	10.00
Juristen	-2.50	-6.25; 0.00	$1.8 * 10^{-1}$	20.00	20.00
Gesamt	-5.00	-7.50; -2.50	$6.0 * 10^{-7}$	20.00	20.00

	selten			selten	
	Kontext – ohne Kontext		p-Wert	ohne Kontext	Kontext
	Mediandifferenz	95% KI	p-Wert	Median	Median
Ärzte	-3.00	-4.55; -1.53	$8.9 * 10^{-5}$	5.00	5.00
Apotheker	-5.50	-7.50; -3.50	$5.1 * 10^{-8}$	5.00	2.75
Juristen	-2.50	-4.95; -0.05	$1.3 * 10^{-2}$	10.00	5.00
Gesamt	-3.75	-4.75; -2.50	$1.9 * 10^{-11}$	5.00	5.00

Zur Berechnung des p- Werts wurde der Wilcoxon- Test verwendet. Die Mediandifferenz entspricht dem Hodges- Lehmann-Punktschätzer.

So ergibt sich beispielsweise für „häufig“ ein Wert von -10 bei den Juristen. Das negative Vorzeichen bedeutet, dass die Prozentwerte für „häufig“ im Kontext niedriger sind als ohne Kontext. Um auf die Grundgesamtheit schließen zu können, wird das

95%-Konfidenzintervall mit angegeben. So läge der Wert der Juristen zu 95% zwischen -15 und -5.

Alle Mediandifferenz-Werte liegen im jeweiligen Konfidenzintervall. Vergleicht man z.B. „häufig“ aus Abschnitt B mit „häufig“ aus Abschnitt A für alle Berufsgruppen insgesamt, ergibt sich ein p-Wert von $8.3 * 10^{-23}$, d.h. der Unterschied ist signifikant und die Nullhypothese ist zugunsten der Alternativhypothese abzulehnen.

Mit einer Ausnahme sind alle Ergebnisse des Wilcoxon-Tests als signifikant anzusehen. Somit werden die Begriffe im Kontext niedriger interpretiert als ohne Kontext. Lediglich bei den Juristen liegt der p-Wert von „gelegentlich“ bei $1.8 * 10^{-1}$ und ist zum Signifikanzniveau $\alpha=0,05$ nicht signifikant. Für „gelegentlich“ gibt es also keine Kontextabhängigkeit in der Berufsgruppe der Juristen.

3.4.5 Unterscheiden sich die drei Berufsgruppen hinsichtlich der Verteilung der Prozentwerte.

Es wurde nun gezeigt, dass die numerische Interpretation vom Kontext abhängt. Nun stellt sich die Frage, ob sich die drei Berufsgruppen in der numerischen Interpretation aus Abschnitt A und B unterscheiden. Hierfür wird der **Kruskal- Wallis- Test** für k unabhängige Stichproben durchgeführt.

Zunächst werden die Prozentangaben der einzelnen Berufsgruppen miteinander verglichen. Beim Begriff „selten“ ergeben sich zum Beispiel p-Werte für „selten“ im Kontext von $2,9*10^{-10}$. Folglich muss die Nullhypothese, es bestehe kein Unterschied zwischen den einzelnen Gruppen, zugunsten der Alternativhypothese verworfen werden. Ein Unterschied ist also anzunehmen.

Für „häufig“ ohne Kontext ergibt sich demnach mit $p = 4.2*10^{-1}$ kein signifikanter Unterschied zwischen den Ärzten, Apothekern und Juristen. Bei allen anderen Wahrscheinlichkeitsbegriffen zeigen die p-Werte (Tabelle 28) einen signifikanten Unterschied.

Tabelle 28 p-Werte des Kruskal-Wallis-Tests für die Wahrscheinlichkeitsbegriffe aus Abschnitt A und B

Wahrscheinlichkeitsbegriff	p-Wert
„häufig“ (Kontext)	$3.4 * 10^{-5}$
„häufig“ (ohne Kontext)	$4.2 * 10^{-1}$
„gelegentlich“ (Kontext)	$2.8 * 10^{-9}$
„gelegentlich“ (ohne Kontext)	$1.2 * 10^{-4}$
„selten“ (Kontext)	$2.9 * 10^{-10}$
„selten“ (ohne Kontext)	$1.2 * 10^{-5}$

Es wurde der asymptotische p- Wert berechnet.

3.4.6 Welche Berufsgruppen unterscheiden sich in der Vergabe der Prozentangaben?

Um jetzt herauszufinden, welche Berufsgruppen sich unterscheiden, wird der **Mann-Whitney-Test** durchgeführt. Es werden jeweils zwei Berufsgruppen (Mediziner versus Apotheker, Mediziner versus Juristen sowie Apotheker versus Juristen) miteinander verglichen.

3.4.6.1 Vergleich der Mediziner mit den Apothekern

Im direkten Vergleich der Ärzte und Apotheker finden sich nicht signifikante p- Werte für die untersuchten Begriffe (siehe Tabelle 29). Somit muss die Nullhypothese, es besteht kein Unterschied zwischen den numerischen Interpretationen zwischen Ärzten und Apothekern, angenommen werden. Einzige Ausnahme ist der Begriff „häufig“ im Kontext interpretiert. Hier findet sich ein signifikanter p- Wert. Ein Unterschied in der Interpretation des Begriffs ist also anzunehmen.

Tabelle 29: Vergleich der Mediziner mit den Apothekern bezüglich der Variablen „häufig“, „gelegentlich“ und „selten“. Zugrunde liegt der Mann-Whitney- Tests.

		Mediziner - Apotheker		
		Mediandifferenz	95%-KI	p-Wert
häufig	(ohne Kontext)	0.00	0.00; 5.00	$1.9 * 10^{-1}$
gelegentlich	(ohne Kontext)	0.00	-4.00; 0.00	$7.9 * 10^{-1}$
selten	(ohne Kontext)	0.00	-0.90; 0.00	$4.99 * 10^{-1}$
häufig	(Kontext)	10.00	0.00; 15.00	$8.6 * 10^{-3}$
gelegentlich	(Kontext)	0.00	0.00; 4.00	$4.2 * 10^{-1}$
Selten	(Kontext)	0.00	0.00; 0.90	$1.6 * 10^{-1}$

3.4.6.2 Vergleich der Mediziner mit den Juristen

Tabelle 30: Vergleich der Mediziner mit den Juristen bezüglich der Variablen „häufig“, „gelegentlich“ und „selten“. Zugrunde liegt der Mann-Whitney- Tests.

		Mediziner - Juristen		
		Mediandifferenz	95%-CI	p-Wert
häufig	(ohne Kontext)	0.00	0.00; 5.00	$5.7 * 10^{-1}$
gelegentlich	(ohne Kontext)	-5.00	-10.00; -1.50	$4.2 * 10^{-5}$
selten	(ohne Kontext)	-3.00	-5.00; -1.00	$1.0 * 10^{-6}$
häufig	(Kontext)	-5.00	-10.00; 0.00	$1.9 * 10^{-2}$
gelegentlich	(Kontext)	-10.00	-10.00; 5.00	$5.4 * 10^{-9}$
Selten	(Kontext)	-4.00	-4.00; -2.00	$1.8 * 10^{-9}$

Bis auf „häufig“ ohne Kontext sind beim Vergleich der Mediziner mit den Juristen alle p-Werte signifikant. Die Nullhypothese muss dementsprechend abgewiesen werden und ein Unterschied zwischen diesen Gruppen ist anzunehmen.

3.4.6.3 Vergleich der Apotheker mit den Juristen

Auch hier errechnen sich bis auf den Begriff „häufig“ ohne Kontext durchweg signifikante p- Werte und somit ist auch hier ein Unterschied zwischen den Apothekern und Juristen anzunehmen.

Tabelle 31: Vergleich der Apotheker mit den Juristen bezüglich der Variablen „häufig“, „gelegentlich“ und „selten“. Zugrunde liegt der Mann- Whitney- Tests.

		Apotheker - Juristen		
		Mediandifferenz	95%-CI	p-Wert
häufig	(ohne Kontext)	0.00	-5.00; 0.00	$4.5 * 10^{-1}$
gelegentlich	(ohne Kontext)	-5.00	-10.00; 0.00	$1.7 * 10^{-3}$
selten	(ohne Kontext)	-2.00	-4.50; 0.00	$3,0 * 10^{-3}$
häufig	(Kontext)	-15.00	-22.50; -10.00	$4.6 * 10^{-6}$
gelegentlich	(Kontext)	-10.00	-10.00; -5.00	$3,0 * 10^{-7}$
Selten	(Kontext)	-4.00	-4.90; -2.50	$5.9 * 10^{-10}$

3.4.7 Unterscheiden sich die drei Berufsgruppen in den Differenzen zwischen der kontextfreier und –abhängiger Interpretation?

Mit dem Wilcoxon- Test wurde gezeigt, dass sich eine Kontextabhängigkeit in der Interpretation besteht. Im Kontext Nebenwirkungen werden niedrigere Prozente vergeben. Nun wird geprüft, ob sich die Berufsgruppen bezüglich der Differenzen zwischen A und B unterscheiden. Nachfolgend sind die Ergebnisse des **Kruskal-Wallis- Test** dargestellt.

Tabelle 32: p-Werte des Kruskal-Wallis-Tests für die neu gebildeten Differenzvariablen von „häufig“, „gelegentlich“ und „selten“

Wahrscheinlichkeitsbegriff	p-Wert
häufig (Kontext) – häufig (ohne Kontext)	$1.9 * 10^{-3}$
gelegentlich (Kontext) – gelegentlich (ohne Kontext)	$1.1 * 10^{-1}$
selten (Kontext) – selten (ohne Kontext)	$1.6 * 10^{-2}$

Für „häufig“ und „selten“ wird die Nullhypothese bei einem α kleiner 0,05 verworfen. Die Berufsgruppen unterscheiden sich also in der Differenz zwischen A und B.

Für den Begriff „gelegentlich“ muss die Nullhypothese allerdings bei einem p- Wert von $1,1 * 10^{-1}$ angenommen werden. Es besteht somit zwischen den Berufsgruppen kein signifikanter Unterschied in den Differenzen.

3.4.8 Welche Berufsgruppen unterscheiden sich in den Differenzen zwischen kontextfreier und –abhängiger Interpretation?

Dass sich die Berufsgruppen in den Differenzen für „selten“ und „häufig“ unterscheiden, konnte nun gezeigt werden. Allerdings kann bis jetzt noch keine Aussage getroffen werden, welche Berufsgruppen sich unterscheiden. Für den paarweisen Vergleich wird erneut der Mann-Whitney-Test verwendet.

3.4.8.1 Vergleich der Mediziner und Apotheker

Für die Begriffe „häufig“ und „gelegentlich“ liegen keine signifikanten p- Werte vor, so dass hier kein Unterschied zwischen den Differenzen besteht. Für den Begriff „selten“ ist allerdings ein signifikanter Unterschied anzunehmen.

Tabelle 33: Vergleich der Mediziner mit den Apotheker bezüglich der Differenzen zwischen kontextfreier und –abhängiger Interpretation.

	Mediziner - Apotheker		
	Mediandifferenz	95%-CI	p-Wert
häufig	0.00	0.00; 10.00	$1.3 * 10^{-1}$
gelegentlich	0.00	0.00; 3.00	$4.8 * 10^{-1}$
selten	0.00	0.00; 1.50	$2.6 * 10^{-2}$

3.4.8.2 Vergleich der Mediziner und Juristen

Für „häufig“ ist ein signifikanter Unterschied anzunehmen. Die p- Werte für die Begriffe „ gelegentlich“ und „selten“ sind nicht signifikant.

Tabelle 34: Vergleich der Mediziner mit den Juristen bezüglich der Differenzen zwischen kontextfreier und –abhängiger Interpretation.

	Mediziner - Juristen		
	Mediandifferenz	95%-CI	p-Wert
häufig	0.00	-10.00; 0.00	$2.8 * 10^{-2}$
gelegentlich	0.00	-3.00; 0.00	$1.3 * 10^{-1}$
selten	0.00	0.00; 0.00	$4.6 * 10^{-1}$

3.4.8.3 Vergleich der Apotheker und Juristen

Die p- Werte für die alle drei Begriffe sind signifikant. Es besteht somit im Vergleich zwischen Apothekern und Juristen ein signifikanter Unterschied in der Differenz.

Tabelle 35: Vergleich der Apotheker mit den Juristen bezüglich der Differenzen zwischen kontextfreier und –abhängiger Interpretation.

	Apotheker - Juristen		
	Mediandifferenz	95%-CI	p-Wert
häufig	-10.00	-15.00; 0.00	$3.3 * 10^{-4}$
gelegentlich	0.00	-5.00; 0.00	$4.7 * 10^{-2}$
selten	0.00	-2.00; 0.00	$5.4 * 10^{-3}$

Zusammenfassend besteht somit zwischen den Apothekern und Juristen für diese drei Begriffe der größte Unterschied.

3.5 Werden Wahrscheinlichkeitsbegriffe im Kontext „Rechtsstreit“ anders interpretiert?

Im Abschnitt D sollten die Begriffe „möglicherweise“, „oft“ und „wahrscheinlich“ im Kontext interpretiert werden. Hier bezog sich dieser jedoch nicht auf Nebenwirkungen, sondern auf einen Rechtsstreit.

Zunächst werden wiederum die Lage- und Streuungsmaße der numerischen Interpretationen der drei Begriffe dargestellt. Die graphische Darstellung der Ergebnisse erfolgt im Anhang. Mit dem Wilcoxon- Test werden dann die Prozentwerte der drei Begriffe aus Abschnitt A und D verglichen.

3.5.1 Numerische Interpretationen von „möglicherweise“, „oft“ und „wahrscheinlich“ im Kontext Rechtsstreit

Als Fallbeispiel diene ein Rechtsstreit wegen eines Behandlungsfehlers. Die in die Rolle des Rechtsanwalts versetzten Probanden vergaben für „möglicherweise“ einen Minimalwert von 0,5%, während das Maximum bei 90% lag.

Tabelle 36 gewählte Prozentwerte für die Begriffe „möglicherweise“, „oft“ und „wahrscheinlich“. Alle Probanden sind zusammengefasst.

	n	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
möglicherweise	406	40	0.50	90	25	50
oft	408	70	1.00	95	60	80
wahrscheinlich	415	80	5.00	100	60	90

n = Anzahl der Antworten und Median, Minimum, Maximum und Quantil (angeben in Prozentwerten)

Schaut man sich nun die Ergebnisse nach Berufsgruppen unterteilt an, ergeben sich z.B. für „möglicherweise“ bei den Ärzten und Apothekern ein Median von 40% und bei den Juristen von 50%.

Tabelle 37: Prozentwerte mit Streuungsmaßen für „möglicherweise“ interpretiert im Kontext Rechtsstreit

	n	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
Ärzte	169	40	0.50	90	20	50
Apotheker	102	40	1.00	80	25	50
Juristen	135	50	5.00	80	30	50

n= die Anzahl der Teilnehmer, Angaben in Prozent

Tabelle 38: Prozentwerte mit Streuungsmaßen für „oft“ interpretiert im Kontext Rechtsstreit

	n	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
Ärzte	171	70	1	95	60	80
Apotheker	103	70	15	95	60	80
Juristen	134	70	30	90	60	76,3

n= die Anzahl der Teilnehmer, Angaben in Prozent

Tabelle 39: Prozentwerte mit Streuungsmaßen für „wahrscheinlich“ interpretiert im Kontext Rechtsstreit

	n	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
Ärzte	176	80	5	100	70	90
Apotheker	103	80	10	99	60	90
Juristen	136	75	25	95	60	80

n= die Anzahl der Teilnehmer, Angaben in Prozent

Bei der Interpretation von „oft“ zum Beispiel fällt auf, dass alle Berufsgruppen einen Median von 70% haben, aber die Minimal und Maximalwerte doch voneinander abweichen. So wurde bei den Ärzten ein Minimalwert von 1% und ein Maximalwert von 95% vergeben. Bei den Juristen lagen der minimal vergebene Wert bei 30% und der Maximalwert bei 90%. Zusammenfassend scheint die Mehrheit „oft“ im oberen Prozentbereich zu sehen (siehe auch Quantile). Die Extremwerte sind wiederum breit gefächert.

Die graphische Gegenüberstellung der Ergebnisse der Interpretationen aus Abschnitt A und D erfolgt im Anhang.

3.5.2 Unterscheiden sich die Prozentangaben von „möglicherweise“, „oft“ und „wahrscheinlich“ wenn sie im Kontext interpretiert werden?

Diese Frage wird mit dem **Wilcoxon-Test** untersucht. Der Test wird für das gesamte Studienkollektiv und für die drei Berufsgruppen einzeln durchgeführt. Wiederum werden der Hodges-Lehmann-Punktschätzer und das dazugehörige Konfidenzintervall gebildet. Es wird die Hypothese zum Signifikanzniveau von $\alpha=0,05$ getestet. Ist der p-Wert kleiner als das Signifikanzniveau, muss die Nullhypothese zugunsten der Alternativhypothese abgelehnt werden.

Tabelle 40: Vergleich von „möglicherweise“, „oft“ und „wahrscheinlich“ im Kontext versus ohne Kontext interpretiert. Dargestellt sind die Mediane aus Abschnitt A und D, sowie die Mediandifferenzen mit dem Konfidenzintervall (KI) jeweils in Prozent und Angabe des p-Werts

möglicherweise					
	Kontext – ohne Kontext			ohne Kontext	Kontext
	Mediandifferenz	95%-CI	p-Wert	Median	Median
Ärzte	7.50	2.50; 12.50	$5.4 * 10^{-3}$	30	40
Apotheker	7.50	0.00; 15.00	$2.7 * 10^{-2}$	30	40
Juristen	0.00	-3.00; 7.50	$6.1 * 10^{-1}$	40	50
Gesamt	5.00	2.50; 9.48	$9.1 * 10^{-4}$	35	40

oft					
	Kontext – ohne Kontext			ohne Kontext	Kontext
	Mediandifferenz	95%-CI	p-Wert	Median	Median
Ärzte	-5.00	-7.50;0.00	$7.2 * 10^{-2}$	70	70
Apotheker	0.00	-5.00;7.50	$9.1 * 10^{-1}$	70	70
Juristen	-1.25	-5.00;2.50	$5.8 * 10^{-1}$	75	70
Gesamt	-2.50	-5.00;0.00	$1.4 * 10^{-1}$	70	70

wahrscheinlich					
	Kontext – ohne Kontext			ohne Kontext	Kontext
	Mediandifferenz	95%-CI	p-Wert	Median	Median
Ärzte	0.00	-3.50; 5.00	$9.6 * 10^{-1}$	80	80
Apotheker	5.00	-2.50;10.00	$2.6 * 10^{-1}$	75	80
Juristen	0.00	-2.50; 7.00	$5.8 * 10^{-1}$	75	75
Gesamt	1.00	-1.00; 5.00	$3.0 * 10^{-1}$	75	80

Zur Berechnung des p- Werts wurde der Wilcoxon-Test verwendet. Die Mediandifferenz entspricht dem Hodges-Lehmann-Punktschätzer.

Für die Begriffe „oft“ und „wahrscheinlich“ liegen keine signifikanten p- Werte vor. Die Nullhypothese muss also angenommen werden. Dies bedeutet, dass zum Beispiel die Apotheker „wahrscheinlich“ im Kontext „Rechtsstreit“ genau den gleichen

Prozentwert geben wie ohne Kontext. Auch für die Juristen und Apothekern ergeben sich für diese zwei Begriffe keine signifikanten Begriffe.

Für „möglicherweise“ liegen signifikante Werte bei den Ärzten und Apothekern vor. Somit muss hier eine Abhängigkeit vom Kontext angenommen werden.

3.6 Vergleich mit der offiziellen Klassifikation für Arzneimittelnebenwirkungen

Nachfolgend sollen die Ergebnisse mit der offiziellen Richtlinie für Arzneimittelnebenwirkungen des Bundesinstituts für Arzneimittel und Medizinprodukte (siehe Tabelle 3) verglichen werden.

Es wird zunächst überprüft, ob „häufig“, „gelegentlich“ und „selten“ im Abschnitt B gemäß der Richtlinie richtig interpretiert wurden. Außerdem soll geklärt werden, ob es einen Unterschied zwischen den Berufsgruppen gibt.

In Abschnitt C des Fragebogens wurde erneut ein Sachverhalt entwickelt, der sich auf Nebenwirkung bezog. Hier waren numerische Wahrscheinlichkeiten wie 1:1000 oder 1:100 zu interpretieren. Die Wahrscheinlichkeitsbegriffe waren dabei frei wählbar. Einige dieser Verhältniszahlen sind in der offiziellen Klassifikation definiert. Somit können die Antworten der Probanden auf ihre Richtigkeit überprüft werden.

3.6.1 Vergleich der Prozentangaben von „häufig“, „gelegentlich“ und „selten“ mit der offiziellen Nebenwirkungstabelle

Die gewählten Prozentwerte aus Abschnitt A und B werden mit der offiziellen Klassifikation verglichen. Wird von den Studienteilnehmern für „häufig“ beispielsweise ein Prozentwert zwischen 1-10% vergeben, ist dies eine korrekte Antwort. Prozentwerte unter 1% beziehungsweise über 10% sind somit inkorrekte Antworten (siehe Tabelle 41).

Tabelle 41 Auszug aus der Richtlinie für die „Häufigkeitsangaben für Nebenwirkungen“ des BfArM

SPC-Richtlinie	deutsch	%-Angaben
common	häufig	$\geq 1\%$ -- $< 10\%$
uncommon	gelegentlich	$\geq 0,1\%$ -- $< 1\%$
rare	selten	$\geq 0,01\%$ -- $< 0,1\%$

Schaut man sich nun die Ergebnisse an (siehe Tabelle 42), so liegen in allen Berufsgruppen nur wenig korrekte Interpretationen vor. Lediglich der Begriff „häufig“ wurde korrekter verwendet. So haben beispielsweise 6 Apotheker (5,8%) und 4 Ärzte (2,2%) den Begriff „häufig“ im Kontext richtig numerisch interpretiert. „Selten“ wurde

sogar nur zwei Mal korrekt verwendet, während 177 inkorrekte Interpretationen bei den Apothekern vorlagen.

Tabelle 42: Anzahl der korrekten und inkorrekten Antworten der Studienteilnehmer unterteilt nach Berufsgruppen.

	häufig		gelegentlich		selten	
	ohne Kontext	Kontext	ohne Kontext	Kontext	ohne Kontext	Kontext
Ärzte						
korrekt	2 (1,1)	4 (2,2)	0 (0,0)	1 (0,6)	2 (1,1)	1 (0,6)
inkorrekt	176 (98,9)	176 (97,8)	178 (100)	180 (99,4)	177 (98,9)	178 (99,4)
Apotheker						
korrekt	1 (1,0)	6 (5,8)	1 (1,0)	2 (1,9)	0 (0,0)	2 (1,9)
inkorrekt	104 (99,0)	98 (94,2)	104 (99,0)	102 (98,1)	105 (100)	102 (98,1)
Juristen						
korrekt	0 (0,0)	1 (0,7)	1 (0,7)	0 (0)	1 (0,7)	1 (0,7)
inkorrekt	140 (100)	139 (99,3)	139 (99,3)	139 (100)	139 (99,3)	138 (99,3)

„Ohne Kontext“ entspricht der Interpretation aus Abschnitt A; „Kontext“ entspricht der Interpretation aus Abschnitt B. In Klammern stehen die absoluten Prozente.

3.6.2 Welche Begriffe wurden den Verhältniszahlen zugeordnet?

Es werden nun die Häufigkeitsverteilungen der Wahrscheinlichkeiten „1:1000“, „1:100“, „1:10“, „1:5“ und „1:2“ des Abschnitt C dargestellt.

3.6.2.1 „1:1000“

146 Probanden haben der Verhältniszahl „1:1000“ den Begriff „selten“ zugeordnet. Weitere 121 Probanden haben „sehr selten“ gewählt um „1:1000“ zu beschreiben. Insgesamt haben somit immerhin 66% diese beiden Begriffe ausgewählt.

Tabelle 43: Anzahl der Nennungen der Begriffe bei „1:1000“. Es wurden alle Begriffe gelistet, die mindestens fünfmal genannt wurden.

1:1000		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
selten	146	36,1 %
sehr selten	121	29,9 %
äußerst selten	21	5,2 %
extrem selten	14	3,5 %
unwahrscheinlich	12	2,97 %
fast nie	11	2,7 %
gelegentlich	10	2,5 %
so gut wie nie	10	2,5 %
ganz selten	6	1,5 %
Andere Begriffe	53	13,1 %
Gesamt	404	100 %

Insgesamt haben 404 Probanden „1:1000“ interpretiert, 26 haben keine Aussage gemacht.

Weitere Nennungen für „1:1000“: Viermal wurde der Begriff „einer von 1000“ verwendet. Auf die Begriffe „vereinzelt“, „denkbar“, „1 Promille“, „nahezu ausgeschlossen“, „nie“ entfielen jeweils 3 Nennungen. Zweimal wurden die Begriffe „sehr, sehr selten“, „kaum“, „ausnahmsweise“, „seltener“, und „gering“ genannt. Einmal wurde die Begriffe „oft“, „möglicherweise“, „manchmal“, „wenig wahrscheinlich“, „nicht selten“, „sehr gering“, „nicht ausgeschlossen“, „fraglich“, „nicht so häufig“, „regelmäßig“, „praktisch nie“, „Promill-Bereich“, „eher selten“, „relativ selten“, „unklar“, „höchst selten“, „äußerst unwahrscheinlich“, „in Ausnahmefällen“, „extrem gering“, „fast niemals“, „mit an Sicherheit grenzender Wahrscheinlichkeit auszuschließen“, „recht selten“, „so gut wie ausgeschlossen“ und „nicht sehr wahrscheinlich“ genannt.

3.6.2.2 „1:100“

Bei „1:100“ wurde der Begriff „selten“ am häufigsten gewählt. Insgesamt 175-mal von 402 Probanden wählten diesen aus. Der Begriff „gelegentlich“ wurde 102-mal verwendet.

Tabelle 44: Anzahl der Nennungen der Begriffe bei „1:100“. Es wurden alle Begriffe gelistet, die mindestens fünfmal genannt wurden.

1:100		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
selten	175	40,7 %
gelegentlich	102	23,7 %
sehr selten	30	7,0 %
manchmal	13	3,0 %
häufig	10	2,3 %
denkbar	10	2,3 %
möglicherweise	6	1,4 %
vereinzelt	5	1,2 %
einer von 100	5	1,2 %

Insgesamt haben 402 Probanden geantwortet, 28 haben keine Angaben gemacht.

Weitere Nennungen für „1:100“: Viermal wurde der Begriff „ab und zu“ genannt, gefolgt von drei Nennungen von „unwahrscheinlich“ und „möglich“. Zwei Nennungen entfielen auf „eher selten“, „1%“, „kommt vor“, „fraglich“, „häufiger“, „wenig wahrscheinlich“ und „öfter“. Eine Nennung gab es jeweils zu den Wörtern „gering“, „eventuell“, „eher unwahrscheinlich“, „kann sein, praktisch ausgeschlossen“, „es kann vorkommen“, „unklar“, „Prozentbereich“, „kaum“, „mäßig“, „relevant“, „nicht auszuschließen“, „überwiegend“, „typischerweise“, „gehäuft“, „fast nie“, „meist“, „wahrscheinlich“, „sehr unwahrscheinlich“, „sehr sehr selten“, „ganz selten“, „sehr oft“ und „oft“.

3.6.2.3 „1:10“

Der Begriff „gelegentlich“ wurde am häufigsten für „1:10“ gewählt (siehe Tabelle 45).

Tabelle 45 Anzahl der Nennungen der Begriffe bei „1:10“. Es wurden alle Begriffe gelistet, die mindestens fünfmal genannt wurden.

1:10		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
gelegentlich	117	29,0 %
häufig	89	22,1 %
oft	29	7,2 %
selten	28	6,9 %
manchmal	26	6,5 %
möglicherweise	16	4,0 %
möglich	9	2,2 %
öfters	7	1,7 %
denkbar	7	1,7 %
sehr häufig	6	1,5 %
wahrscheinlich	6	1,5 %
jeder 10.	6	1,5 %
häufiger	5	1,2 %

Insgesamt haben 403 Personen geantwortet, 27 haben keine Angaben gemacht.

Weitere Nennungen für „1:10“: Die Begriffe „überwiegend“, „nicht selten“, „fraglich“ wurden viermal genannt. Zwei Nennungen gab es jeweils zu den Begriffen „öfter“, „meist“, „ziemlich häufig“, „relativ häufig“, „auszuschließen“, „ab und zu“, „10%“, „teilweise“ und „bisweilen“. Einmal wurden die Begriffe „wenig wahrscheinlich“, „einer von 10“, „gehäuft“, „unwahrscheinlich“, „zum Teil“, „mit höherer Wahrscheinlichkeit“, „sehr bedeutsam“, „kann auftreten“, „kaum“, „hin und wieder“, „unter Umständen“, „nicht auszuschließen“, „relativ oft“, „unsicher“, „damit rechnen – ich denke nicht“, „kann schon mal vorkommen“, „eventuell“, „mittleres Risiko“, „ziemlich hoch“, „des Öfteren“, „kann schon sein, aber selten“, „in einigen Fällen“.

3.6.2.4 „1:5“

Insgesamt haben 100 Probanden den Begriff „häufig“ gewählt, um die Verhältniszahl „1:5“ zu beschreiben. Weitere 76 Probanden wählten „oft“ zur Interpretation aus. Wiederum werden die fünf häufigsten Begriffe tabellarisiert dargestellt.

Tabelle 46 Anzahl der Nennungen der Begriffe bei „1:5“. Es wurden alle Begriffe gelistet, die mindestens fünfmal genannt wurden.

1:5		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
häufig	100	25,0 %
oft	76	19,0 %
gelegentlich	50	12,5 %
sehr häufig	28	7,0 %
möglicherweise	15	3,8 %
manchmal	13	3,3 %
sehr oft	12	3,0 %
wahrscheinlich	9	2,3 %
öfter	8	2,0 %
jeder 5.	6	1,5 %
relativ häufig	6	1,5 %
öfters	5	1,3 %
überwiegend	5	1,3 %
regelmäßig	5	1,3 %

Insgesamt haben 400 Probanden geantwortet, 30 haben keine Angaben gemacht.

Weitere Nennungen für „1:5“: Jeweils vier Nennungen fielen auf die Begriffe „meist“ und „denkbar“. Dreimal wurden die Wahrscheinlichkeitsbegriffe „selten“, „möglich“, „sehr wahrscheinlich“, „häufiger“ und „ab und zu“ genannt. Zwei Nennungen entfielen auf „recht häufig“, „typischerweise“, „fast immer“, „20%“, und „hin und wieder“. Einmal wurden die verbalen Wahrscheinlichkeiten „sehr selten“, „gehäuft“, „ziemlich oft“, „nicht auszuschließen“, „oft möglich“, „unvertretbar hoch“, „großes Risiko“, „extrem häufig“, „oftmals“, „kommt vor“, „große Wahrscheinlichkeit“, „immer wieder (mal)“, „viele“, „teilweise“, „wiederholt“, „zu oft“, „kann sein“, „eventuell“, „hohes Risiko“, „hoch“, „vermehrt“, „sehr möglich“, „mit gewisser Wahrscheinlichkeit“, „relativ wahrscheinlich“, „recht häufig“, „mehrfach“, „schon möglich“, „gut möglich“ und „meist nicht“ genannt.

3.6.2.5 „1:2“

Es haben 25% der Befragten den Begriff „häufig“ zur Beschreibung von „1:2“ gewählt. Alle Begriffe, die weniger als fünfmal genannt wurden, repräsentieren in ihrer Gesamtheit 15,3% der Werte. Im Anschluss an die Tabelle erfolgt die Nennung dieser Begriffe.

Tabelle 47 Anzahl der Nennungen der Begriffe bei „1:2“. Es wurden alle Begriffe gelistet, die mindestens fünfmal genannt wurden.

1:2		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
häufig	100	25,0 %
oft	52	13,0 %
sehr häufig	40	10,0 %
wahrscheinlich	29	7,3 %
sehr oft	27	6,8 %
jeder 2./einer von 2	18	4,5 %
fast immer	14	3,5 %
meist	10	2,5 %
typischerweise	9	2,3 %
regelmäßig	8	2,0 %
möglicherweise	8	2,0 %
meistens	6	1,5 %
überwiegend	6	1,5 %
50:50 Chance	6	1,5 %
möglich	6	1,5 %

Insgesamt haben 400 Personen dem „1:2“ einen Prozentwert gegeben, während 30 Werte fehlen.

Weitere Nennungen für „1:2“: Es wurden jeweils viermal die Begriffe „höchst wahrscheinlich“, „denkbar“, „sehr wahrscheinlich“ vergeben. Drei Nennungen entfielen auf „stets“ und „in der Hälfte der Fälle“. Zweimal wurden die Ausdrücke „in der Regel“, „mit hoher Wahrscheinlichkeit“, „es ist damit zu rechnen“, „unsicher“, „immer wieder (mal)“, „50%“ und „immer“ genannt. Einmal wurden die Begriffe „so gut wie immer“, „fast regelmäßig“, „sicher“, „bedrohlich oft“, „jeder dritte Fall“, „vermehrt“, „sehr hoch“, „hoch“, „alle 2“, „sehr hohes Risiko“, „fast sicher“, „mit Wahrscheinlichkeit“, „wird sein“, „unklar“, „risikoreich“, „unvertretbar oft“, „wiederholt“, „unter Umständen“, „mit großer Wahrscheinlichkeit“, „oftmals“, „extrem häufig“, „fast regelhaft“, „lassen sie es sein“, „sehr großes Risiko“, „unvertretbar hoch“, „gehäuft“, „fast nie“, „recht häufig“ und „sehr selten“ genannt.

3.6.3 Welche Begriffe wurden am häufigsten verwendet?

Die Verhältniszahl „1:1000“ haben die Ärzte jeweils 63-mal mit „sehr selten“ oder „selten“ beschrieben. Die Apotheker wählten „sehr selten“ und die Juristen „selten“ am häufigsten. Alle Berufsgruppen haben bei „1:2“ „häufig“ am meisten gewählt (siehe Tabelle 48). Schaut man sich die Ergebnisse an, so fällt auf, dass die am häufigsten gewählten Begriffe Teil der offiziellen Klassifikation sind.

Tabelle 48 Vergleich der meist gewählten Begriffe mit Angabe der jeweiligen Anzahl (d)

	Ärzte	d	Apotheker	d	Juristen	d
1:1000	sehr selten/selten	je 63	sehr selten	34	selten	50
1:100	selten	78	selten	48	selten	49
1:10	gelegentlich	55	gelegentlich	29	gelegentlich	33
1:5	häufig	52	häufig	26	oft	29
1:2	häufig	48	häufig	19	häufig	33

3.6.4 Vergleich der gewählten Begriffe von „1:1000“, „1:100“ und „1:10“ mit der offiziellen Nebenwirkungstabelle

Die Wahrscheinlichkeiten „1:1000“, „1:100“, „1:10“, „1:5“ und „1:2“ aus Abschnitt C lassen sich wie folgt der offiziellen Nebenwirkungstabelle zuordnen:

Tabelle 49: Einordnung der numerischen Wahrscheinlichkeiten in die offizielle Nebenwirkungstabelle

Wahrscheinlichkeit	offizielle Definition	verbale Wahrscheinlichkeit
1:1000	entspricht 0,1%	gelegentlich
1:100	entspricht 1%	häufig
1:10	entspricht 10%	sehr häufig
1:5	entspricht 20%	sehr häufig
1:2	entspricht 50%	sehr häufig

Die Wahrscheinlichkeiten „1:10“, „1:5“ und „1:2“ entsprechen alle dem Begriff „sehr häufig“. „1:100“ wurde offiziell mit „häufig“ und „1:1000“ mit „gelegentlich“ klassifiziert.

Die Anzahl der korrekten und inkorrekten Antworten wurden in Tabelle 50 veranschaulicht. Die Wahrscheinlichkeit „1:1000“ wurde von 2,3% der Ärzte, 2,0% der Apotheker und von 3,0% der Juristen mit dem richtigen Begriff „gelegentlich“ belegt. „Häufig“ wurde für die Verhältniszahl „1:100“ von 4,1% der Ärzte, 1,0% der Apotheker und 1,5% der Juristen richtig gewählt. „Sehr häufig“ entspricht der Wahrscheinlichkeit „1:10“. Dies wurde immerhin etwas häufiger richtig interpretiert (siehe Tabelle 50).

Tabelle 50: Korrekte und inkorrekte Antworten der Studienteilnehmer für jede Berufsgruppe einzeln und als Gesamtergebnis dargestellt

		1:1000	1:100	1:10	1:5	1:2
Ärzte	korrekt	4 (2.3)	7 (4.1)	4 (2.3)	7 (4.1)	27 (15.9)
	inkorrekt	168 (97.7)	165 (95.9)	168 (97.7)	164 (95.9)	143 (84.1)
Apothek	korrekt	2 (2.0)	1 (1.0)	1 (1.0)	9 (9.4)	10 (10.3)
	inkorrekt	97 (98.0)	96 (99.0)	96 (99.0)	87 (90.6)	87 (87.7)
Juristen	korrekt	4 (3.0)	2 (1.5)	1 (0.8)	12 (9.1)	2 (1.5)
	inkorrekt	128 (97.0)	130 (98,5)	132 (99.2)	120 (90.9)	130 (98.5)
Gesamt	korrekt	10 (2.5)	10 (2.5)	6 (1.5)	28 (7.0)	40 (10.0)
	inkorrekt	394 (97,5)	392 (97.5)	397 (85.5)	372 (93.0)	360 (90.0)

In den Klammern sind die Prozentwerte angegeben.

Zusammenfassend lässt sich feststellen, dass die Wahrscheinlichkeit „1:2“ mit 10.0%, bezogen auf alle Berufsgruppen, am häufigsten mit dem korrekten verbalen Wahrscheinlichkeitsbegriff nach der SPC-Richtlinie belegt wurde.

Die Zuordnung von frei wählbaren Wahrscheinlichkeitsbegriffen zu den 5 Verhältniszahlen zeigt den Reichtum der deutschen Sprache an diesen Begriffen. Insgesamt wurden, von allen Teilnehmern der Studie zusammengerechnet, 172 verschiedene Begriffe ausgewählt.

4 Diskussion

Als Untersucher und Autor dieser Arbeit stellt man sich nun die Frage, wie die Qualität dieser Studie zu bewerten ist. Was kann man für Schlussfolgerungen aus diesen Ergebnissen ziehen und welche Konsequenzen ergeben sich daraus für den Arzt, Apotheker oder Juristen? Konnten die initialen Forschungsfragen mit den errechneten Daten beantwortet werden?

4.1 Wie gut und relevant ist diese Studie?

Wie wurde nun diese Befragung von den Teilnehmern angenommen? Konnte das Interesse der Probanden geweckt werden?

Eine Ausschöpfungsquote von insgesamt 62% lässt auf ein großes Interesse der Teilnehmer an dieser Studie schließen. In der Literatur werden Quoten von mehr als 40% als hervorragend angesehen [37]. Werte über 80% gelten als unrealistisch [40]. Der gute Rücklauf dieser Studie wurde durch das ansprechende Anschreiben und den Fragebogen gefördert. Wichtig in dieser Hinsicht waren auch die eigenhändige Unterschrift und die persönliche Bitte zur Unterstützung einer Dissertation. Mit dem Erinnerungsschreiben konnte der Rücklauf noch erhöht werden. Im zweiten Durchlauf fiel der geplante Zeitpunkt des Erinnerungsschreibens in die Sommerferien. Hier wurde die Aussendung unter der Vorstellung, man möge in seinem Urlaub nicht mit Fragebögen konfrontiert werden, verschoben.

Durch den bereits adressierten und frankierten Rückumschlag wurde die Rücksendung deutlich erleichtert und die Teilnahme blieb darunter für die Probanden kostenlos. Es ist anzunehmen, dass auch diese Maßnahmen die Antwortbereitschaft erhöhten.

Bezüglich der Rücksendedauer fiel ein rasches Antworten auf. So waren in beiden Durchgängen bereits nach sieben bzw. acht Werktagen mehr als 50 % der Antworten im Institut für Medizinische Biomathematik eingetroffen. Der zweite Durchlauf dauerte mit 110 Werktagen bis zum Eintreffen des letzten Briefs aber deutlich länger. Dies könnte einerseits an den Sommerferien liegen, andererseits könnte der später verschickte Erinnerungsbogen auch dazu beigetragen haben.

Schaut man sich nun das Antwortverhalten der einzelnen Berufsgruppen genauer an, so erkennt man einen deutlichen Unterschied zwischen den einzelnen Berufsgruppen. So wurden Ausschöpfungsquoten von 53,5% bei Apothekern, 61,1% bei Ärzten und 71,9% bei Juristen gesehen. Die Juristen waren engagierter im Beantworten dieses Fragebogens als Ärzte und Apotheker. Dies kann zum einen am Fragebogen liegen, wenn sich z.B. die Apotheker weniger angesprochen fühlen. Zum anderen werden besonders Berufsgruppen wie Apotheker und Ärzte oft mit Studien und Befragungen überhäuft, so dass durchaus ein Überdruß bestehen könnte. Dieser Fragebogen ist in seiner Fragestellung für Apotheker und Ärzte relativ fachspezifisch, während Juristen

seltener an medizinisch orientierten Studien teilnehmen. Vielleicht war auch der Unterschied dieser Fragen zum juristischen Alltag interessant. Wer lässt sich nicht mal gerne vom Alltäglichen ablenken?

Es fiel auch auf, dass insbesondere die Juristen zusätzliche schriftliche Anmerkungen machten. Teilweise wurden persönliche Wünsche für ein gutes Gelingen mitgeteilt. Die Ärzte und Apotheker waren da eher zurückhaltend. Einige Probanden baten sogar um ein Zusenden der Auswertung, da Interesse an den Ergebnissen bestehe.

Insgesamt sind die Ausschöpfungsquoten sehr gut, was für die Qualität dieser Umfrage spricht. Die Rücklaufquote übertraf bei weitem der, vor Beginn der Studie, angenommenen Quote. Es konnte das Interesse für die Darstellung von Wahrscheinlichkeiten in allen Berufsgruppen geweckt werden.

4.2 Gibt es berufsgruppenspezifische Rangreihen der Wahrscheinlichkeitsbegriffe und unterscheiden sich diese voneinander?

Schaut man sich zunächst die freie numerische Interpretation der 20 Wahrscheinlichkeitsbegriffe an, so fällt auch hier eine große Spannweite in der Interpretation der einzelnen Begriffe auf. Wenn „häufig“ beispielsweise von einer Person mit 0,5 % und von einer anderen Person mit 100% belegt wird, so sind Missverständnisse im Gespräch vorprogrammiert. Ein adäquates Gespräch mit fundierten Entscheidungen erscheint unter diesen Umständen kaum möglich zu sein.

Begriffe wie „unklar“ oder „fraglich“ wurden sogar von einigen Probanden gar nicht erst numerisch interpretiert. Diese Begriffe weisen außerdem in ihren Streuungsmaßen Auffälligkeiten auf. So liegen die Quantile sehr weit auseinander. Immerhin 25% der Apotheker sagten „unklar“ sei weniger als 10%, während 25% der Meinung waren, es müsse mehr als 50% bedeuten (siehe Tabelle 11). Die restlichen Apotheker (also 50%) meinen der Wert von „unklar“ liege zwischen 10 und 50%. Die Verweigerung diese Begriffe zu interpretieren, wie auch die große Spannweite, lassen auf eine Unsicherheit in der Interpretation schließen.

Insgesamt ergeben sich auch in dieser Studie für alle 20 Begriffe sehr große Spannweiten ähnlich zu vorherigen Untersuchungen [1, 2, 4, 12, 13, 49]. So liegen die minimalen und maximalen Prozentwerte oft weit auseinander (siehe oben). Allerdings scheinen einige Begriffe auch klarer zu sein. 75% aller Probanden haben zum Beispiel „selten“ mit einem Wert kleiner als 10 % interpretiert.

Diese Ergebnisse untermauern noch einmal Studien der letzten Jahre, die klar eine alleinige Darstellung von Wahrscheinlichkeiten mittels verbalen Wahrscheinlichkeitsbegriffen ablehnen [7, 34].

Da es für den Menschen aber natürlicher ist, Worte zu gebrauchen und Zahlen häufig den Redefluss und damit auch das Verständnis unterbrechen [9], wird der Gebrauch von verbalen Wahrscheinlichkeitsbegriffen nicht zu vermeiden sein. Häufig liegen auch keine genaueren Zahlen von Wahrscheinlichkeiten vor oder differieren von Buch zu Buch, von Professor zu Professor, von Sachkundigem zu Sachkundigem. Um keine Genauigkeiten vorzutauschen, die es eventuell gar nicht gibt, wird die Verwendung von verbalen Wahrscheinlichkeitsbegriffen nicht zu unterbinden sein.

Ein Ansatz ist der Versuch Rangreihen zu bilden, um wenigstens eine Tendenz zu ermitteln [9, 16, 23]. Wie in Material und Methodik beschrieben, konnte in einer früheren Dissertation eine „ideale Rangreihe von Wahrscheinlichkeitsbegriffen“ für die Berufsgruppe der Ärzte (siehe Tabelle 1) nachgewiesen werden. Diese „ideale Rangreihe“ zeigte ein enges Maß an Zusammenhang zur englischsprachigen Literatur.

In dieser Studie wurde nun für jede Berufsgruppe eine Rangreihe der Wahrscheinlichkeitsbegriffe erstellt. Der Konkordanzkoeffizient Kendall W (siehe 3.3.2 Rangreihenfolgen) wurde zum Vergleich der Rangreihen innerhalb einer Berufsgruppe verwendet. So wurden zum Beispiel alle Rangreihen der Ärzte miteinander verglichen, um den Grad der Übereinstimmung zu ermitteln. Mit Werten von 0,684 für Ärzte, 0,696 für Apotheker und 0,709 für Juristen kann für jede Berufsgruppe die Nullhypothese bei einem Signifikanzniveau von $p < 0,0001$ zurückgewiesen werden.

Da kein Verfahren zur direkten Berechnung des Konfidenzintervalls für den Kendall Konkordanzkoeffizienten W vorliegt, wurde hier das Bootstrap-Verfahren verwandt. Die hiermit berechneten Konkordanzkoeffizienten der jeweiligen Berufsgruppen zeigten eine hohe Übereinstimmung mit den Konkordanzkoeffizienten aus den Rohdaten. Das bedeutet, dass bei 1000 neuen zufälligen Ziehungen annähernd dieselben Werte für Kendall W errechnet wurden und damit eine ähnlich hohe Übereinstimmung in den neu errechneten Rangreihenfolgen, wie die der ursprünglichen Stichprobe, vorliegt. Mit dem Bootstrap-Verfahren konnten die initialen Hinweise für das Vorliegen von berufsgruppenspezifischen Rangreihen noch untermauert werden. Für jede Berufsgruppe ist somit eine „echte berufsgruppenspezifische“ Rangreihe anzunehmen.

Die in dieser Arbeit neu erstellten Rangreihen der einzelnen Berufsgruppen wurden anschließend mit der „idealen Rangreihe“ aus der deutschsprachigen Literatur verglichen. Dies geschah unter Zuhilfenahme der Rangkorrelationskoeffizienten Spearman's ρ und Kendall τ . Die Ergebnisse lassen für alle Berufsgruppen auf ein hohes Maß an Zusammenhang schließen. So dass jeweils die Nullhypothesen, es gäbe keinen Zusammenhang, verworfen wurden. Die drei neuen echten Rangreihen unterscheiden sich somit nur wenig von der idealen Rangreihe der Literatur.

Mit dem Permutationstest konnte gezeigt werden, dass die Rangreihen nicht zufällig angeordnet sind. Somit ist von einer Konsistenz in der Vergabe der Rangplätze auszugehen. Wenn schon keine einheitliche Interpretation vorliegt, so muss man aber eine einheitliche Rangreihenfolge der Wahrscheinlichkeitsbegriffe unterstellen. So

scheint es klar zu sein, dass „typischerweise“ häufiger eintritt als „möglicherweise“ oder „gelegentlich“.

4.3 Besteht eine Kontextabhängigkeit der Wahrscheinlichkeitsbegriffe und gibt es einen Unterschied zwischen den Berufsgruppen?

Die Probanden dieser Studie waren aufgefordert die Begriffe „häufig“, „gelegentlich“ und „selten“ im Kontext mit Nebenwirkungen zu interpretieren. Es bot sich an diese drei Begriffe auszuwählen, da die Ergebnisse vergleichbar mit der offiziellen Klassifikation von Nebenwirkungen sind.

Schaut man sich zunächst nur die Minimal- und Maximalwerte der Begriffe an, so fällt die große Spannbreite auf. Dabei gibt es keinen Unterschied zwischen der freien Interpretation der Begriffe (siehe Tabelle 13) und der Interpretation im Kontext (siehe Tabelle 20). Der Begriff „häufig“ zum Beispiel erhielt im Kontext Werte von 1% bis 99%. Dies entspricht ungefähr dem Ergebnis der freien Interpretation aus Abschnitt A. Der Kontext hat somit keinen Einfluss auf die Minimal- oder Maximalwerte.

Wie genau verteilen sich nun aber die Prozentwerte? Dies ist sehr anschaulich den graphischen Gegenüberstellungen der Begriffe zu entnehmen (siehe Abbildung 1- Abbildung 12). So wirkt die Abbildung von „häufig“ ohne Kontext (Abbildung 1) optisch kompakter und stellt sich eher im höheren Prozentbereich dar. Im Kontext verlagert sich „häufig“ aber eher in den niedrigeren Prozentbereich. Dies spiegelt sich auch in den Ergebnissen der Mediane und Quantile wider. So beträgt das 25%- Quantil kontextfrei 60% und im Kontext 40% und das 75%- Quantil bei 80% bzw. 75%.

Mithilfe des Wilcoxon-Tests konnte dann auch rechnerisch eine Abhängigkeit vom Kontext bestätigt werden. Es wurden in allen Berufsgruppen niedrigere Prozentwerte als in der freien Interpretation vergeben. Mit dieser Studie kann somit nachgewiesen werden, dass auch bei den Juristen und Apothekern eine Kontextabhängigkeit in der Interpretation von Wahrscheinlichkeitsbegriffen vorliegt.

Dieses Ergebnis unterstützt frühere Studien, die bereits eine Kontextabhängigkeit bewiesen [9, 13, 15, 18, 20, 50]. Es konnte gezeigt werden, dass negative Endpunkte wie Pneumonien oder Tod zu einer Kontextabhängigkeit führten. Die verbalen Wahrscheinlichkeiten wurden in diesem Zusammenhang mit niedrigeren Prozentwerten belegt [18, 51].

Auch der Begriff Nebenwirkungen an sich impliziert ein negatives Gefühl: Etwas, was man als Patient nicht haben will, was sich kein Arzt für seinen Patienten wünscht, was möglichst selten auftreten soll. Somit lässt sich vermutlich auch erklären, warum die numerischen Interpretationen in Zusammenhang mit Nebenwirkungen niedriger

ausfallen als in der freien Interpretation. Es ist leicht vorstellbar, dass Begriffe wie Küssen, Lachen, Flirten, Erfolg und viele weitere positive Begriffe mit deutlich höheren Prozentwerten belegt werden würden, wenn man sie in diesem Test verwendet hätte.

Dieses Wissen um die Kontextabhängigkeit erleichtert die Kommunikation leider nicht. Parameter wie das Alter und die Erfahrung der Gesprächspartners sind ebenfalls mit ausschlaggebend [18, 51, 52]. Gerade bei einem heiklen Thema wie Nebenwirkungen möchten viele Patienten hinreichend aufgeklärt sein. Wenn also die umfassende Aufklärung im Vordergrund steht, reichen allein verbale Wahrscheinlichkeitsbegriffe, wie dies *Paling* [34] beschreibt, nicht aus. Auch in dieser Studie sind die Spannweiten der Begriffe viel zu groß.

Doch nicht jeder Patient ist gleich in seinem Wunsch nach vollständiger Aufklärung. So gibt es Patienten, die alles bis ins Kleinste genau wissen wollen, die sich umfangreich mit Therapien, Nebenwirkungen und Alternativen beschäftigen. Wenn nicht schon mit dem I-Phone in der Praxis, so werden doch spätestens zu Hause, am PC, alle nötigen Informationen im Internet nachgelesen. Hier sollten definitiv numerische Wahrscheinlichkeiten, soweit sie bekannt sind, gebraucht werden. Diese Patienten würden schwammige verbale Wahrscheinlichkeiten als Versuch empfinden, sie nicht ausreichend aufklären zu wollen. Das Arzt-Patienten-Verhältnis wäre schon vor Therapiebeginn gestört. Diese Patienten verlangen geradezu nach einer numerischen Darstellung der Nebenwirkungen. Diese werden sich am Ende eines Gesprächs auch mit den Worten verabschieden: „Danke, dass Sie so offen waren.“

Viele ängstlichere Patienten dagegen wünschen gar keine umfassende Aufklärung. Häufig wird schon am Anfang eines Gesprächs geäußert, man wolle gar nicht alles so genau wissen. Für diese Patienten ist es lediglich wichtig, dass es etwas gibt, was ihnen helfen kann. Diese eher emotional denkenden und handelnden Personen, welche nach Erfahrung der Doktorandin in der Überzahl sind, sind froh und zufrieden mit einem Begriff wie „selten“. Es lässt sie in der Hoffnung leben, dass sie es schon nicht sein werden, die diese schlimme Nebenwirkung bekommen. Zu hören, man könne einer von 3000 sein („selten“ entspricht $\geq 0,01\%$ bis $< 0,1\%$), der an der Therapie stirbt, verursacht bei diesen emotionalen Menschen eine Abwehrhaltung. Ja, sie führt vielleicht sogar zu einem Verzicht auf eine dringend nötige Therapie.

Es bleibt also die Aufgabe des Arztes herauszufinden, welcher Typ Mensch vor ihm sitzt und welche Art der Kommunikation er wählen muss.

Die zweite Frage war, ob sich die Berufsgruppen bezüglich der Kontextabhängigkeit unterscheiden. Zunächst konnte für alle Berufsgruppen eine Abhängigkeit vom Kontext Nebenwirkungen festgestellt werden. Doch es existieren auch Unterschiede. So haben die Juristen die Begriffe im Kontext zwar niedriger als in der freien Interpretation gewählt. Insgesamt liegen die Mediane und Quantile im Kontext aber etwas über denen der Ärzte und Apotheker. Unterscheiden sich die Mediziner und Apotheker im

paarweisen Vergleich (siehe 3.4.6) nur wenig voneinander, so fällt doch ein signifikanter Unterschied zur Gruppe der Juristen auf.

Über die Gründe kann nur spekuliert werden. Zunächst sind die ähnlichen Ergebnisse zwischen den Ärzten und Apothekern sicherlich in der Ähnlichkeit der Berufe zu finden. Mediziner müssen sich zwangsläufig mit der Pharmazie beschäftigen und die Apotheker kommen auch nicht ohne medizinisches Wissen aus. Diese Berufe sind also miteinander verbunden. Das Thema Nebenwirkungen beschäftigt diese beiden Berufsgruppen täglich und so liegen sicherlich auch Erfahrungswerte vor, die in die Interpretation mit einfließen. Wenn man z.B. allergische Reaktionen auf ein Penicillin schon mehrfach in seinem Leben gesehen und behandelt hat, so kann man die Häufigkeit anders einschätzen als vielleicht ein Jurist. Dieser besitzt sicherlich das Allgemeinwissen, dass Penicilline allergische Reaktionen verursachen können. Da aber die Erfahrungswerte fehlen, ist eine Häufigkeitseinschätzung schwieriger. Einem Mediziner würde es ähnlich gehen, wenn er über die Häufigkeiten von Vaterschaftsklagen mutmaßen sollte.

Vielleicht sind Juristen aber auch skeptischer; skeptischer gegenüber Arzneimitteln, Behandlungen und damit auch Nebenwirkungen. In den Medien wurde in letzter Zeit über eine Zunahme der Arzthaftungsprozesse berichtet. Im deutschen Ärzteblatt erscheinen regelmäßig Artikel zu Arzthaftungsfragen. Auch Juristen müssen sich immer mehr mit diesem Thema beschäftigen. Vielleicht trägt auch eine unterschiedliche Wahrnehmung zu den Unterschieden zwischen den Berufsgruppen bei. Ein Beispiel: Ein Arzt verschreibt Gentamicin leitliniengerecht bei einer Endokarditis. Er sieht vielleicht 1000 Patienten ohne ototoxische Nebenwirkungen und einen Patienten mit einem Innenohrschaden. Ein Rechtsanwalt, der nun aber schon zehn Patienten vertreten hat, die über ototoxische Nebenwirkungen klagten, wird auch eine andere Wahrnehmung der Häufigkeiten haben. Denn dieser hat nicht die 1000 gesunden Patienten, denen das Medikament auch verabreicht wurde, gesehen.

Diese Überlegungen sind allerdings spekulativ und beruhen auf Erfahrungen der Autorin und auf Gespräche mit angehenden Juristen. Die genaueren Hintergründe müssten in weiteren Studien erfragt werden. In dieser Studie konnte zwar ein Unterschied nachgewiesen werden, die Gründe konnten aber nicht erfragt werden. Dies ist der Nachteil einer schriftlichen Befragung. Außerdem fällt auf, dass einige Teilnehmer exakt die gleichen Zahlen im Abschnitt B wie in A verwandt haben. Einerseits wäre es denkbar, dass die Begriffe prozentual so verinnerlicht sind, dass für beide Interpretationen der gleiche Wert verwendet wird. Andererseits ist aber auch ein Zurückblättern und Ablesen der Zahlen nicht auszuschließen. Um diesen Fehler zu vermeiden, müsste ein Verfahren angewandt werden, welches ein Zurückblättern verhindert. Dies könnte z.B. mittels einer computer-gesteuerten Befragung erfolgen. Allerdings würde dann die Teilnehmerquote vermutlich sinken.

Aus den Ergebnissen dieser Arbeit lassen sich bezüglich der Kontextabhängigkeit folgende Schlussfolgerungen ziehen: Die Ergebnisse der Mediziner unterstützen frühere Studien, die bereits eine Kontextabhängigkeit postulierten [9, 18, 20]. Neu ist, dass auch Apotheker und Juristen Wahrscheinlichkeiten im Kontext anders interpretieren.

4.4 Kennen Ärzte, Apotheker und Juristen die Richtlinie für „Häufigkeitsangaben von Nebenwirkungen“?

Vielen Autoren haben bereits eine Kodierung von Wahrscheinlichkeitsbegriffen empfohlen und so wurden einige Vorschläge erarbeitet [4, 15, 17, 21, 53]. Bezüglich Nebenwirkungen wurde vom Bundesinstitut für Arzneimittel und Medizinprodukte eine offizielle Richtlinie veröffentlicht. Die fünf Begriffe „sehr häufig“, „häufig“, „gelegentlich“, „selten“ und „sehr selten“ sind somit klar definiert. Diese Richtlinie ist mittlerweile Bestandteil jedes Medikamentenbeipackzettels.

Umso erstaunlicher sind die Ergebnisse dieser Studie. Nur wenige der Probanden haben den Begriffen „häufig“, „gelegentlich“ und „selten“ im Kontext Nebenwirkungen den richtigen Prozentwert zugeordnet (siehe Tabelle 42). Von 413 Teilnehmern haben nur 11 Probanden „häufig“ richtig numerisch interpretiert. Bei „gelegentlich“ waren es nur drei und bei „selten“ vier Personen. Die korrekten numerischen Interpretationen sind also nicht bekannt.

Bei der Interpretation der Verhältniszahlen fiel zunächst einmal die Masse von Wahrscheinlichkeitsbegriffen auf. Insgesamt wurden 172 verschiedene Ausdrücke verwendet, um die fünf Zahlen zu beschreiben. Mit diesem Wissen braucht man sich fast gar nicht mehr die Frage nach der korrekten Interpretation stellen. Nur 10 von insgesamt 394 Probanden haben „1:1000“ den richtigen Begriff, nämlich „gelegentlich“, zugeordnet. „1:100“ wurde zehnmal richtig mit „häufig“ und „1:10“ sechsmal richtig mit „sehr häufig“ interpretiert. Dabei scheint es völlig egal zu sein, welcher Berufsgruppe man angehört. Die einzelnen Häufigkeiten unterscheiden sich nur wenig und liegen durchweg unter fünf Prozent.

Die Häufigkeiten „1:2“ und „1:5“ sind nicht explizit in der SPC- Richtlinie standardisiert. Sie fallen aber in die Kategorie >10% und entsprechen damit dem Begriff „sehr häufig“. Lediglich diese Werte wurden in sieben bzw. zehn Prozent der Fälle richtig gewählt. Diese Ergebnisse sprechen dafür, dass die genaue Klassifikation nicht geläufig ist.

Immerhin sind „häufig“, „gelegentlich“, „selten“ und „sehr selten“ die meist gewählten Begriffe (siehe Tabelle 48). Dass es eine Klassifikation der Wahrscheinlichkeiten gibt, scheint somit doch bekannt zu sein. Die Begriffe sind geläufig und werden auch angewandt. Nur leider nicht in der exakten Definition. Dies stellte schon die Arbeitsgruppe um *Knapp und Berry* [27, 29, 30] in verschiedenen Experimenten an Patienten fest. In dieser Studie konnte nun gezeigt werden, dass auch Ärzte, Apotheker

und Juristen in Deutschland die Richtlinie zu den Häufigkeitsangaben nicht adäquat verinnerlicht haben.

Nun bleibt zu klären, wo die Gründe hierfür liegen. Die Verhältniszahlen wurden eher niedriger verbal umschrieben. So wurde „1:1000“ von der Mehrzahl mit „sehr selten“ und „selten“ beschrieben (siehe Tabelle 48). Der korrekte Ausdruck wäre aber „gelegentlich“ gewesen. Für „1:100“ haben viele Probanden „selten“ gewählt, obwohl „häufig“ korrekt gewesen wäre.

Allein anhand dieser Daten würde man unterstellen, dass die Prozentwerte in der Definition zu niedrig gewählt sind. Dem Begriff „gelegentlich“ würde man eher einen höheren Wahrscheinlichkeitswert als „1:1000“ zuordnen. Eher „1:100“ oder „1:10“. Selbst unter Apothekern und Ärzten ist einer von Tausend eher ein „seltener“ oder „sehr seltener“ als ein „gelegentlicher“ Fall.

Schon die Arbeitsgruppe um *Berry* [27] bewies, dass Patienten diese fünf Begriffe der Richtlinie überinterpretierten, d.h. mit höheren Prozentwerten belegten. So zum Beispiel wurden Nebenwirkungen, die „selten“ auftraten, mit deutlich höheren Prozentwerten belegt. Die Ergebnisse dieser Studie beweisen nun, dass auch Ärzte, Apotheker und Juristen diese verbalen Begriffe überschätzen. Auch sie nehmen höhere Wahrscheinlichkeiten an, als eigentlich durch die Definition festgelegt wurde.

Warum kennen die Mediziner, Apotheker und Juristen diese Richtlinie nicht? Die Juristen haben sich natürlich im Studium nie mit dieser Richtlinie beschäftigt. Dies ist somit auch am ehesten die Ursache für die falsche Interpretation. Es kann somit unterstellt werden, dass die Ergebnisse der Juristen auch auf andere Nicht-Mediziner oder Nicht-Pharmakologen übertragbar sind. Von den Ärzten würde man erwarten diese Richtlinie zu kennen. Doch da scheitert es an mehreren Problemen. Zum einen wird diese Klassifikation im Studium nicht gelehrt. Auch *Paling* [34] fordert eine bessere Schulung der Ärzte zu diesem Thema. Zum anderen sind die Häufigkeiten von Nebenwirkungen nicht immer geläufig. Dies soll am Beispiel des Gentamicins verdeutlicht werden. Jedem Medizinstudenten werden die ototoxischen Nebenwirkungen dieses Medikaments eingetrichtert. Doch selbst in gängigen Pharmakologie-Büchern der Mediziner [54] steht nichts über die Häufigkeit des Auftretens von Innenohrschäden. Woher soll also das Wissen kommen? Andererseits kann jeder Patient unter Wikipedia oder im Beipackzettel nachlesen, dass 1-3% der Behandelten Innenohrschäden davontragen. Der Mediziner muss sich dieses Wissen also mühevoll erarbeiten; Beipackzettel wälzen; im Internet oder in Fachzeitschriften recherchieren. Bei der Vielzahl der Medikamente und der ewigen Zeitnot ein schier unlösbarer Kraftaufwand. Schon im Studium sollten die Häufigkeiten der Nebenwirkungen nicht vernachlässigt werden. Des Weiteren würde man sich eine zentrale Datenbank wünschen, die frei zugänglich wäre und laufend aktualisiert werden würde.

Was kann man nun aus diesen Ergebnissen ableiten?

Diese Studie zeigt deutlich die Probleme der Interpretation von verbalen Wahrscheinlichkeiten. Wie schon bei Ärzten und Patienten gesehen, so haben auch Apotheker und Juristen Schwierigkeiten in der numerischen Interpretation dieser Begriffe. Missverständnisse sind also nicht nur im Arzt-Patienten-Gespräch vorprogrammiert, sondern auch im Patient-Apotheker oder Jurist-Kläger-Gespräch. Ja selbst ein Apotheker kann nicht sicher sein, seinen Kollegen richtig verstanden zu haben, wenn dieser über „häufige“ Nebenwirkungen eines Medikaments spricht. Von einer alleinigen Darstellung von Nebenwirkungen mit verbalen Wahrscheinlichkeiten muss demnach abgeraten werden. Selbst dann, wenn die fünf Begriffe der offiziellen Klassifikation gebraucht werden. Die Klassifikation an sich ist sinnvoll und die Begriffe sind den Probanden geläufig. Leider kennen diese die numerische Bedeutung nicht. Einerseits liegt dies vermutlich an einer Schwäche in der Lehre. Andererseits werden die Begriffe überinterpretiert. Gegebenenfalls sollte auch über eine Änderung der Referenzwerte der Klassifikation nachgedacht werden. Wann immer Ärzte oder Apotheker Patienten über Nebenwirkungen aufklären, sollten Prozentwerte, Verhältniszahlen oder visuelle Skalen genutzt werden [34]. Wenn sich der Gebrauch von verbalen Wahrscheinlichkeitsbegriffen nicht vermeiden lässt, so sollten diese zumindest durch numerische Angaben ergänzt werden. Weitere Studien sind nötig, um die Definitionen der fünf Wahrscheinlichkeitsbegriffe anzupassen.

5 Zusammenfassung

Verstehen Ärzte, Apotheker und Juristen einander? In dieser Studie sollten verschiedene Berufsgruppen hinsichtlich des Gebrauchs von verbalen Wahrscheinlichkeitsausdrücken miteinander verglichen werden. Für Mediziner wurden bereits Rangreihen von verbalen Wahrscheinlichkeiten und eine Kontextabhängigkeit nachgewiesen. Gibt es diese auch bei Apothekern und Juristen? Des Weiteren stellte sich die Frage, ob klar definierte Wahrscheinlichkeitsbegriffe wie „häufig“ und „gelegentlich“ von Ärzten, Apothekern und Juristen richtig numerisch interpretiert werden.

Ziel dieser Studie war es zu prüfen, ob die offizielle Klassifikation für Nebenwirkungen bekannt ist und richtig gebraucht wird.

Im Rahmen dieser empirischen Erhebung erfolgte die Befragung postalisch. Insgesamt 700 Ärzte, Apotheker und Juristen aus Schleswig- Holstein erhielten einen Fragebogen, der in vier thematische Abschnitte unterteilt war. Im ersten Abschnitt sollten 20 Wahrscheinlichkeitsbegriffe kontextfrei numerisch interpretiert werden. Aus diesen Daten wurden Rangreihen der Berufsgruppen erstellt und mit der idealen Rangreihe aus der Literatur verglichen. In den folgenden Abschnitten wurden Fallbeispiele konstruiert, in denen Begriffe oder Verhältniszahlen interpretiert werden sollten. Es wurde geprüft, ob die Interpretationen vom Kontext abhängen und ob sich die Berufsgruppen unterscheiden. Außerdem wurden die Ergebnisse mit der offiziellen Richtlinie für „Häufigkeitsangaben für Nebenwirkungen“ des Bundesinstituts für Arzneimittel und Medizinprodukte verglichen.

Von 700 Angeschriebenen haben insgesamt 430 Probanden an der Untersuchung teilgenommen. Es konnte eine Ausschöpfungsquote von insgesamt 62 % erzielt werden. Dies ist ein sehr guter Rücklauf und spricht für das Interesse an dieser Studie. Mit einem eigenhändig unterschriebenen Anschreiben, dem frankierten Rückumschlag und dem Erinnerungsschreiben konnte der Rücklauf gesteigert werden.

Auch für die Apotheker und Juristen ließen sich Rangreihen der Wahrscheinlichkeitsbegriffe erstellen, die nicht zufällig angeordnet sind und sich nur wenig von der idealen Rangreihe der Literatur unterscheiden. Außerdem wurden im Kontext „Nebenwirkungen“ die Wahrscheinlichkeitsbegriffe mit anderen Prozentwerten belegt als in der kontextfreien Interpretation. Diese Kontextabhängigkeit ließ sich auch bei den Juristen und Apothekern nachweisen.

Die erhobenen Daten beweisen, dass die offizielle Klassifikation der „Häufigkeitsangaben für Nebenwirkungen“ nicht ausreichend bekannt ist. Selbst Apotheker und Ärzte haben die verbalen Wahrscheinlichkeiten nicht richtig interpretiert. Die konsequente Darstellung der Richtlinie, in jedem Beipackzettel eines jeden Medikaments, konnte scheinbar nicht dazu beitragen, diese Definition zu verinnerlichen.

Es war bereits bekannt, dass Patienten die verbalen Wahrscheinlichkeitsbegriffe falsch interpretieren. Auch Ärzte, Apotheker und Juristen haben den fünf Begriffen eher höhere Prozentwerte zugeordnet und somit die Wahrscheinlichkeiten überinterpretiert.

Was muss man nun daraus schließen? Steht eine umfassende Aufklärung über Nebenwirkungen im Vordergrund eines Gesprächs, so sollten, wann immer möglich, die exakten numerischen Wahrscheinlichkeiten dargestellt werden. Sind diese nicht bekannt oder erscheint die numerische Darstellung nicht sinnvoll, zum Beispiel, weil dies vom Patienten nicht gewünscht ist, so sollten nur wenige klar definierte Begriffe verwendet werden. Die bestehende Richtlinie zur Darstellung von „Häufigkeiten von Nebenwirkungen“ ist eine sinnvolle Alternative zur alleinigen numerischen Darstellung. Da aber sowohl Patienten, wie auch Ärzte, Apotheker und Juristen diese Begriffe überinterpretieren, sind weitere Studien nötig, um die optimalen Definitionsbereiche zu ermitteln.

6 Literaturverzeichnis

1. Simpson, R.: Stability in Meanings for Quantitative Terms: A Comparison over 20 Years. *Quarterly Journal of Speech* 46, S. 146-151 (1963)
2. Stone, D.R. und Johnson, R.T.: A study of words indicating frequency. *Journal of Educational Psychology* 50, S. 224-227 (1959)
3. Damrosch, S.P. und Soeken, K.: Communicating probability in clinical reports: nurses' numerical associations to verbal expressions. *Research in Nursing and Health* 6, S. 85-87 (1983)
4. Merz, J.F., Druzdzal, M.J. und Mazur, D.J.: Verbal expressions of probability in informed consent litigation. *Medical Decision Making* 11, S. 273-281 (1991)
5. Marteau, T.M., Saidi, G., Goodburn, S., Lawton, J., Michie, S. und Bobrow, M.: Numbers or words? A randomized controlled trial of presenting screen negative results to pregnant women. *Prenatal Diagnosis* 20, S. 714-718 (2000)
6. Bryant, G.D. und Norman, G.R.: Expressions of probability: words and numbers. *New England Journal of Medicine* 302, S. 411 (1980)
7. Nakao, M.A. und Axelrod, S.: Numbers are better than words. Verbal specifications of frequency have no place in medicine. *American Journal of Medicine* 74, S. 1061-1065 (1983)
8. Mazur, D.J. und Hickam, D.H.: Patients' preferences for risk disclosure and role in decision making for invasive medical procedures. *Journal of General Internal Medicine* 12, S. 114-117 (1997)
9. Ohnishi, M., Fukui, T., Matsui, K., Hira, K., Shinozuka, M., Ezaki, H., Otaki, J., Kurokawa, W., Imura, H., Koyama, H. und Shimbo, T.: Interpretation of and preference for probability expressions among Japanese patients and physicians. *Family Practice* 19, S. 7-11 (2002)
10. Mazur, D.J., Hickam, D.H. und Mazur, M.D.: How patients' preferences for risk information influence treatment choice in a case of high risk and high therapeutic uncertainty: asymptomatic localized prostate cancer. *Medical Decision Making* 19, S. 394-398 (1999)
11. Shaw, N.J. und Dear, P.R.: How do parents of babies interpret qualitative expressions of probability? *Archives of Disease in Childhood* 65, S. 520-523 (1990)
12. Lichtenstein, S. und Newman, J.R.: Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic Science* 9, S. 562-564 (1967)
13. Sutherland, H.J., Lockwood, G.A., Trichter, D.L., Sem, F., Brooks, L. und Till, J.E.: Communicating probabilistic information to cancer patients: is there 'noise' on the line? *Social Science and Medicine* 32, S. 725-731 (1991)
14. Cohen, J., Dearnley, E.J. und Hansel, C.E.M.: A Quantitative Study of Meaning. *British Journal of Educational Psychology* 28, S. 141-148 (1958)
15. Kong, A., Barnett, G.O., Mosteller, F. und Youtz, C.: How medical professionals evaluate expressions of probability. *New England Journal of Medicine* 315, S. 740-744 (1986)
16. Budescu, D.V. und Wallsten, T.S.: Consistency in Interpretation of Probabilistic Phrases. *Organizational Behavior and Human Decision Processes* 36, S. 391-405 (1985)
17. Hay, J.M., Fingerhut, A., Michot, F., Flamant, Y. und Leblanc, I.: French vocabulary terms of sensitivity expressed as percentages. *Theoretical Surgery* 8, S. 181-193 (1993)

18. Mazur, D.J. und Merz, J.F.: How age, outcome severity, and scale influence general medicine clinic patients' interpretations of verbal probability terms. *Journal of General Intern Medicine* 9, S. 268-271 (1994)
19. Ferguson, W.J. und Candib, L.M.: Culture, Language, and the Doctor-Patient Relationship. *Family Medicine* 34, S. 353-361 (2002)
20. Pepper, S. und Prytulak, L.S.: Sometimes Frequently Means Seldom: Context Effects in the Interpretation of Quantitative Expressions. *Journal of Research in Personality* 8, S. 95-101 (1974)
21. Mosteller, F. und Youtz, C.: Quantifying Probabilistic Expressions. *Statistical Science* 5, S. 2-34 (1990)
22. Girone, J.A.: Further elucidation of medical jargon. *New England Journal of Medicine* 286, S. 382 (1972)
23. Hammerton, M.: How much is a large part? *Applied Ergonomics* 7, S. 10-12 (1976)
24. Suß, J.: Ist die Rangordnung von umgangssprachlichen Wahrscheinlichkeitsausdrücken in der umgangssprachlichen Kommunikation unter Ärzten konsistent? *Med. Diss. Fachbereich Humanmedizin, Philipps-Universität Marburg* (1997)
25. EU-Guideline: A guideline of Summary of Product Characteristics. (1999) <http://mednet3.who.int/prequal> (Tag des letzten Zugriffs: 23.06.2012)
26. Bundesinstitut für Arbeit und Medizinprodukte: Darstellung des aktuellen wissenschaftlichen Erkenntnisstandes über die Umsetzung der SPC-Guideline "Häufigkeitsangaben zu Nebenwirkungen". (2002) <http://www.bfarm.de> (Tag des letzten Zugriffs: 23.06.2012)
27. Berry, D.C., Raynor, D.K., Knapp, P. und Bersellini, E.: Patients' understanding of risk associated with medication use: impact of European Commission guidelines and other risk scales. *Drug Safety* 26, S. 1-11 (2003)
28. Knapp, P., Gardner, P.H., Carrigan, N., Raynor, D.K. und Woolf, E.: Perceived risk of medicine side effects in users of a patient information website: a study of the use of verbal descriptors, percentages and natural frequencies. *British Journal of Health Psychology* 14, S. 579-594 (2009)
29. Knapp, P., Raynor, D.K. und Berry, D.C.: Comparison of two methods of presenting risk information to patients about the side effects of medicines. *Quality & safety in health care* 13, S. 176-180 (2004)
30. Knapp, P., Gardner, P.H., Raynor, D.K., Woolf, E. und McMillan, B.: Perceived risk of tamoxifen side effects: a study of the use of absolute frequencies or frequency bands, with or without verbal descriptors. *Patient education and counseling* 79, S. 267-271 (2010)
31. Berry, D.C., Knapp, P. und Raynor, D.K.: Provision of information about drug side-effects to patients. *Lancet* 359, S. 853-854 (2002)
32. Moore, R.A., Derry, S., McQuay, H.J. und Paling, J.: What do we know about communicating risk? A brief review and suggestion for contextualising serious, but rare, risk, and the example of cox-2 selective and non-selective NSAIDs. *Arthritis Research & Therapy* 10, S. R20 (2008)
33. Armstrong, K., FitzGerald, G., Schwartz, J.S. und Ubel, P.A.: Using survival curve comparisons to inform patient decision making can a practice exercise improve understanding? *Journal of General Internal Medicine* 16, S. 482-485 (2001)
34. Paling, J.: Strategies to help patients understand risks. *British Medical Journal* 327, S. 745-748 (2003)
35. Timmermans, D., Molewijk, B., Stiggelbout, A. und Kievit, J.: Different formats for communicating surgical risks to patients and the effect on choice of treatment. *Patient Education and Counseling* 54, S. 255-263 (2004)

36. Timmermans, D.R., Ockhuysen-Vermeij, C.F. und Henneman, L.: Presenting health risk information in different formats: the effect on participants' cognitive and emotional evaluation and decisions. *Patient Education and Counseling* 73, S. 443-447 (2008)
37. Böcker, F. und Helm, R.: *Marketing* 7. Aufl., S. 228-239, Lucius & Lucius Verlagsgesellschaft mbH, Stuttgart (2003)
38. Diekmann, A.: *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen* 11. Aufl., S. 439 ff, Rowohlt Verlag, Reinbek bei Hamburg (2004)
39. Dillman, D.: *Mail and Telephone Surveys. The Total Design Methods*, S. 360, Wiley, New York (1983)
40. Porst, R.: Ausschöpfung bei sozialwissenschaftlichen Umfragen. Die Sicht der Institute., ZUMA- Arbeitsbericht 07/96 Mannheim. (1996)
41. Barth, S.: Die schriftliche Befragung. (1998) <http://www.stephan-barth.de> (Tag des letzten Zugriffs: 14.05.2012)
42. Siegel, S.: *Nichtparametrische statistische Methoden* 4. Aufl., S. 73-81; 193-227; 233-240, Verlag Dietmar Klotz GmbH, Frankfurt am Main (1997)
43. Bortz, J., Lienert, G.A. und Boehnke, K.: *Verteilungsfreie Methoden in der Biostatistik* 2. Aufl., S. 279-283; 329-338; 422-427; 443-444, Springer-Verlag, Berlin, Heidelberg, New York (2000)
44. Kendall, M.G.: *Rank correlation methods* 4. Aufl., S. 94-106, Griffin, London (1970)
45. Kendall, M.G. und Gibbons, J.D.: *Rank correlation methods*, S. 3-9, Oxford University Press, New York (1990)
46. Diaconis, P.: *Group Representations in Probability and Statistics. Lecture Notes-Monograph Series* 11. Aufl., S. 11 ff, Institute of Mathematical Statistics Hayward, Harvard, California (1988)
47. Everitt, B.S.: *The Cambridge Dictionary of Statistics*, CUP (2002)
48. Lienert, G.A.: *Verteilungsfreie Methoden in der Biostatistik*, S. 327-330, Meisenheim/Clau, Hain (1973)
49. Mapes, R.E.: Verbal and numerical estimates of probability in therapeutic contexts. *Social Science and Medicine* 13A, S. 277-282 (1979)
50. Man-Son-Hing, M., O'Connor, A.M., Drake, E., Biggs, J., Hum, V. und Laupacis, A.: The effect of qualitative vs. quantitative presentation of probability estimates on patient decision-making: a randomized trial. *Health Expect* 5, S. 246-255 (2002)
51. Woloshin, K.K., Ruffin, M.T.t. und Gorenflo, D.W.: Patients' interpretation of qualitative probability statements. *Archives of Family Medicine* 3, S. 961-966 (1994)
52. Cohn, L.D., Schydlower, M., Foley, J. und Copeland, R.L.: Adolescents' misinterpretation of health risk probability expressions. *Pediatrics* 95, S. 713-716 (1995)
53. Mazur, D.J. und Hickam, D.H.: Patients' interpretations of probability terms. *Journal of General Intern Medicine* 6, S. 237-240 (1991)
54. Karow, T. und Lang-Roth, R.: *Allgemeine und Spezielle Pharmakologie und Toxikologie- Vorlesungsorientierte Darstellung und klinischer Leitfaden* 16. Aufl., S. 744-746, Thomas Karow, Köln (2008)

7 Anhang

7.1 SPSS- Syntax zum Bootstrap

Programm zur Erzeugung von Bootstrap Files in SPSS/Windows:

```
GET FILE = 'c:\test.sav'.
COMPUTE ID = $CASENUM.
SAVE outfile = 'c:\test2.sav'.
INPUT PROGRAM.
LOOP SAMP = 1 to 100.
LOOP V = 1 to 182.
COMPUTE ID = TRUNC(UNIFORM(182)) + 1
END CASE.
LEAVE SAMP.
END LOOP.
END LOOP.
END FILE.
END INPUT PROGRAM.
SORT CASES BY ID.
MATCH FILES /FILE* /TABLE 'c:\test2.sav' /BY ID.
SORT CASES BY SAMP.
SPILT FILE BY SAMP.
```

NPAR TESTS /KENDALL = selten niemals unsicher sicher /MISSING LISTWISE.

7.2 Fragebogen, Anschreiben und Erinnerungsschreiben

Fragebogen

Wahrscheinlichkeiten, Prozentwerte, Abschätzungen - im Beruf wie auch im Privatleben gehen wir ganz selbstverständlich mit Ausdrücken um, die eine Wahrscheinlichkeit ausdrücken sollen. Wer gebraucht nicht Wörter wie „selten“, „häufig“, „gelegentlich“, „immer“? Doch was genau bedeutet denn nun „gelegentlich“? Viele Menschen benutzen Zahlen, um ihre Aussagen zu präzisieren, und gehen so möglichen Missverständnissen aus dem Weg. Wenn ein Arzt uns sagt, dass das Risiko, an einer Erkrankung zu leiden, bei 95% liegt, dann meint jeder, sich etwas darunter vorstellen zu können.

Sie bekommen jetzt eine Liste mit typischen umgangssprachlichen Begriffen. Ergänzen Sie bitte spontan denjenigen Prozentwert („x%“) hinter den Ausdrücken, den Sie im Sinn hätten, wenn Sie selbst im Gespräch mit Kollegen den betreffenden Begriff verwenden würden. Bei der Beantwortung des Fragebogens geht es nicht um „richtige“ oder „falsche“ Antworten, sondern um die Streuung bzw. Übereinstimmung der gegebenen Wahrscheinlichkeitswerte in bestimmten Situationen.

Antworten Sie spontan!

selten _____%	häufig _____%
niemals _____%	unklar _____%
sicher _____%	normalerweise _____%
unsicher _____%	fraglos _____%
oft _____%	stets _____%
möglicherweise _____%	nie _____%
gelegentlich _____%	wahrscheinlich _____%
überwiegend _____%	immer _____%
meist _____%	fraglich _____%
denkbar _____%	typischerweise _____%

Wollen wir einen umgangssprachlichen Begriff ausdrücken, so verwenden wir dafür entweder Worte oder Zahlen. Wie wir Wahrscheinlichkeitsbegriffe beschreiben, hängt möglicherweise von der Situation ab. Bedeutet z.B. „häufig“ in jeder Situation dasselbe für Sie? Versetzen Sie sich nun bitte in folgende Situationen. Geben Sie Ihre Antworten spontan!

Montagsmorgen, 7.30 Uhr, eine Arztpraxis irgendwo in Schleswig-Holstein... Sie sind der Hausarzt von Herrn Meier, 58 Jahre, von Beruf Mathematiklehrer. Bei der jährlichen Routineuntersuchung stellen Sie einen behandlungsbedürftigen Bluthochdruck fest. Im letzten Ärzteblatt wurde das neue Medikament „Senkefix“ vorgestellt. Da Ihr Patient sämtliche Nebenwirkungen zu diesem Medikament wissen möchte, rufen Sie Theo, einen Segelfreund und Apotheker, zu diesem Thema an:

„Senkefix“? Ja, gib ihm das Medikament ruhig. Gelegentlich tritt Schwindel auf, selten Durchfall. Sag mal - wie alt war der Patient? 58 Jahre? Naja, häufig treten Potenzstörungen auf!“

Im folgenden Gespräch mit Ihrem Patienten möchte dieser die Nebenwirkungen als Zahlenangaben von Ihnen, da er mit Worten wenig anfangen könne.

Welche Prozentangaben machen Sie zu:

häufig	_____ %
gelegentlich	_____ %
selten	_____ %

Ein Jahr später, ein Krankenhaus irgendwo in Bayern...

Sie sind Anästhesist und klären Frau Müller, die einen Unfall beim Bergsteigen erlitten hat, über die morgen anstehende OP auf. Das Medikament „Senkefix“ macht Sie hellhörig. Im letzten Ärzteblatt haben Sie etwas über OP-Komplikationen im Zusammenhang mit diesem Medikament gelesen.

Sie fragen sich, ob und wie Sie die Patientin über Risiken aufklären, die in folgenden Größenordnungen auftreten. Welche Wahrscheinlichkeitsbegriffe, also Worte, würden Sie hierfür verwenden?

1 : 1000	_____
1 : 100	_____
1 : 10	_____
1 : 5	_____
1 : 2	_____

Mittwochnachmittag, 16.00 Uhr, eine Anwaltskanzlei irgendwo in Deutschland...

Sie sind der Rechtsanwalt von Prof. Dr. Schmidt, dem ein Behandlungsfehler vorgeworfen wird. Nach genauer Prüfung der Sachlage teilen Sie ihm mit, dass der Fall möglicherweise nicht vor Gericht verhandelt werden muss. Diese Art von Rechtsstreit haben Sie schon oft im gegenseitigen Einvernehmen regeln können. Sie beruhigen Ihren Mandanten und sagen ihm, dass das Ganze wahrscheinlich in einem Monat vergessen sein wird.

Um Ihnen glauben zu können, will er diese Wahrscheinlichkeitsangaben von Ihnen in Prozentwerten hören:

möglicherweise _____%

oft _____%

wahrscheinlich _____%

Bitte geben Sie abschließend noch einige Eckdaten zu Ihrer Person an!

Geschlecht: männlich
 weiblich

Alter: _____ Jahre

Fachrichtung: Innere Medizin
 Anästhesie
 Apotheker
 Rechtsanwalt
 Richter

Einrichtung: Praxis
 Klinik
 Apotheke
 Kanzlei
 Gericht
 sonstiges _____

In welchem Bundesland sind Sie aufgewachsen?

- | | | |
|---|---|---|
| <input type="checkbox"/> Baden- Württemb. | <input type="checkbox"/> Hessen | <input type="checkbox"/> Sachsen |
| <input type="checkbox"/> Bayern | <input type="checkbox"/> Meckl.-Vorp. | <input type="checkbox"/> Sachsen-Anhalt |
| <input type="checkbox"/> Berlin | <input type="checkbox"/> Niedersachsen | <input type="checkbox"/> Schleswig-Holstein |
| <input type="checkbox"/> Brandenburg | <input type="checkbox"/> Nordrhein-Westf. | <input type="checkbox"/> Thüringen |
| <input type="checkbox"/> Bremen | <input type="checkbox"/> Rheinland-Pfalz | <input type="checkbox"/> sonstige_____ |
| <input type="checkbox"/> Hamburg | <input type="checkbox"/> Saarland | |

In welchem Bundesland haben Sie überwiegend studiert?

- | | | |
|---|---|---|
| <input type="checkbox"/> Baden- Württemb. | <input type="checkbox"/> Hessen | <input type="checkbox"/> Sachsen |
| <input type="checkbox"/> Bayern | <input type="checkbox"/> Meckl.-Vorp. | <input type="checkbox"/> Sachsen-Anhalt |
| <input type="checkbox"/> Berlin | <input type="checkbox"/> Niedersachsen | <input type="checkbox"/> Schleswig-Holstein |
| <input type="checkbox"/> Brandenburg | <input type="checkbox"/> Nordrhein-Westf. | <input type="checkbox"/> Thüringen |
| <input type="checkbox"/> Bremen | <input type="checkbox"/> Rheinland-Pfalz | <input type="checkbox"/> sonstige_____ |
| <input type="checkbox"/> Hamburg | <input type="checkbox"/> Saarland | |

In welchem Bundesland haben Sie Ihren Abschluss gemacht?

- | | | |
|---|---|---|
| <input type="checkbox"/> Baden- Württemb. | <input type="checkbox"/> Hessen | <input type="checkbox"/> Sachsen |
| <input type="checkbox"/> Bayern | <input type="checkbox"/> Meckl.-Vorp. | <input type="checkbox"/> Sachsen-Anhalt |
| <input type="checkbox"/> Berlin | <input type="checkbox"/> Niedersachsen | <input type="checkbox"/> Schleswig-Holstein |
| <input type="checkbox"/> Brandenburg | <input type="checkbox"/> Nordrhein-Westf. | <input type="checkbox"/> Thüringen |
| <input type="checkbox"/> Bremen | <input type="checkbox"/> Rheinland-Pfalz | <input type="checkbox"/> sonstige_____ |
| <input type="checkbox"/> Hamburg | <input type="checkbox"/> Saarland | |

In welchem Bundesland arbeiten Sie zur Zeit?

- Bayern
 Meckl.-Vorp.
 Schleswig-Holstein
 sonstige_____

..... Nochmals vielen Dank für Ihre Teilnahme!



Institut für Medizinische Biometrie und Statistik
UNIVERSITÄTSKLINIKUM Schleswig-Holstein
Campus Lübeck Ratzeburger Allee 160 Haus 4 23538 Lübeck

Anschrift

UNIVERSITÄTSKLINIKUM Schleswig-Holstein

Campus Lübeck

Institut für Medizinische Biometrie und Statistik

Direktor: Univ.-Prof. Dr. rer. nat. Andreas Ziegler

Tel: 0451 / 500-2780

Fax: 0451 / 500-2999

E-Mail: gockel@imbs.uni-luebeck.de

Internet: www.imbs.uni-luebeck.de

Bearbeiter: Steffi Gockel

Datum:

Betreff: Bitte um Studienteilnahme im Rahmen einer Doktorarbeit

Sehr geehrte Damen und Herren,

Hiermit bitte ich Sie, an einer Befragung im Rahmen meiner Doktorarbeit teilzunehmen, in der ich klären möchte, ob „nie“ auch immer „nie“ bedeutet.

So könnte Ihr Arzt z.B. sagen: „Einen Hirntumor haben Sie wahrscheinlich nicht.“ oder Ihr Apotheker: „Bei diesem Medikament bekommen Sie häufig Juckreiz.“ oder Ihr Anwalt sagt Ihnen: „Selten habe ich so einen Fall verloren.“ Doch was bedeutet „selten“ hier konkret? Ich möchte herausfinden, wie Ärzte, Apotheker und Juristen Wahrscheinlichkeitsbegriffe definieren und gebrauchen. Gibt es Gemeinsamkeiten und Unterschiede? Reden wir alle aneinander vorbei?

Ich studiere Medizin an der Universität zu Lübeck und führe diese Studie im Rahmen meiner Doktorarbeit durch. Die Leitung unterliegt Prof. Dr. rer. nat. Andreas Ziegler, Direktor am Institut für Medizinische Biometrie u. Statistik an der Universität zu Lübeck. Unterstützt wird die Studie u.a. von den Ärzte-kammern der Länder Schleswig-Holstein, Mecklenburg-Vorpommern und Bayern sowie den Apothekerverbänden dieser Länder. Insgesamt befrage ich ca. 2000 Ärzte, 1000 Apotheker und 1000 Juristen. Ziel ist es, die Ergebnisse in wissenschaftlichen Zeitschriften zu publizieren.

Ich bitte Sie, mich bei dieser Studie zu unterstützen, denn jede Rückantwort bessert die Aussagekraft meiner Befragung und bringt mich ein Stück weiter in Richtung Dissertation. Für die Beantwortung des Fragebogens benötigen Sie nur ca. 5 Minuten. Benutzen Sie bitte für die Rücksendung den beiliegenden Umschlag. Ihre Daten werden anonym ausgewertet und unterliegen selbstverständlich den gängigen Datenschutzregelungen. Im Übrigen verwende ich zur besseren Übersichtlichkeit ausschließlich die männliche Berufsbezeichnung.

Wenn Sie Interesse oder Fragen jeglicher Art haben, melden Sie sich bitte bei mir unter: gockel@imbs.uni-luebeck.de; Tel.: 0451-5002781

Ich danke Ihnen sehr im Voraus für Ihre Teilnahme!



Institut für Medizinische Biometrie und Statistik
UNIVERSITÄTSKLINIKUM Schleswig-Holstein
Campus Lübeck Ratzeburger Allee 160 Haus 4 23538 Lübeck

Anschrift

UNIVERSITÄTSKLINIKUM Schleswig-Holstein

Campus Lübeck

Institut für Medizinische Biometrie und Statistik

Direktor: Univ.-Prof. Dr. rer. nat. Andreas Ziegler

Tel: 0451 / 500-2780

Fax: 0451 / 500-2999

E-Mail: konrad@imbs.uni-luebeck.de

Internet: www.imbs.uni-luebeck.de

Bearbeiter: Steffi Gockel

Datum

Betreff: Erinnerungsschreiben

Sehr geehrte Damen und Herren,

Vor einer Woche haben Sie von mir einen Fragebogen zugesandt bekommen. Thema war der Gebrauch von Wahrscheinlichkeitsbegriffen unter Ärzten, Apothekern und Rechtsanwälten.

Sollten Sie den Fragebogen bereits ausgefüllt und an mich zurückgeschickt haben, bedanke ich mich sehr für Ihre Teilnahme. Falls nicht, möchte ich Sie hiermit nochmals daran erinnern. Nehmen Sie sich bitte diese Zeit – es dauert nur ca. 5 Minuten! Bedenken Sie, dass jede Rückantwort die Aussagekraft meiner Befragung bessert, und mich ein Stück weiter in Richtung Dissertation bringt.

Wenn Sie Interesse oder Fragen jeglicher Art haben, melden Sie sich einfach bei mir unter: gockel@imbs.uni-luebeck.de; Tel.: 0451-5002781

Ich danke Ihnen nochmals sehr für Ihre Teilnahme!

Mit freundlichem Gruß

7.3 Rücklaufdauer

Tabelle 51: Rücklauf des ersten Durchlaufs nach Werktagen (WT) aufgelistet

Ankunftstag	Anzahl	kumulative Anzahl	Kumulative Prozente
1.WT	0	0	0
2.WT	0	0	0
3.WT	0	0	0
4.WT	29	29	13,6
5.WT	31	60	28
6.WT	20	80	37,4
7.WT	4	84	39,3
8.WT	28	112	52,3
9.WT	0	112	52,3
10.WT	1	113	52,8
11.WT	12	125	58,4
12.WT	6	131	61,2
13.WT	3	134	62,6
14.WT	0	134	62,6
15.WT	5	139	65
16.WT	4	143	66,8
17.WT	0	143	66,8
18.WT	2	145	67,8
19.WT	4	149	69,6
20.WT	1	150	70,1
21.WT	5	155	72,4
22.WT	11	166	77,6
23.WT	12	178	83,2
24.WT	5	183	85,5
25.WT	11	194	90,7
26.WT	6	200	93,5
27.WT	0	200	93,5
28.WT	4	204	95,3
29.WT	0	204	95,3
30.WT	1	205	95,8
31.WT	1	206	96,3
32.WT	2	208	97,2
33.WT	1	209	97,7
34.WT	1	210	98,1
35.WT	0	210	98,1
36.WT	0	210	98,1
37.WT	0	210	98,1
38.WT	0	210	98,1
39.WT	0	210	98,1
40.WT	1	211	98,6
55.WT	1	212	99,1
62.WT	1	213	99,5
67.WT	1	214	100

Tabelle 52: Rücklauf des zweiten Durchlaufs nach Werktagen (WT) aufgelistet

Ankunftstag	Anzahl	Kumulative Anzahl	Kumulative Prozente
1.WT	0	0	0
2.WT	1	1	0,5
3.WT	0	1	0,5
4.WT	28	29	13,4
5.WT	35	64	29,6
6.WT	33	97	44,9
7.WT	22	119	55,1
8.WT	0	119	55,1
9.WT	0	119	55,1
10.WT	10	129	59,7
11.WT	0	129	59,7
12.WT	2	131	60,6
13.WT	7	138	63,9
14.WT	0	138	63,9
15.WT	3	141	65,3
16.WT	4	145	67,1
17.WT	3	148	68,5
18.WT	1	149	69
19.WT	0	149	69
20.WT	0	149	69
21.WT	1	150	69,4
22.WT	2	152	70,4
23.WT	0	152	70,4
24.WT	1	153	70,8
25.WT	0	153	70,8
26.WT	1	154	71,3
27.WT	1	155	71,8
28.WT	2	157	72,7
29.WT	21	178	82,4
30.WT	1	179	82,9
31.WT	2	181	83,8
32.WT	8	189	87,5
33.WT	4	193	89,4
34.WT	14	207	95,8
35.WT	0	207	95,8
36.WT	2	209	96,8
37.WT	0	209	96,8
38.WT	0	209	96,8
39.WT	0	209	96,8
40.WT	0	209	96,8
41.WT	1	210	97,2
42.WT	0	210	97,2

43.WT	1	211	97,7
44.WT	2	213	98,6
52.WT	1	214	99,1
59.WT	1	215	99,5
110.WT	1	216	100

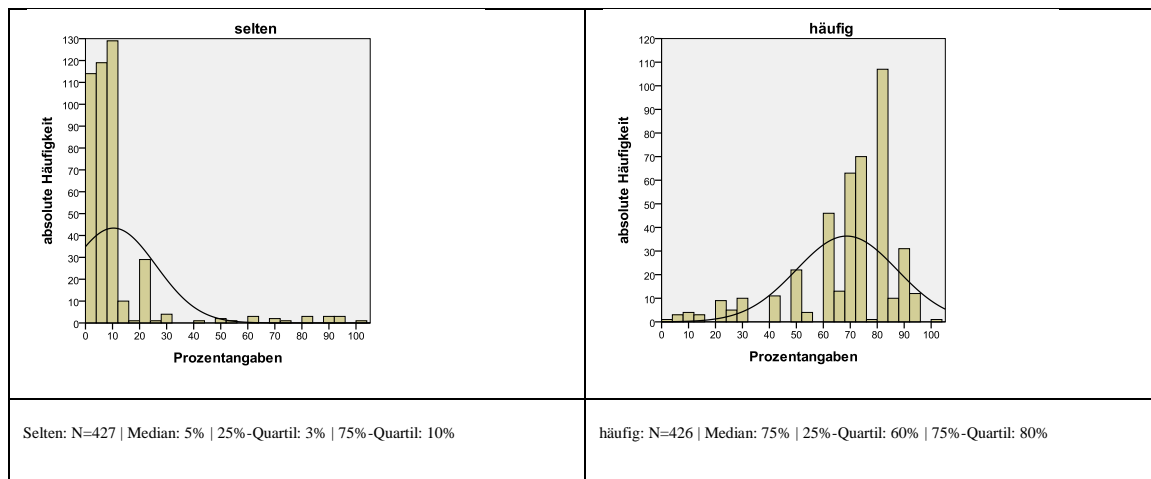
7.4 Altersverteilung nach Berufsgruppen unterteilt

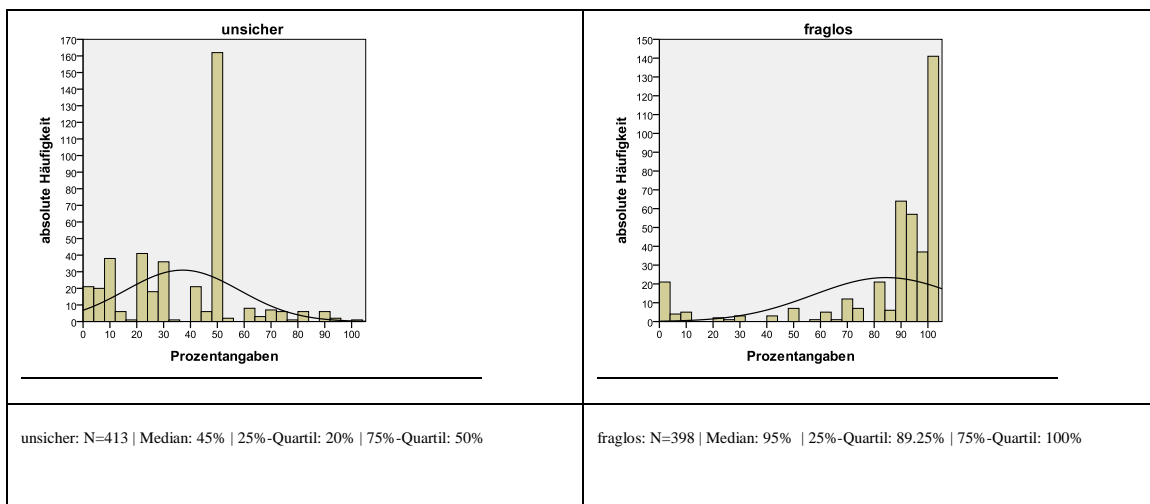
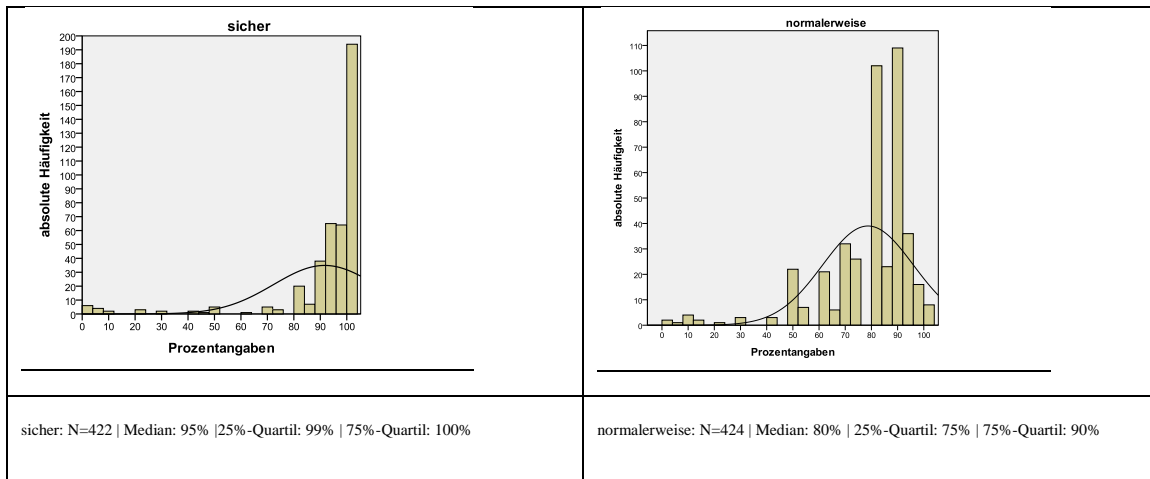
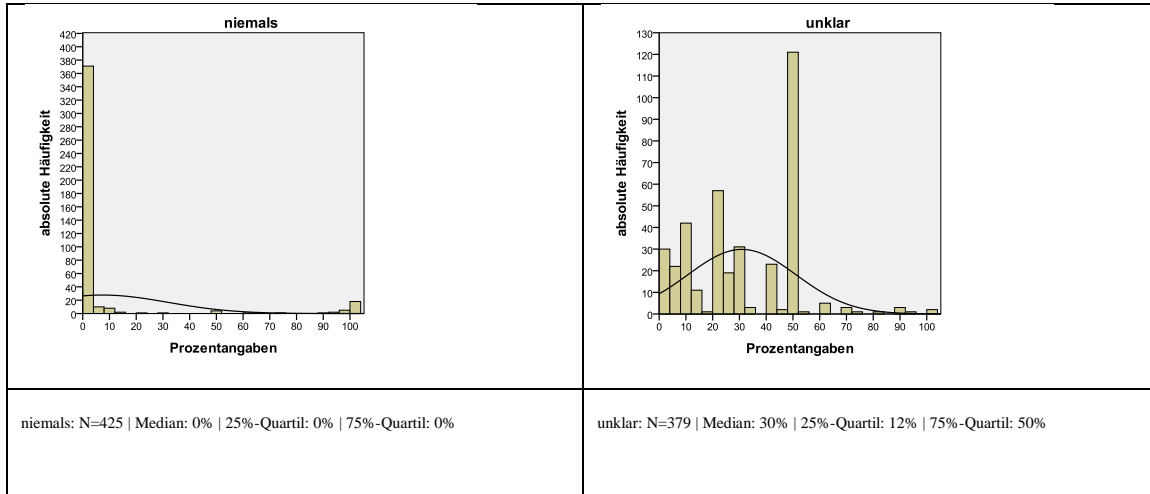
Tabelle 53: Lage- und Streuungsmaße für das Alter nach Berufsgruppen

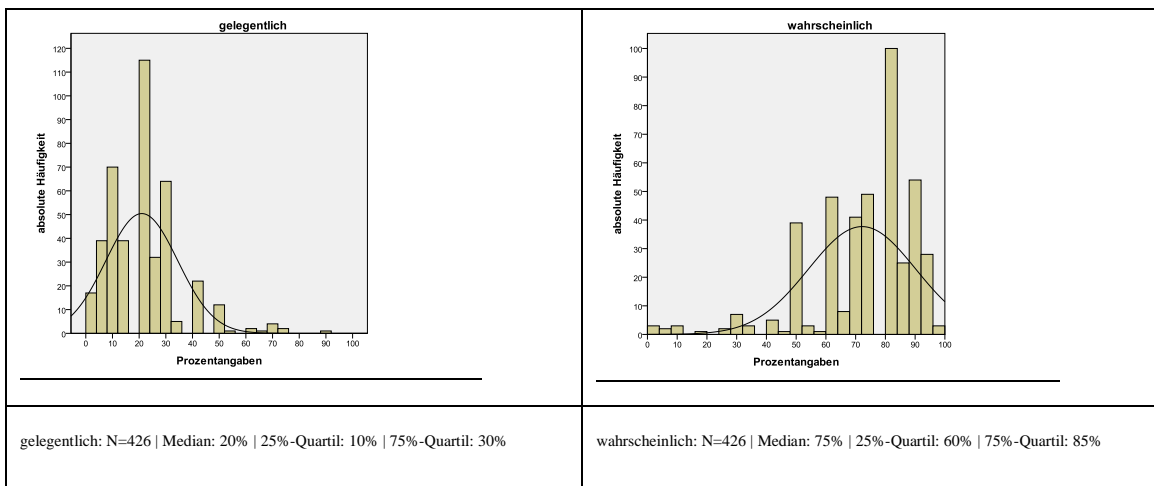
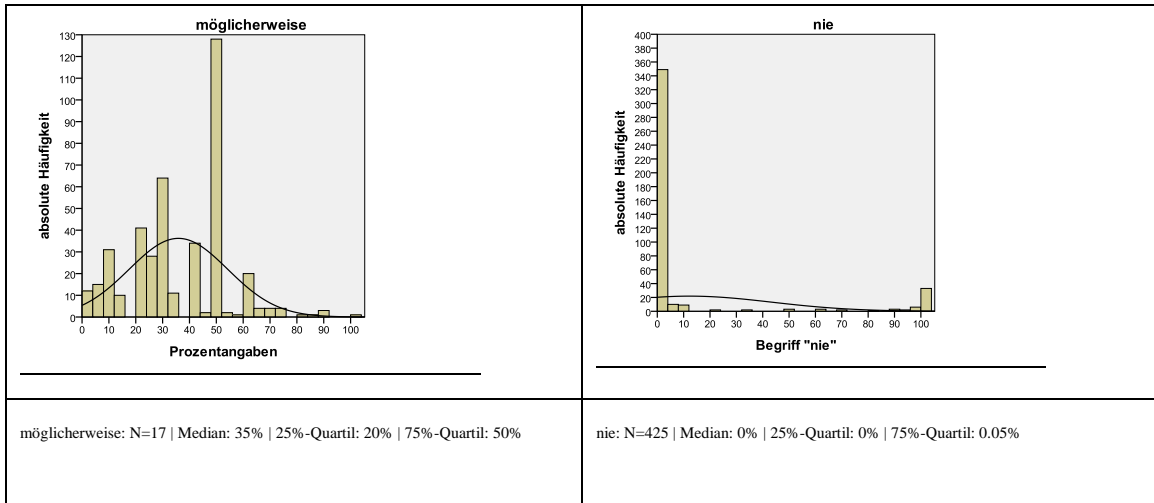
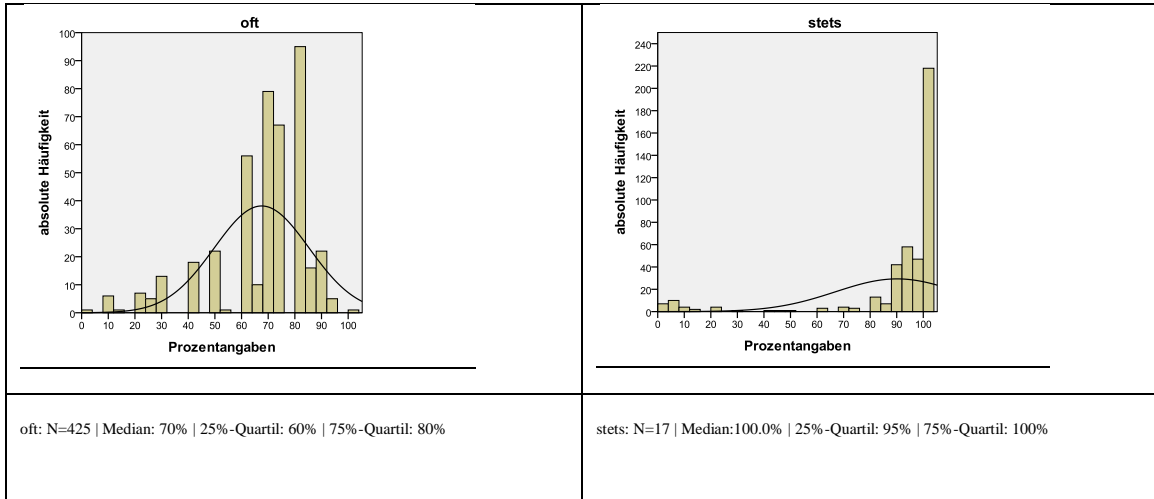
	n = Anzahl der Studienteilnehmer	Mittelwert	Median	Minimum	Maximum	25%-Quartil	75%-Quartil
Ärzte	178	50	50	33	69	45	56
Apotheker	104	49	48	27	74	41	58
Juristen	137	47	46	30	73	40	55

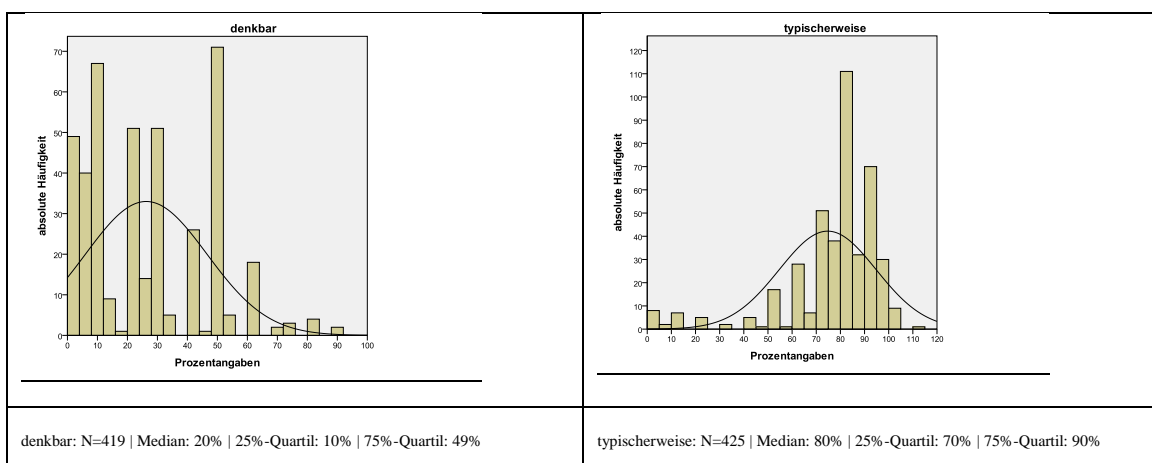
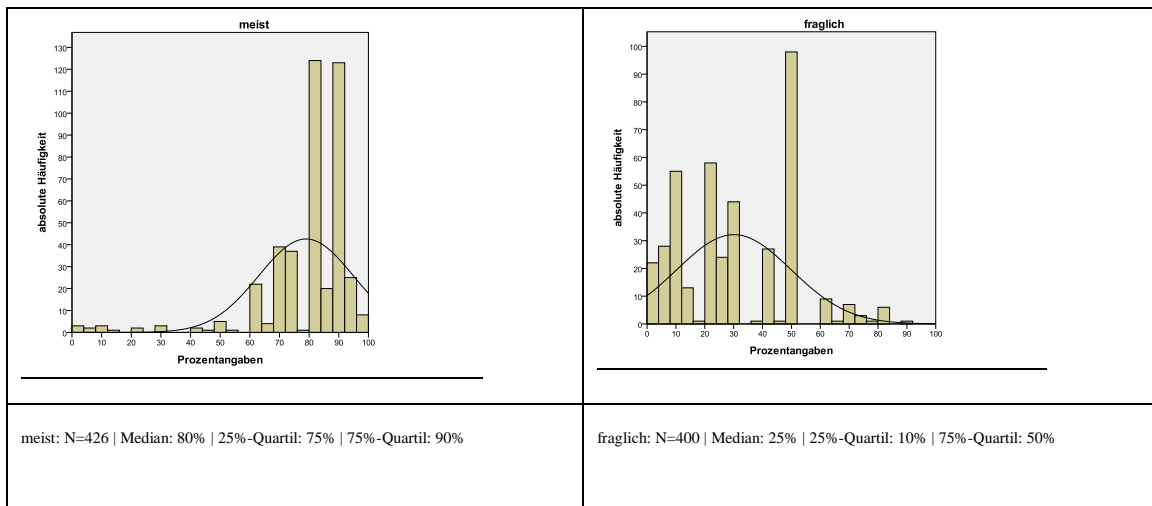
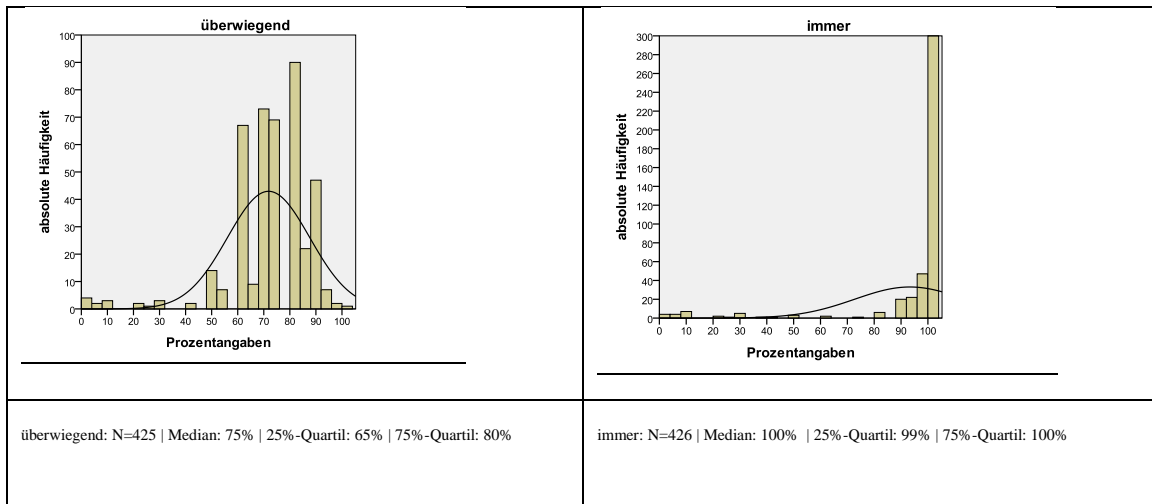
4 Ärzte haben keine Angaben zu ihrem Alter gemacht. Bei den Apothekern hat eine Personen ihr Alter nicht genannt. Von den Juristen haben drei Teilnehmer ihr Alter nicht angegeben. Die Dezimalstellen beim Mittelwert wurden nicht berücksichtigt.

7.5 Graphische Darstellung der numerischen Interpretation der 20 Wahrscheinlichkeitsbegriffe; alle Teilnehmer zusammengefasst









7.6 Vergleich der drei neuen Rangreihen mit der idealen Rangreihe der Literatur

Tabelle 54 Vergleich der neuen Rangreihen der Ärzte, Apotheker und Juristen mit der idealen Rangreihe gemäß ihrer numerischen Rangplätze

Rangplatz	ideale Reihe	Ärzte	Apotheker	Juristen
1	niemals	1	1	1
2	selten	3	3	2
3	nie	2	2	3
4	gelegentlich	4	6	4
5	unsicher	6	7	6
6	denkbar	7	4	9
7	fraglich	9	9	5
8	möglicherweise	8	5	7
9	unklar	5	8	8
10	oft	10	16	10
11	wahrscheinlich	12	13	12
12	häufig	13	12	13
13	überwiegend	11	10	11
14	meist	16	11	16
15	normalerweise	14	15	15
16	typischerweise	15	14	14
17	fraglos	17	17	17
18	stets	18	18	19
19	sicher	19	19	18
20	immer	20	20	20

7.6.1 Permutationsberechnung:

Tabelle 55 Vergleich der Rangreihen der Ärzte, Apotheker und Juristen und Berechnung der Permutationen

Rangplatz	ideale Reihe	Rangplätze			Permutationen		
		Ärzte	Apotheker	Juristen	Ärzte	Apotheker	Juristen
1	niemals	1	1	1	0	0	0
2	selten	3	3	2	1	1	0
3	nie	2	2	3	1	1	0
4	gelegentlich	4	6	4	0	2	0
5	unsicher	6	7	6	1	2	1
6	denkbar	7	4	9	1	2	3
7	fraglich	9	9	5	2	2	2
8	möglicherweise	8	5	7	0	3	1
9	unklar	5	8	8	4	1	1
10	oft	10	16	10	0	6	0
11	wahrscheinlich	12	13	12	1	2	1
12	häufig	13	12	13	1	0	1
13	überwiegend	11	10	11	2	3	2
14	meist	16	11	16	2	3	2
15	normalerweise	14	15	15	1	0	0
16	typischerweise	15	14	14	1	2	2
17	fraglos	17	17	17	0	0	0
18	stets	18	18	19	0	0	1
19	sicher	19	19	18	0	0	1
20	immer	20	20	20	0	0	0
Gesamt:					18	30	18

7.7 Graphische Gegenüberstellung der kontextfreien und der kontextabhängigen Interpretation von „möglicherweise“, „oft“ und „wahrscheinlich“

7.7.1 Gesamtes Studienkollektiv

Abbildung 13: Vergleich „möglicherweise“ für gesamtes Studienkollektiv

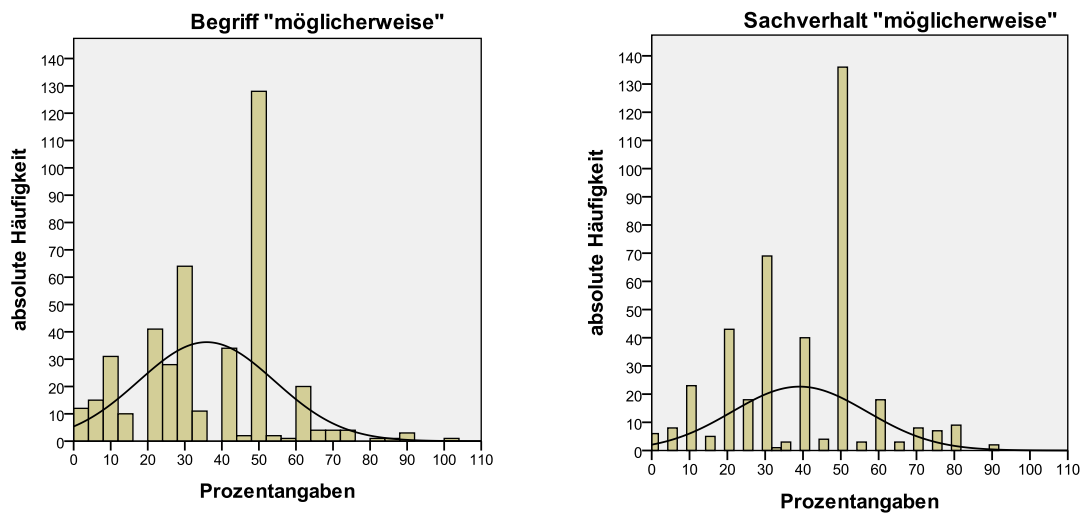


Abbildung 13: Anzahl der Nennungen von „möglicherweise“. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung 14: Vergleich „oft“ für gesamtes Studienkollektiv

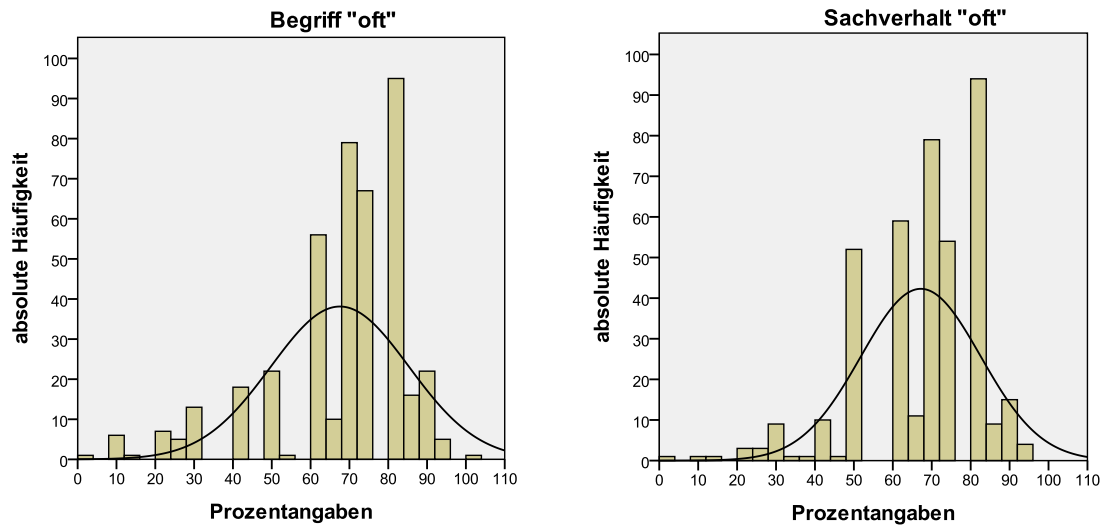


Abbildung 14: Anzahl der Nennungen von „oft“. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung 15: Vergleich „wahrscheinlich“ für gesamtes Studienkollektiv

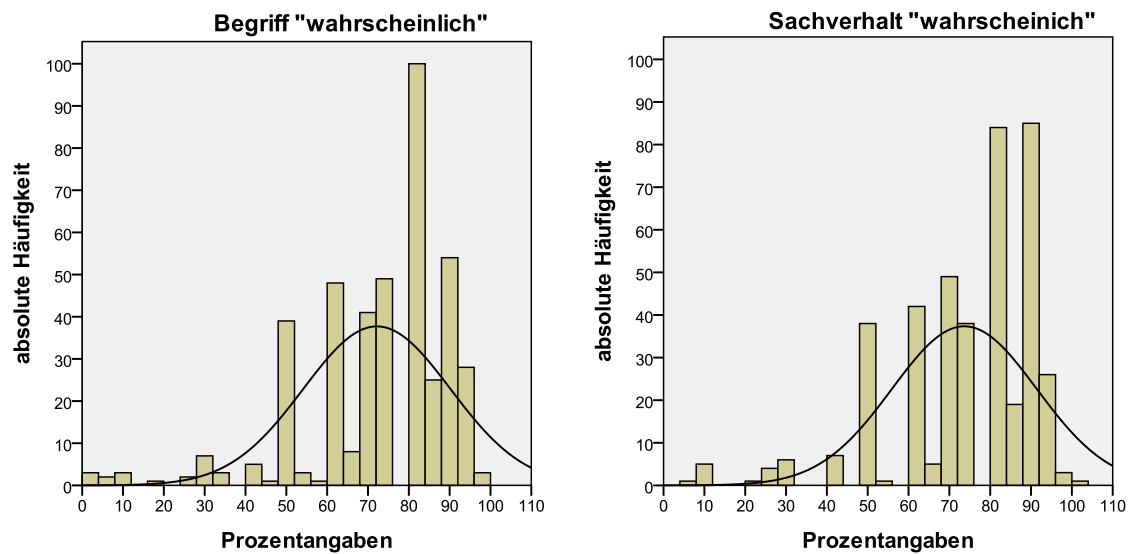


Abbildung 15: Anzahl der Nennungen von „wahrscheinlich“. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

7.7.2 Berufsgruppenspezifische Darstellung

7.7.2.1 Anzahl der Nennungen der Prozentwerte von „wahrscheinlich“

Abbildung 16: Vergleich der **Ärzte** für „wahrscheinlich“

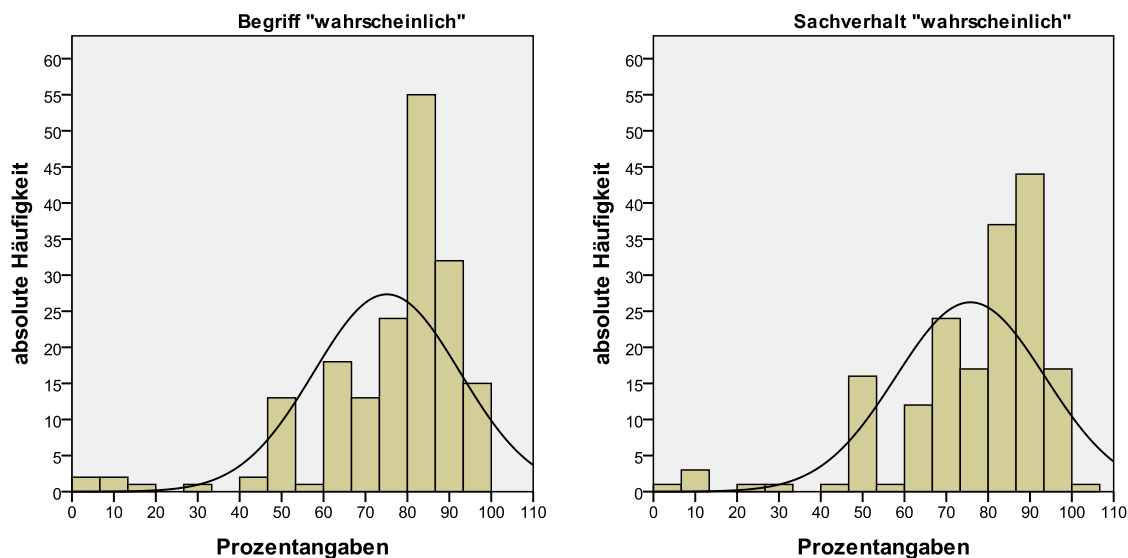


Abbildung 16: Anzahl der Nennungen von „wahrscheinlich“ in der Gruppe der **Ärzte**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung 17: Vergleich der **Apotheker** für „wahrscheinlich“

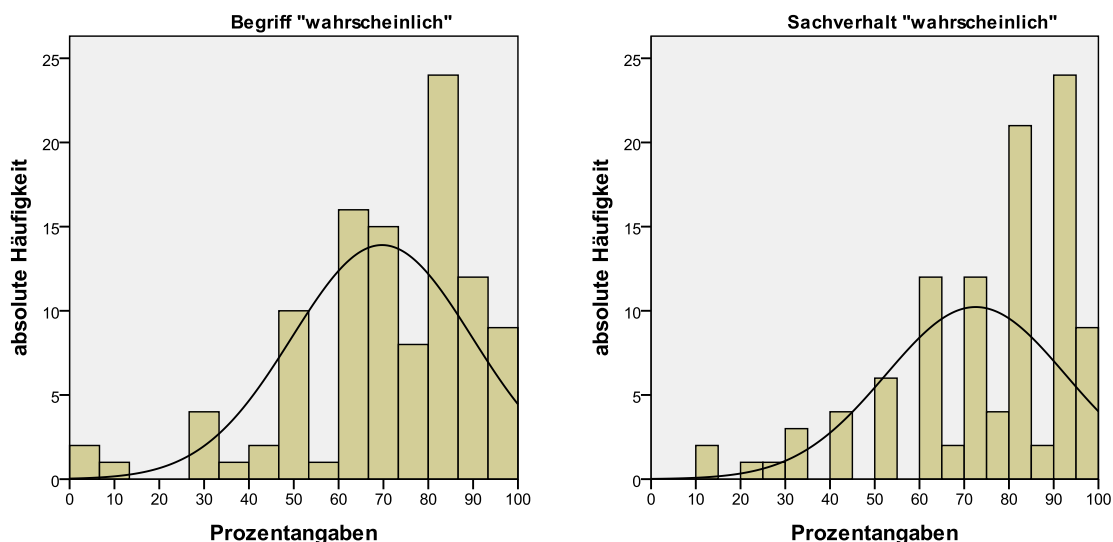


Abbildung 17: Anzahl der Nennungen von „wahrscheinlich“ in der Gruppe der **Apotheker**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

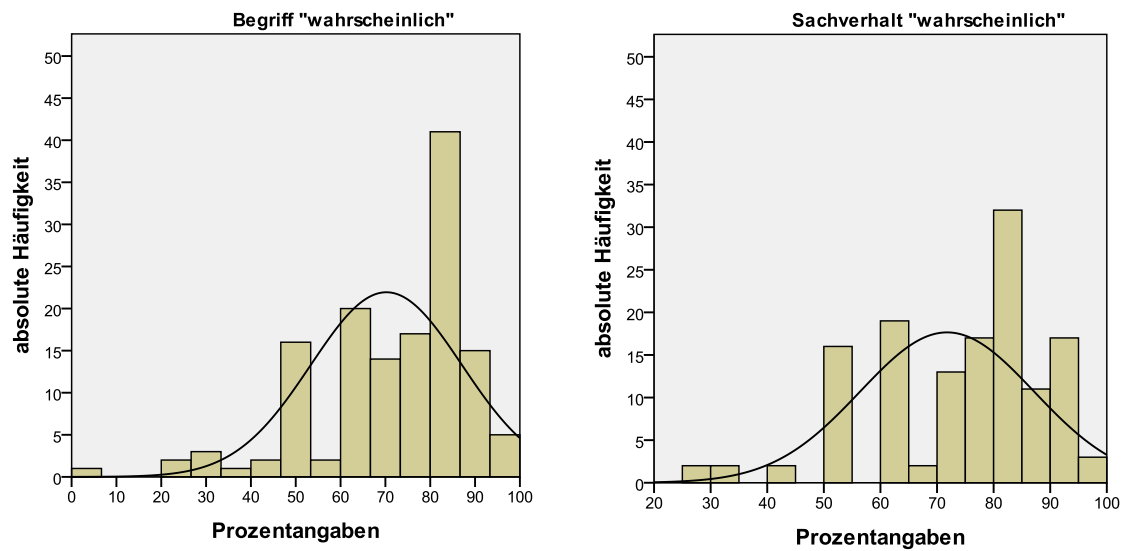
Abbildung 18: Vergleich der **Juristen** für „wahrscheinlich“

Abbildung 18: Anzahl der Nennungen von „wahrscheinlich“ in der Gruppe der **Juristen**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

7.7.2.2 Anzahl der Nennungen der Prozentwerte von „oft“

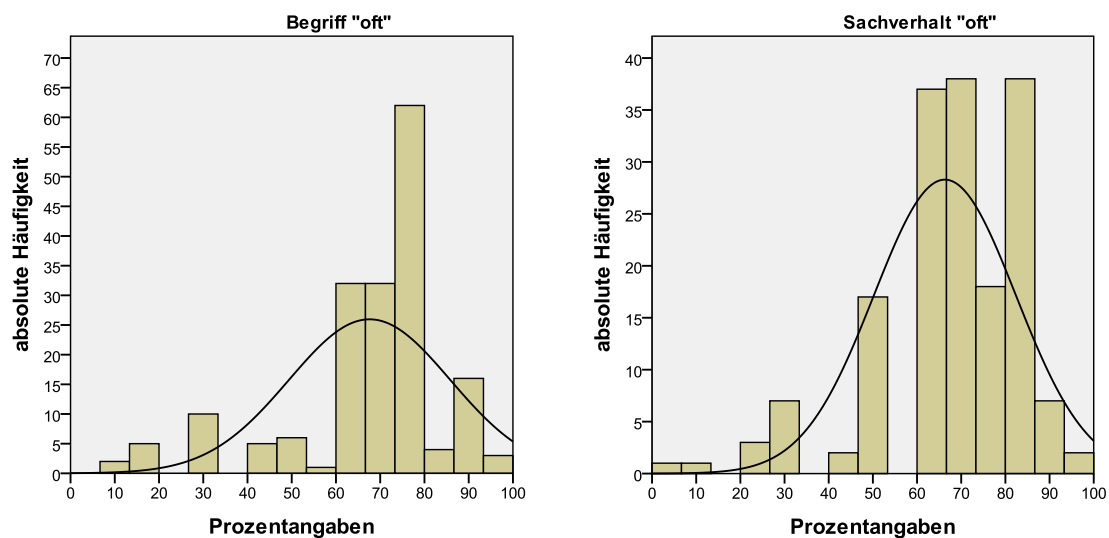
Abbildung 19: Vergleich der **Ärzte** für „oft“

Abbildung 19: Anzahl der Nennungen von „oft“ in der Gruppe der **Ärzte**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

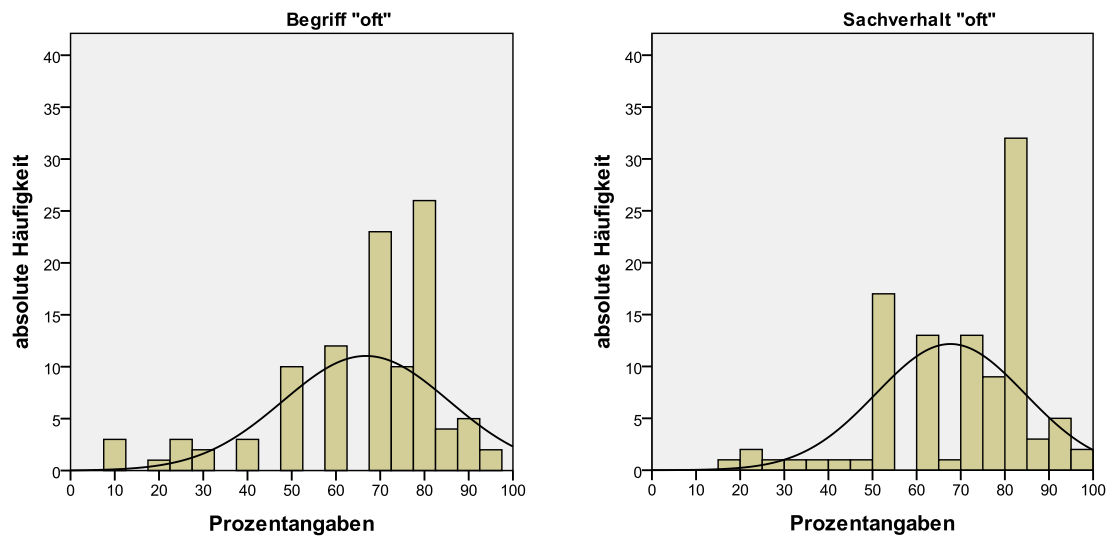
Abbildung 20: Vergleich der **Apotheker** für „oft“

Abbildung 20: Anzahl der Nennungen von „oft“ in der Gruppe der **Apotheker**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

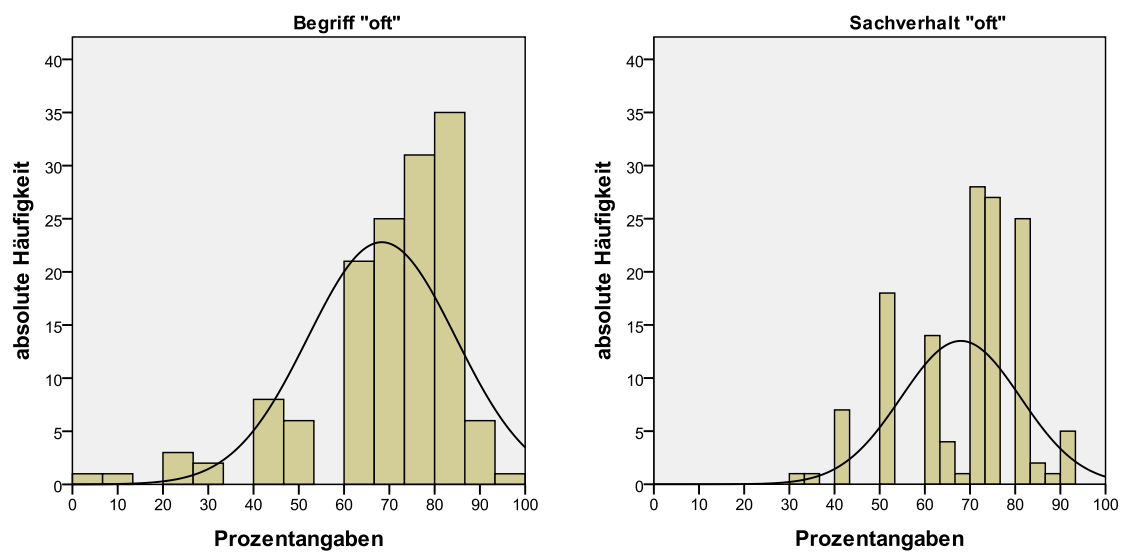
Abbildung 21: Vergleich der **Juristen** für „oft“

Abbildung 21 : Anzahl der Nennungen von „oft“ in der Gruppe der **Juristen**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

7.7.2.3 Anzahl der Nennungen der Prozentwerte von „möglicherweise“

Abbildung 22: Vergleich der **Ärzte** für „möglicherweise“

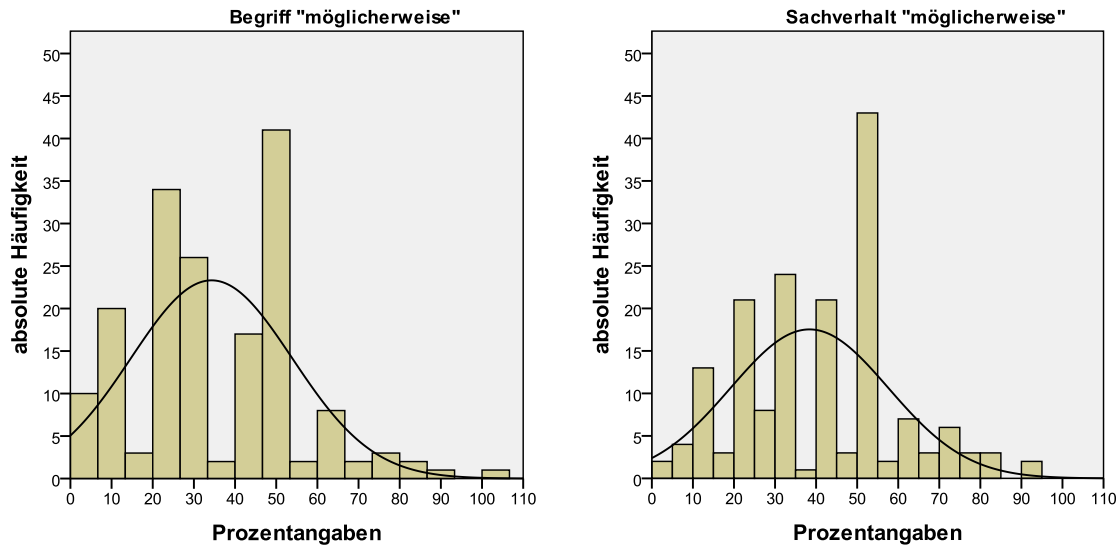


Abbildung 22: Anzahl der Nennungen von „möglicherweise“ in der Gruppe der **Ärzte**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x- Achse sind die numerischen Interpretationen in Prozent angegeben.

Abbildung 23: Vergleich der **Apotheker** für „möglicherweise“

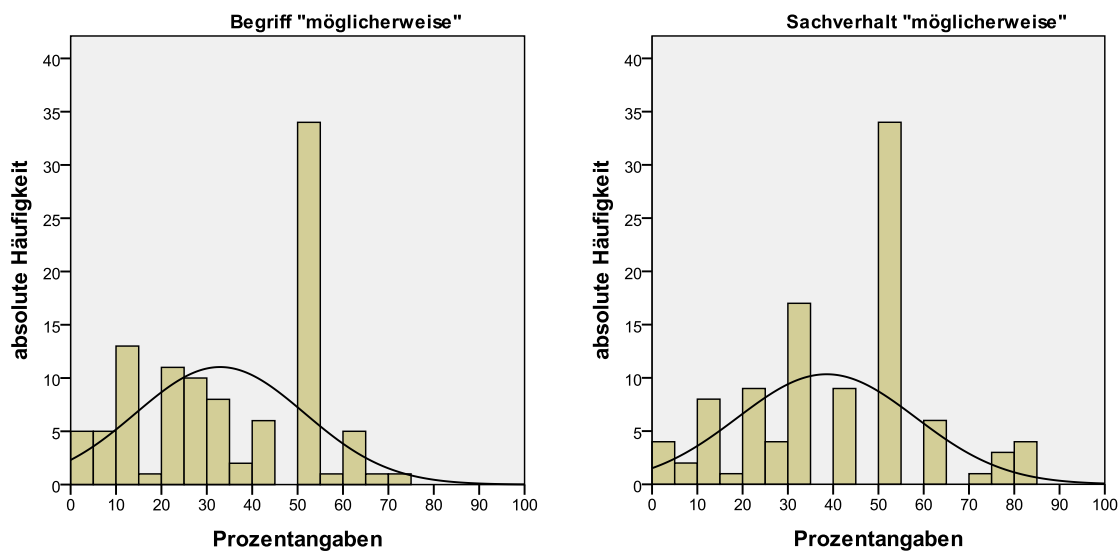


Abbildung 23: Anzahl der Nennungen von „möglicherweise“ in der Gruppe der **Apotheker**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x- Achse sind die numerischen Interpretationen in Prozent angegeben.

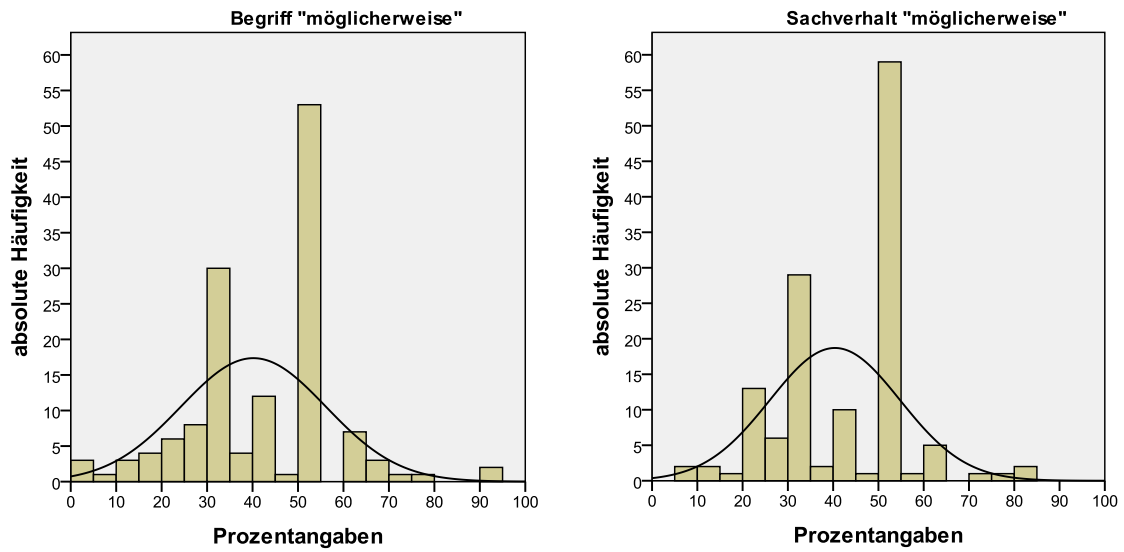
Abbildung 24: Vergleich der **Juristen** für „möglicherweise“

Abbildung 24: Anzahl der Nennungen von „möglicherweise“ in der Gruppe der **Juristen**. Links ohne Kontext, rechts kontextabhängig. Die absoluten Häufigkeiten entsprechen der Anzahl der Studienteilnehmer. Auf der x-Achse sind die numerischen Interpretationen in Prozent angegeben.

8 Danksagung

Nun möchte ich mich bei den vielen Menschen bedanken, die mich in den letzten Jahren begleitet und unterstützt haben und mir somit die Chance zur Fertigstellung dieser Arbeit gegeben haben.

Zuerst möchte ich mich bei meinem Doktorvater Herrn Prof. Dr. rer. nat. A. Ziegler für die Überlassung dieser Arbeit bedanken: Ihre Begeisterung für dieses Thema war einfach ansteckend und motivierend. Durch Ihre vielen fachlichen z.T. auch kritischen Anregungen konnten Sie auch in schwierigen Phasen des Schreibens meinen Ehrgeiz immer wieder anfeuern. Ich möchte mich herzlich für Ihre Unterstützung und Geduld in der Begleitung dieser Arbeit bedanken.

Des Weiteren möchte ich mich bei den Mitarbeitern des IMBS bedanken. Ich fühlte mich im Institut stets willkommen. Namentlich möchte ich insbesondere Frau Prof. Dr. rer. hum. biol. Inken König, Herrn Dr. Andre Kleensang, Herrn Dr. Daniel Franke, sowie Frau Christiane Wichmann, Frau Maren Frankenberger und Frau Gabrielle Schatton erwähnen.

Mischa Marchlewski, der als Dokumentationsassistent im Rahmen seines Studiums am Institut für Medizinische Biometrie und Statistik arbeitete, möchte ich für die abschließende Hilfestellungen bei der Vollendung der Statistik danken. Seine hilfreichen Ratschläge haben maßgeblich zur Fertigstellung der Arbeit beigetragen.

Außerdem möchte ich mich bei Anka Hadlak für die Zusammenarbeit zu Beginn dieser Arbeit bedanken. Besonders einprägsam blieb das stundenlange gemeinsame Ausdrucken, Frankieren und Verpacken der unzähligen Anschreiben, Fragebögen und Erinnerungsschreiben. Durch interessante und unterhaltsame Gespräche konnten wir auch diesen Teil der Arbeit meistern.

Ich möchte mich außerdem herzlich bei meinem Mann Ulf- Arne, meiner „kleinen“ und „großen“ Familie und meinen Freunden für die Unterstützung während meines Medizinstudiums und dieser Arbeit bedanken: Wenn der kleine Teufel auf meiner Schulter mal wieder die Oberhand gewonnen hatte, so habt Ihr mit Geduld, motivierenden Gesprächen und dem Glauben an mich den rotschwarzen Quälgeist auf meiner Schulter vertrieben. Ich danke Euch für Eure Unterstützung und Eure Liebe.

9 Lebenslauf

Persönliche Daten:

Steffi Mehlbeer
 geb: 10.10.1980
 geb. Gockel
 verheiratet, 2 Kinder



Schulausbildung:

1987- 1991	Grundschule Badow
1991- 1993	Gymnasiale Orientierungsstufe an der Realschule Lützow
1993- 1999	Gymnasium Gadebusch mit Abschluss der Hochschulreife

weiterer Bildungsweg:

10/99- 01/00	Psychologie-Studium an der Freien Universität zu Berlin
02/00- 09/00	Freiwilliges Soziales Jahr beim DRK Mecklenburg-Vorpommern; Klinikum Schwerin

Hochschulausbildung:

10/00	Beginn des Medizinstudiums an der Universität zu Lübeck
08/02	Physikum
08/03	1. Staatsexamen
08/05	2. Staatsexamen
10/05- 01/06	HNO-ärztliches Tertian in der HNO- Klinik, UKSH Schleswig-Holstein, Campus Lübeck
02/06- 05/06	Chirurgisches Tertian im Gesundheitszentrum Fricktal in Rheinfelden (Schweiz)
06/06- 09/06	Internistisches Tertian in der Schön Klinik Neustadt in Holstein
12/06	3. Staatsexamen

Berufsausbildung:

01/07- 01/08	Mutterschutz und Elternzeit
02/08- 11/11	Ärztin in Weiterbildung in der Inneren Medizin Schön Klinik Neustadt in Holstein
11/11- vorr. 05/13	Mutterschutz und Elternzeit