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

Der Gebrauch von umgangssprachlichen Wahrscheinlichkeitsausdrücken in der ärztlichen Tätigkeit – Sprechen Ärzte in Schleswig-Holstein die gleiche medizinische Sprache wie Ärzte in Bayern?

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der Universität zu Lübeck
-Aus der Medizinischen Fakultät-

vorgelegt von

Anka Hadlak, geb. Konrad
Wismar

Lübeck
2011

1. Berichterstatter:	Prof. Dr. rer. nat. Andreas Ziegler
2. Berichterstatter:	Prof. Dr. med. Dr. h. c. (BSMU) Peter Sterk
Tag der mündlichen Prüfung:	04.10.2011
Zum Druck genehmigt. Lübeck, den	04.11.2011

Inhaltsverzeichnis

1. Einleitung 1
   1.1 Allgemeines 1
   1.2 Überblick über die Studienlage zum Gebrauch von Wahrscheinlichkeitsbegriffen 1
   1.3 Ziel 3
2. Material und Methoden 5
   2.1 Allgemeines 5
   2.2 Aufbau des Fragebogens 5
   2.3 Auswahl der Teilnehmer 7
      2.3.1 Wahl der Bundesländer 7
      2.3.2 Wahl der Ärzte 8
   2.4 Ablauf der Studie 9
      2.4.1 Pre-Test 9
      2.4.2 Hauptfeld 10
      2.4.3 Erinnerungsschreiben 10
   2.5 Methodik 10
      2.5.1 Grundlagen und Hypothesen 10
      2.5.2 Vergleich der einzelnen Rangreihenfolgen mit Kendall`s W 12
      2.5.3 Vergleich der Rangreihen mit einer idealen Rangreihe 14
      2.5.3.1 Ideale Rangreihe 14
      2.5.3.2 Rangkorrelation 14
      2.5.3.3 Permutationstest nach Diaconis 16
      2.5.4 Vergleich der Angaben zu Begriffen in der Liste und im Kontext 16
      2.5.4.1 Wilcoxon-Test 16
      2.5.4.2 Mann-Whitney-Test 17
      2.5.5 Auswertung der Angaben zu den vorgegebenen Wahrscheinlichkeitsquoten 18
      2.5.5.1 Auswertung der Begriffe der Ärzte aus Schleswig-Holstein und Bayern 18
      2.5.5.2 Vergleich der ausgewerteten Begriffe mit der Richtlinie 19
3. Ergebnisse 21
   3.1 Ergebnisse zum Fragebogenrücklauf 21
   3.2 Soziodemographische Daten 24
   3.3 Ergebnisse zum Rangreihenvergleich – Hypothese 1 26
   3.4 Ergebnisse im Vergleich mit einer idealen Rangreihe – Hypothese 2 29

3.5 Ergebnisse zur Bewertung der Begriffe im Kontext – Hypothese 3 31
3.6 Ergebnisse zur Bewertung der Begriffe durch die Ärzte in Schleswig-Holstein und Bayern – Hypothese 4 35
3.7 Ergebnisse der Auswertung zu den Wahrscheinlichkeitsquoten 38
3.7.1 Angaben der Ärzte aus Schleswig-Holstein und Bayern 38
3.7.2 Vergleich der ausgewerteten Begriffe mit der Richtlinie 39
4. Diskussion 41
4.1 Aufbau und Ablauf der Studie 41
4.2 Ergebnisse 44
4.3 Diskussion im Überblick 47
5. Zusammenfassung 49
6. Literaturverzeichnis 51
7. Abbildungsverzeichnis 56
8. Tabellenverzeichnis 57
9. Anhang 58
9.1 Material zum Vertiefen des Themas 58
9.1.1 Anschreiben 58
9.1.2 Fragebogen 59
9.1.3 Erinnerungsschreiben 61
9.1.4 Was bedeutet die Zweite Lautverschiebung? 62
9.1.5 SPSS-Syntax zum Bootstrap 62
9.1.6 Rücklauf 63
9.1.7 Übersicht Lage- und Streuungsmaße für die 20 Wahrscheinlichkeitsbegriffe 65
9.1.8 Angaben zu Wahrscheinlichkeitsquoten 70
9.2 Danksagungen 74
9.3 Lebenslauf 75

1. Einleitung

1.1 Allgemeines

Wahrscheinlichkeiten, Prozentwerte, Abschätzungen. Im Beruf wie auch im Privatleben gehen wir ganz selbstverständlich mit Begriffen um, die eine Wahrscheinlichkeit ausdrücken sollen. Wer benutzt nicht täglich Begriffe wie „häufig“, „wahrscheinlich“ oder „selten“? Aber wie häufig ist denn nun häufig – 5%, 15% oder gar 50%?

Ein Beispiel dazu: Als Nebenwirkung bei der Einnahme von Cox-2-Inhibitoren, speziell von Rofecoxib, war das Auftreten von Myokardinfarkten mit einer Häufigkeit von 1 zu 1000 pro Jahr beschrieben worden. Es stellte sich jedoch im Verlauf heraus, das diese Nebenwirkung bei 4 von 1000 Patienten im Jahr auftrat. Damit war das absolute Risiko auf 3 zu 1000 Fälle im Jahr angestiegen. Das war „häufig“ genug, um 2004 das Medikament vom Markt zu nehmen (1).

Da wir als Ärzte täglich mit den unterschiedlichsten Menschen und ihren Anforderungen an uns zu tun haben, sind wir besonders gefordert, uns klar auszudrücken. Der Patient sollte unsere Risikoaufklärung verstehen, die Fachgespräche unter den Kollegen sollten auf der gleichen Basis stattfinden, und auch die Erklärung des Klinikapothekers zu den Nebenwirkungen eines speziell angeforderten Medikamentes sollten wir gut interpretieren können.

Aber wie werden Wahrscheinlichkeiten, speziell unter Ärzten am verständlichsten ausgedrückt – als Begriffe, als Prozentangaben oder als Quoten? Spielt der Kontext im Verständnis der Wahrscheinlichkeiten eine Rolle? Ist zum Beispiel der Beipackzettel von Medikamenten zur Häufigkeit von Nebenwirkungen, der 2003 vom Bundesinstitut für Arzneimittel und Medizinprodukte (2) konzipiert wurde, eine gute Lösung?

1.2 Überblick über die Studienlage zum Gebrauch von Wahrscheinlichkeitsbegriffen.

Seit vielen Jahren beschäftigen sich Wissenschaftler mit der Frage, wie Wahrscheinlichkeiten vor allem im medizinischen Bereich am verständlichsten dargestellt werden können.

Eine der ersten Untersuchungen wurde von *Simpson* (3) durchgeführt. Er befragte 1942 über 300 Studenten und gab ihnen eine Liste von 20 Wahrscheinlichkeitsausdrücken vor, die numerisch interpretiert werden sollten. Eine ähnliche Befragung an über 300 Studenten

erfolgte im Jahre 1962 mit 25 Wahrscheinlichkeitsbegriffen. Er stellte fest, dass die Zahlenangaben (Mittelwerte) zu bestimmten Begriffen über die Spanne von 20 Jahren stabil waren, obwohl es sich um unterschiedliche Probandengruppen handelte.

Auf Grundlage vieler weiterer Studien gab *Girone* 1972 (4) eine Liste mit der Kodierung von Wahrscheinlichkeitsausdrücken heraus.

In den Jahren nach 1980 war immer deutlicher geworden, dass Begriffe bei der Umschreibung von Wahrscheinlichkeiten vor allem im medizinischen Bereich zu ungenau waren und zu Schwierigkeiten im Verständnis insbesondere zwischen Arzt und Patient führten (5, 6). *Kong* (7) hatte allerdings bemerkt, dass Begriffe individuell bevorzugt benutzt wurden, da eine gewisse Ungenauigkeit gewünscht sei. Dadurch wären allerdings Missverständnisse zwischen Arzt und Patient vorprogrammiert. *O`Brien* (8) und *Reagan* (9) fanden in ihren Studien 1989 eine große Konsistenz in der Zuordnung von Wahrscheinlichkeitsbegriffen zu numerischen Ausdrücken, zumindest hinsichtlich der Mediane – die Variationsbreiten waren jedoch sehr hoch. Zudem war *O`Brien* aufgefallen, dass zwar „niemals“ mit 0% bewertet wurde, „sicher“ jedoch nur mit 95%!

In den neunziger Jahren war die große Variationsbreite bei der Zuordnung von Zahlen zu Begriffen weiterhin ein spannendes Thema (10). Den Wissenschaftlern war aufgefallen, dass es vor allem für geringe Wahrscheinlichkeiten und Darstellung von kleinen Differenzen keine eindeutigen verbalen Umschreibung gab (11). *Mazur* fand 1994 in einer weiteren Studie (12) heraus, dass es für Wahrscheinlichkeiten unter 1%, die in der Medizin regelmäßig vorkommen, keinen eindeutigen Begriff gab. Weiterhin spielten Alter, persönliche Lebenserfahrung, ethnische Herkunft und die Muttersprache eine zu große Rolle bei der Interpretation von Begriffen (13). *Mazur* (14) und *Edwards* (15) bestätigten diese Ergebnisse 1997, *Ferguson* kam mit seinen Untersuchungen 2002 zum gleichen Ergebnis (16).

Eine weitere entscheidende Rolle spielt der Kontext, in welchem die Begriffe benutzt werden. Eine wichtige Studie hierzu gab es bereits 1974 von *Pepper* (17). Als Beispiel führte er ein Erdbeben an, welches beim Auftreten zweimal im Monat am selben Ort als „häufig“ zu bezeichnen sei, wohingegen dieses für eine Schießerei in einem Western nicht zutreffen würde, sofern sie nur zweimal stattfand. Die Kontextabhängigkeit wurde immer wieder in vielen Studien bestätigt (z.B. 7, 13, 18, 19, 20, 21).

Deshalb wurde in vielen Studien bereits die gemeinsame Verwendung von verbalen und entsprechenden numerischen Begriffen gefordert, da eine Rangreihenfolge der Begriffe im Verständnis der Anwender vorhanden war (3, 22, 23, 24).

Als Ergebnis dieser Bemühungen war 2002 vom Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) die EU-Richtlinie Summary of Product Characteristics (SPC) von 1999 zu „Häufigkeitsangaben von Nebenwirkungen“ umgesetzt worden (2, 25). Die Angabe von Nebenwirkungen erfolgt seitdem sowohl in Zahlen (Quoten) als auch verbalisiert anhand von fünf Graduierungen. Dieses war die Grundlage für weitere Studien. *Paling* (26) hielt den Ansatz dieser Form der Beschreibung für gut, hinterfragte jedoch das Verständnis kritisch. *Berry* (27) ließ das Verständnis der Zahlenangaben in der EU-Richtlinie ebenfalls überprüfen. Auch hier gab es große Missverständnisse in der Übersetzung. 2007 wurden die Angaben vom BfArM erneut überarbeitet. Positiv wurde diese Auswertung zum Beispiel von *Burkiewicz* interpretiert (28). Als Problem sah *Moore* dagegen, dass die dargestellten Risiken in Form der genannten EU-Richtlinie durch diese Darstellung überbewertet werden (29). Entscheidungen würden eher auf Emotion als auf Fakten basieren, deshalb werden gewöhnliche Risiken eher unterschätzt und seltene überschätzt.

Epstein machte den Vorschlag, Wahrscheinlichkeiten zusätzlich mit Hilfe von Graphiken darzustellen (30). Auch *Timmermanns* untersuchte diese Möglichkeit und bestätigte sie als gute Lösung (31, 32). Letztendlich stellte er jedoch auch heraus, dass völlig unklar sei, welches das beste Format ist. *Lipkus* forderte deshalb alle drei Formate in Kombination miteinander einzusetzen (33).

Eine sehr gute Übersichtsarbeit zu allen oben genannten Themen wurde von *Visschers* verfasst (34).

Risikokommunikation ist komplex (35). Funktioniert sie zum Beispiel im medizinischen Bereich zwischen Arzt und Patient, so kann das entscheidend für sämtliche Behandlungen sein. Gibt es im medizinischen Bereich Missverständnisse, kann dieses unter Umständen sogar Leben kosten.

Was bedeutet schon „sicher“ (36) oder nach *Moore* `The only certainty is that there is uncertainty´ (29)?

1.3 Ziel

Die folgende Arbeit beschäftigt sich mit dem Verständnis und dem Gebrauch von Wahrscheinlichkeitsausdrücken in der ärztlichen Tätigkeit.

Es wurden Ärzte der Fachrichtungen Anästhesiologie und Innere Medizin aus Schleswig-Holstein und Bayern mittels eines Fragebogens befragt.

Ich habe das Verständnis von 20 Wahrscheinlichkeitsbegriffen, Prozentangaben von Häufigkeiten und die Verständlichkeit von den entsprechenden Quoten durch die verschiedenen ärztlichen Fachrichtungen untersucht. Die Begriffe wurden sowohl kontextunabhängig in einer Liste vorgegeben, als auch zum Vergleich in einem bestimmten Sachverhalt abgefragt.

In dieser Studie wurde auch überprüft, ob es tatsächlich eine einheitliche Reihenfolge im Sinne von Rängen (Rangreihenfolge) für die vorgegebenen Begriffe unter den verschiedenen Ärzten gibt und ob diese Rangreihenfolge mit einer in der Literatur berichteten idealen Rangreihenfolge übereinstimmt. Eine weitere Frage war, ob es einen Unterschied darin gibt, dass die Angaben von Ärzten in Bayern oder von Ärzten in Schleswig-Holstein gemacht werden und inwieweit die Antworten von einem vorgegebenen Kontext abhängen. Schließlich wurde das Verständnis der Quotenangaben zu den Häufigkeiten von Nebenwirkungen im Sinne der EU-Richtlinie unter den verschiedenen Ärztegruppen untersucht.

Diese Arbeit soll einen Beitrag zur besseren Verständlichkeit im Umgang mit Wahrscheinlichkeitsausdrücken leisten sowie eventuell bestehende Probleme, vor allem im deutschsprachigen Raum, aufdecken.

2. Material und Methoden

2.1 Allgemeines

Diese Arbeit ist Teil einer umfangreicheren Studie am Institut für Medizinische Biometrie und Statistik der Universität zu Lübeck, in welcher insgesamt Ärzte aus Schleswig-Holstein und Bayern sowie Apotheker und Rechtsanwälte beziehungsweise Richter aus Schleswig-Holstein befragt wurden. Die vorliegende Arbeit beschäftigte sich ausschließlich mit dem Antwortverhalten der Ärzte.

Für die vorliegende Studie ist die schriftliche Befragung als Methode der Wahl zum Einsatz gekommen. Aufwand und Kosten sind hierbei im Allgemeinen geringer als bei einer persönlichen oder telefonischen Befragung. Außerdem gestaltet sich die logistische und personelle Umsetzung einfacher, das Antreffen der zu Befragten ist sichergestellt, und die Beantwortung der Fragen geschieht unabhängig von einem Interviewer. Eine mündliche oder telefonische Befragung hätte wahrscheinlich trotz höherer Kosten inhaltlich kein anderes Ergebnis erbracht (37).

Die schriftliche Befragung erfolgte nach erfolgreichem Pre-Test in zwei Wellen im Abstand von sechs Wochen. Die erste Welle startete am 02.08.2004, die zweite Welle folgte am 16.09.2004.

Es wurden jeweils 300 Fragebögen verschickt, davon 150 an Ärzte in Schleswig-Holstein und 150 an Ärzte in Bayern. Bei der Erstellung des Fragebogens wurde sich an den internationalen Richtlinien zur Fragebogenerstellung orientiert (38). Vor allem wurde großer Wert auf die Übersichtlichkeit sowie die Verständlichkeit der Fragen und Sachverhalte gelegt. Die Fragebögen wurden mit einem persönlichen Anschreiben versehen und eigenhändig unterschrieben (Anlage 9.1.1). Weiterhin wurde einen Rückumschlag für die portofreie Rücksendung des Fragebogens beigelegt.

Die Rücklaufdauer für die Auswertung war auf 50 Werktage festgesetzt worden.

2.2 Aufbau des Fragebogens

Der Fragebogen wurde in Anlehnung an die Arbeit von *Suß* 1997 (39) weiterentwickelt und teilnehmerfreundlicher gestaltet. Wichtig für die Beantwortung sämtlicher Fragenstellungen im Bogen war, dass die Teilnehmer ihre Antworten möglichst spontan geben sollten.

Der Fragebogen gliederte sich in vier Abschnitte:

Im ersten Abschnitt (Abschnitt A) wurde eine Liste von 20 Wahrscheinlichkeitsbegriffen vorgegeben. Die Auswahl der Liste dieser Begriffe basierte auf Arbeiten, die sich ebenfalls mit der Anwendung und dem Gebrauch von Wahrscheinlichkeitsbegriffen beschäftigt haben (39, 40).

Folgende Wahrscheinlichkeitsbegriffe wurden vorgegeben:

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

Für diese Begriffe sollte eine äquivalente Prozentangabe (0 – 100%) gemacht werden, wie sie zum Beispiel auch im täglichen Gespräch mit Kollegen benutzt werden könnte. Zum Begriff „selten“ könnte zum Beispiel der Wert von 5% eingetragen werden, zum Begriff „immer“ beispielsweise 100%.

Im zweiten Abschnitt wurden drei Sachverhalte des täglichen Lebens beschrieben.

Im ersten Sachverhalt (Abschnitt B) wurde eine Situation in einer Arztpraxis geschildert. Dabei ging es um die Häufigkeit von Nebenwirkungen von einem zu verschreibenden Medikament. Zu diesem Sachverhalt waren die drei Begriffe „häufig“, „gelegentlich“ und „selten“ vorgegeben. Es wurde in Anlehnung an den ersten Abschnitt nach der Interpretation der Begriffe durch die Angabe in Prozentzahlen gefragt.

Im zweiten Sachverhalt (Abschnitt C) wurde eine Situation in einer Klinik dargestellt. Es handelte sich um eine Aufklärung zur Operation und das Auftreten von Komplikationen bei einem bestimmten Eingriff. Zu diesem Zweck wurden Verhältniszahlen im Sinne von Chancen oder Chancenverhältnissen (im Deutschen ist dieser Begriff ungebräuchlich, im Englischen wir der Begriff der `odds´ oder `odds ratio´ verwendet) vorgegeben. Mit solchen „Quoten“ wird zum Beispiel bei Pferdewetten gearbeitet, oder man findet sie mittlerweile auch auf jedem Beipackzettel von verschreibungspflichtigen Medikamenten. Im vorgegebenen Sachverhalt wurde gefragt, wie die Größenordnungen von „1:1000“, „1:100“, „1:10“, „1:5“ und „1:2“ mit Hilfe von äquivalenten Wahrscheinlichkeitsbegriffen im Sinne einer Risikowahrscheinlichkeit umschrieben werden können.

Im dritten Sachverhalt (Abschnitt D) wurde eine Situation zu einem medizinischen Rechtsstreit beschrieben. Es ging hierbei allgemein um den Gebrauch von Wahrscheinlichkeitsbegriffen. Hierzu wurden analog zum ersten Sachverhalt drei Häufigkeitsbegriffe angeboten, die durch äquivalente Prozentangaben umschrieben werden sollten. In diesem Sachverhalt waren die Begriffe „möglicherweise“, „oft“ und „wahrscheinlich“ vorgegeben.

Im letzten Abschnitt (Abschnitt E) wurden demographische Daten der Teilnehmer erhoben. Gefragt wurde nach dem Alter, Geschlecht, der Fachrichtung und Einrichtung, in der die Teilnehmer tätig sind. Zusätzlich wurde danach gefragt, in welchem Bundesland die Teilnehmer aufgewachsen sind, wo sie überwiegend studiert und ihren Abschluss gemacht haben und in welchem Bundesland sie derzeit arbeiten.
Der vollständige Fragebogen ist im Anhang enthalten (Anlage 9.1.2).

2.3 Auswahl der Teilnehmer

2.3.1 Wahl der Bundesländer

Die Daten habe ich durch Befragung von Ärzten in Schleswig-Holstein und Bayern erhoben. Schleswig-Holstein ist das Bundesland, in dem ich selbst studiert habe und derzeit arbeite, so dass diese Wahl nahe liegend war. Bayern wurde zum einen deshalb gewählt, weil es flächenmäßig das größte Bundesland ist und somit am ehesten repräsentative Antworten bieten kann und zum anderen, weil die allgemeinen Sprachgewohnheiten gänzlich zum Norden verschieden sind. Sprachgeschichtlich könnten diese beiden Bundesländer nicht verschiedener sein. An dieser Stelle erfolgt ein kurzer Exkurs in diese sprachlich unterschiedliche Entwicklung, um die Wahl der beiden Bundesländer zu begründen. Grundsätzlich gehen alle deutschen Mundarten als Elemente des sogenannten Dialektkontinuums aus den germanischen Sprachen hervor (41). Diese reichen vom äußersten Norden mit der Sprachgrenze zum Dänischen bis zum äußersten Süden mit der Sprachgrenze zum Italienischen, zum Osten mit der Sprachgrenze zu den slawischen Sprachen und zum Westen mit der Sprachgrenze zum Französischen hin. Daraus haben sich im Laufe der Jahre unter anderem eine niederdeutsche (meint niedergelegenere Landschaften) und eine hochdeutsche (meint höhergelegenere Landschaften) Standardsprache entwickelt. Die niederdeutsche Sprache war vor allem die Handelssprache der Hanse (42). Durch Sprachwandlungen, die als „Zweite Lautverschiebung“ bezeichnet werden, kam es im Mittelalter im heutigen hochdeutschen Raum zu einer starken Umgestaltung, so dass das ursprüngliche (Nieder-)Deutsche durch das Hochdeutsche als Standardsprache verdrängt wurde. Das heutige moderne Standarddeutsch ist eine Variante des Hochdeutschen, weshalb die beiden Begriffe oft synonym verwendet werden. Hochdeutsch ist genau wie Ober-, Mittel- und Niederdeutsch ein topographisch motivierter Ausdruck. Dialekte stehen eigenständig im Gegensatz zur Hochsprache. Mundart ist ein künstliches Gelehrtenwort in der Sprachwissenschaft und meint im Süden den Dialekt und im Norden das Platt (-deutsche) (42). In den

niederdeutschen Dialekten werden zumindest in der Alltagssprache bestimmte Laute noch weitgehend so ausgesprochen wie vor 2000 Jahren, da diese ohne die Zweite Lautverschiebung blieben, während die oberdeutschen Dialekte relativ nah an die Standardsprache heran reichen (41). Zur besseren Verständlichkeit der Sprachunterschiede findet sich eine Karte im Anhang (Abschnitt 9.1.4). Im Rahmen dieser ganzen Entwicklung wurden im Norden die niederdeutschen Dialekte durch das Hochdeutsche als Standarddeutsch überlagert und überwiegend verdrängt. Hinzu kam, dass im niederdeutschen Dialektraum das mundartliche Sprechen einen relativ niedrigen Status hatte und die Sprechweise der Mittelschicht sich stark am Standard orientierte, während sich diese Entwicklung im oberdeutschen Dialektraum nicht durchsetzte. Angehörige der Mittelschicht sprachen weiterhin dialektnah, da das mundartliche Sprechen in diesen Regionen eine hohe Wertschätzung besaß (41). Heute ist es so, dass im Norden das Platt in größeren öffentlichen Situationen völlig ausgeschaltet ist und nur noch in wirklich ländlichen Regionen gesprochen wird. Es kann in der Schule als Fremdsprache erlernt werden, da der Unterschied zwischen Mundart und Schriftsprache zu groß ist. Im Süden ist die gesprochene Hochsprache deutlich mundartlich gefärbt und die Sprechweisen gehen ineinander über (42). Bairisch als Fremdsprache in der Schule gibt es nicht.

Im öffentlichen Leben herrscht eine einheitliche Hoch- beziehungsweise Standardsprache vor (Duden), die bereits jeder in der Schule erlernt, und an der sich auch die Medizinerausbildung und fachliche Fort- und Weiterbildung orientiert. Trotzdem gibt es die unterschiedlichen Dialekte, und durch die Wahl genau dieser beiden Bundesländer sollten Unterschiede und Gemeinsamkeiten besonders deutlich werden.

2.3.2 Wahl der Ärzte

Ich habe Ärzte aus den Fachrichtungen der Inneren Medizin und der Anästhesie befragt. Durch die Wahl genau dieser beiden Fachrichtungen sollte ein ausgewogenes Verhältnis zwischen Praxis- und Klinikärzten erreicht werden. Für die Auswahl der Internisten war der ständige und enge Patientenkontakt ausschlaggebend, für die Auswahl der Anästhesisten vor allem der innerklinische Personenumgang.

Die Adressen der Ärzte aus Bayern wurden freundlicherweise von der Bayerischen Landesärztekammer als Adresslisten zugeschickt. Es handelte sich hierbei um weiterbildungsbefugte Ärzte. Insgesamt waren es circa 900 Adressen für die Internisten und etwa 300 Adressen für die Anästhesisten. Von der Landesärztekammer in Schleswig-Holstein wurde lediglich auf deren Homepage verwiesen (43). Über diese Seite konnten

rund 450 Adressen für die Internisten und circa 130 Adressen für die Anästhesisten ausfindig gemacht werden.

Für die Befragung wurden 600 Adressen aus diesem Pool ausgewählt. Die Auswahl erfolgte zufällig mit Hilfe von Microsoft Excel, Version 2003. Im ersten Durchgang wurden jeweils 75 Internisten und Anästhesisten pro Bundesland angeschrieben. Im zweiten Durchgang war die Verteilung aufgrund fehlender Adressen zugunsten der Internisten in Schleswig-Holstein verschoben. Somit wurden 98 Internisten und nur 52 Anästhesisten in Schleswig-Holstein angeschrieben, in Bayern waren es wiederum jeweils 75 Ärzte der entsprechenden Fachgruppe.

2.4 Ablauf der Studie

2.4.1 Pre-Test

Der Pre-Test ist bei umfangreichen empirischen Forschungsprojekten Voraussetzung für eine optimale Gestaltung des Hauptfeldes. Dennoch gibt es trotz genauer Vorbereitung und Auswahl der besten Methode keine Garantie, dass die Studie zum Beispiel hinsichtlich des erwarteten Rücklaufs gelingt oder dass der Befragte die Befragungsunterlagen auch wirklich versteht und selbständig beantwortet. Üblicherweise werden die Befragungsunterlagen im Pre-Test einem Teil der ausgewählten Grundgesamtheit zugestellt, von dieser bearbeitet und zurückgesandt. Bei der Auswahl der Pre-Test-Stichprobe ist auf die demographische Bandbreite hinsichtlich entscheidender soziodemographischer Merkmale der ausgewählten Grundgesamtheit zu achten. Die Größe der Pre-Test-Stichprobe sollte mindestens ein Prozent der Gesamtstichprobe betragen (44).

Der Zweck von Pre-Tests besteht unter anderem in der Ermittlung der durchschnittlichen Befragungszeit, der Prüfung der Verständlichkeit von Fragen, der Prüfung von Fragekontexteffekten und alternativen Frageformulierungen, und letztendlich kann der Pre-Test Hinweise auf die zu erwartende spätere Rücklaufquote geben (45).

Zu Beginn der Studie wurde ein verkürzter Pre-Test durchgeführt. Das bedeutet, dass ich zwar den inhaltlichen Kontext der Studie getestet habe, nicht jedoch den organisatorischen Ablauf. Den Inhalt und die Verständlichkeit des Fragebogens habe ich an 20 Personen unterschiedlichen Alters, Ausbildungstandes und verschiedener Herkunft getestet. Dazu zählten Ärzte der Fachrichtungen Anästhesie und Innere Medizin, Kommilitonen und Nichtmediziner aus dem Freundes- und Familienkreis. Das entspricht etwa 2% der Gesamtstichprobe. Aufgrund der Ergebnisse dieser Vortestung wurde der Fragebogen hinsichtlich der Verständlichkeit und des Zeitaufwandes modifiziert.

2.4.2 Hauptfeld

Das Hauptfeld entspricht üblicherweise der endgültigen Version einer Befragung. Der zeitliche und organisatorische Rahmen wurde zunächst für zwei Wellen festgelegt. In jeder Welle wurden insgesamt 300 Briefen versandt, davon 150 an Ärzte in Schleswig-Holstein und 150 Briefe an Ärzte in Bayern. Es folgte mit etwas Abstand ein Erinnerungsschreiben. Vor Absenden der zweiten Welle wurde die Rücklaufquote der ersten Welle abgewartet, um gegebenenfalls hinsichtlich der Ausfälle Gegenmaßnahmen ergreifen zu können. Somit wurde diese erste Welle als eigentlicher und vollständiger ‚Pre-Test' verstanden, weshalb ich im Folgenden auf die Begrifflichkeiten ‚Pre-Test' und ‚Hauptfeld' verzichte. Aufgrund der Rücklaufergebnisse mussten für die zweite Welle hinsichtlich Organisation und Inhalt keine wesentlichen Änderungen vorgenommen werden.

2.4.3 Erinnerungsschreiben

Um einen möglichst hohen Rücklauf zu erzielen, wurden für beide Wellen Erinnerungsschreiben an diejenigen Ärzte verschickt, die bis zu einem festgesetzten Zeitpunkt noch nicht geantwortet hatten. Für die erste Welle wurde das Erinnerungsschreiben nach drei Wochen (25.08.2004) verschickt, für die zweite Welle nach vier Wochen (20.10.2004). Das Erinnerungsschreiben war personalisiert und eigenhändig unterschrieben (siehe Anlage 9.1.3). Es wurden nochmals der Fragebogen sowie ein portofreier Rückumschlag beigelegt, um die Antwortwahrscheinlichkeit zu erhöhen.

2.5 Methodik

2.5.1 Grundlagen und Hypothesen

Wie bereits erläutert, waren im Rahmen dieser Studie insgesamt 600 Fragebögen an Ärzte in Schleswig-Holstein und Bayern verschickt worden. Ein objektives Maß für die Qualität der Studie ist die Ausschöpfung. Dabei setzt sich die Ausschöpfung, die in diesem Fall der Anzahl der ausgewerteten Fragebögen entspricht, folgendermaßen zusammen (56):

Bruttostichprobe – stichprobenneutrale Ausfälle = bereinigte Stichprobe,

Bereinigte Stichprobe – nicht stichprobenneutrale Ausfälle = Ausschöpfung.

Die Ausschöpfungsquote errechnet sich daraus wie folgt:

Ausgewertete Interviews / Bereinigte Stichprobe x 100 = Ausschöpfungsquote.

Dabei ist zu beachten, dass das Ergebnis der bereinigten Stichprobe gleich 100% gesetzt wird.

Spricht man von den oben genannten Ausfällen, so unterscheiden sich diese in stichprobenneutrale und systematische sowie nicht stichprobenneutrale Ausfälle. Zu den stichprobenneutralen Ausfällen zählen zum Beispiel Fälle, in denen der Brief zwar zurückkommt, aber eine Beantwortung nicht möglich war, weil der Empfänger verzogen ist. Verweigert dagegen ein Arzt die Teilnahme, so spricht man von einem nicht stichprobenneutralen Ausfall (38).

Die Auswertung der beantworteten Fragebögen erfolgte unter verschiedenen Gesichtspunkten. Dabei ging es vor allem um Unterschiede und Gemeinsamkeiten im Gebrauch der Wahrscheinlichkeitsbegriffe unter den Ärzten in Schleswig-Holstein und den Ärzten in Bayern. Zunächst wurden die absoluten Werte der einzelnen Teilnehmer miteinander verglichen. Für die Auswertung der Daten war aber nicht nur der absolute Wert von Bedeutung, sondern vielmehr die Reihenfolge im Gebrauch der Wahrscheinlichkeitsausdrücke. Aus diesem Grund wurden die Zuordnungen der quantitativen Wahrscheinlichkeitswerte zu den verbalen Wahrscheinlichkeitsbegriffen und umgekehrt für jede Person in ordinal skalierte Werte überführt.

Unter der Verwendung der gewonnenen Daten aus der Liste mit den 20 Wahrscheinlichkeitsbegriffen (Abschnitt A) sollten die folgende Fragestellungen und Hypothesen überprüft werden.

1. Es gibt einen Zusammenhang in der prozentualen Bewertung der 20 Wahrscheinlichkeitsbegriffe aus der Liste im Sinne einer Reihenfolge von Rängen (Rangreihenfolge) unter den Ärzten aus Schleswig-Holstein und den Ärzten aus Bayern.

Sollte sich dieser Zusammenhang bezüglich einer Reihenfolge, nach der diese 20 Begriffe nach Rängen eingeordnet werden können, bestätigen, so ist zu überprüfen, ob die beiden aktuellen Rangreihenfolgen der Ärzte aus Schleswig-Holstein und Bayern (echte Rangreihen) mit einer ‚idealen' Rangreihenfolge aus der Literatur übereinstimmen.

2. Die errechneten echten Rangreihenfolgen der Ärzte aus Schleswig-Holstein sowie der Ärzte aus Bayern stimmen mit einer vorgegebenen ‚idealen' Rangreihenfolge bezüglich der 20 Wahrscheinlichkeitsbegriffe aus der Liste überein.

Zusammen mit den Daten aus dem ersten und dritten Sachverhalt (Abschnitt B und D) waren diese Daten aus Abschnitt A Grundlage der dritten und vierten Hypothese, in denen es vor allem um die Frage der Kontextabhängigkeit von Wahrscheinlichkeitsbegriffen geht. Dabei wurden aus Abschnitt B die Begriffe „häufig – gelegentlich – selten" und aus

Abschnitt D die Begriffe „möglicherweise – oft – wahrscheinlich“ mit denselben sechs Begriffen aus der Liste in Abschnitt A verglichen.

3. Es gibt Unterschiede in der prozentualen Bewertung von verbalen Wahrscheinlichkeitsbegriffen insbesondere hinsichtlich des Median der Paardifferenzen, abhängig davon, ob diese Begriffe einzeln in einer Liste aufgeführt werden oder in einen bestimmten Sachverhalt eingebettet sind.

Insbesondere interessierte neben der genannten Fragestellung auch, wie groß die Unterschiede und Gemeinsamkeiten bezüglich der Bewertungen der sechs Begriffe unter den Ärzten in Schleswig-Holstein und Bayern sind.

4. Ärzte in Schleswig-Holstein und Bayern unterscheiden sich bei der Bewertung in der Vergabe von Prozentzahlen bezüglich der sechs vorgegebenen Wahrscheinlichkeitsbegriffe.

Die Daten aus dem zweiten Sachverhalt (Abschnitt C) wurden mit den aktuellen Angaben zu Wahrscheinlichkeitsbegriffen bezüglich unerwünschter Nebenwirkungen von Medikamenten, wie sie in jedem Beipackzettel zu finden sind, verglichen und anhand der Vorgaben aus der EU-Richtlinie deskriptiv ausgewertet.

2.5.2 Vergleich der einzelnen Rangreihenfolgen mit Kendall`s W

Hypothese 1 wurde unter Verwendung des Konkordanz-Koeffizienten von Kendall (Kendall`s W) überprüft.

Die zu prüfende Hypothese lautet:

Es gibt einen Zusammenhang in der prozentualen Bewertung der 20 Wahrscheinlichkeitsbegriffe aus der Liste im Sinne einer Reihenfolge von Rängen (Rangreihenfolge) unter den Ärzten aus Schleswig-Holstein und den Ärzten aus Bayern.

H_0: Die Bewertung der Begriffe stimmt zufällig überein ($W = 0$)

H_1: Die Bewertung der Begriffe ist konkordant ($W > 0$)

Kendall`s W ist ein Maß für den Zusammenhang zwischen mehreren Rangreihen von n Datenträgern (hier speziell n Wahrscheinlichkeitsausdrücken). Bei Vorliegen von m Rangreihen kann mit Hilfe dieses Konkordanzkoeffizienten der Zusammenhang zwischen ihnen festgestellt werden (46, 47).

Kendall`s W lässt sich wie folgt berechnen: $W = \frac{12S}{m^2(n^3-n)}$.

Dabei ist W der Konkordanzkoeffizient, S die Summe der quadrierten Abweichungen, m die Anzahl der Rangreihen, z.B. die Anzahl der befragten Ärzte und n die Anzahl der in einer Rangreihe angeordneten Einheiten.

Mit Hilfe von Kendall`s W kann somit die Stärke eines Zusammenhanges zwischen mehreren Rangreihen überprüfen werden. W erreicht bei perfekter Konkordanz den Wert 1 und bei maximaler Diskordanz den Wert 0. Je größer also die Ähnlichkeit der Rangreihen der befragten Ärzte ist, desto größer wird der Wert von W (47). Negative Werte oder Werte > 1 sind bei der Berechnung des Konkordanzkoeffizienten nicht möglich.

Die Berechnung des Konkordanzkoeffizienten W nach Kendall wurde mit Hilfe von SPSS/Windows (Version 12.8/18) durchgeführt. Zusätzlich zum berechneten Wert von W wurde eine ‚echte' Rangreihe auf Basis der verwendeten Daten ermittelt. Am Ende der Berechnung lagen somit jeweils für die Ärzte aus Schleswig-Holstein als auch für die Ärzte aus Bayern eine eigene ‚echte' Rangreihe sowie ein geschätzter Konkordanzkoeffizient W vor.

Es wurde zunächst für jedes W die Hypothese

H_0: W = 0, das heißt, die Bewertung der Begriffe stimmt zufällig überein versus

H_1: W > 0, das heißt, die Bewertung der Begriffe ist konkordant

mit dem in der Referenz (47) beschriebenen asymptotischen χ^2-Test überprüft. Für diesen und alle weiteren Tests wurde ein Signifikanzniveau von $\alpha = 0{,}05$ festgelegt.

Da die asymptotische Verteilung von Kendall`s W nicht bekannt ist, bedient man sich bestimmter Resamplingmethoden um für Kendall`s W ein approximatives Konfidenzintervall zu schätzen. Eine konservative Abschätzung der Zufälligkeit für das Maß der Übereinstimmung ist mit Hilfe des Bootstrap- oder Jacknife-Verfahrens möglich. Einen Überblick über diese und einige andere Resamplingmethoden findet man zum Beispiel in den Referenzen (48) und (49).

Für die folgende Berechnung wurde das Bootstrapverfahren angewendet, das heißt, die Zufalls-Stichproben wurden mit Zurücklegen gezogen.

Die Syntax für die Bootstrapberechnung ist im Anhang (Anlage 9.1.5) abgebildet. Es wurde separat für die Fragebögen der Ärzte aus Schleswig-Holstein und der Ärzte aus Bayern ein Bootstrapping mit n = 1000 Ziehungen durchgeführt.

Als konservative Abschätzung wurde der mit dem Bootstrap berechnete Konkordanzkoeffizient W verwendet. Als untere Grenze für das Konfidenzintervall wurde das 2,5te Perzentil, und für die obere Grenze das 97,5te Perzentil gewählt.

2.5.3 Vergleich der Rangreihen mit einer idealen Rangreihe

2.5.3.1 Ideale Rangreihe

Als ideale Rangreihe wird die in Tabelle 1 dargestellte Rangreihe der 20 Variablen bezeichnet, die von *Suß* (39) mit Hilfe von verschiedenen statistischen Verfahren im Vergleich mit einer Rangreihe aus der Literatur als die optimale Rangreihe ermittelt wurde.

Tabelle 1: Die ideale Rangreihe der 20 Wahrscheinlichkeitsbegriffe

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

Somit war die ideale Rangreihe nicht zufällig ermittelt worden, sondern die Rangfolge wurde bewusst gewählt. Die hier errechneten echten Rangreihenfolgen für die Ärzte in Schleswig-Holstein und Bayern waren dagegen zufällig ermittelt.

2.5.3.2 Rangkorrelation

Die ideale Rangreihe und die errechneten Rangreihen sollten Grundlage von Hypothese 2 sein:

> Die errechneten echten Rangreihenfolgen der Ärzte aus Schleswig-Holstein sowie der Ärzte aus Bayern stimmen mit einer vorgegebenen ‚idealen' Rangreihenfolge bezüglich der 20 Wahrscheinlichkeitsbegriffe aus der Liste überein.
>
> H_0: Es besteht kein Zusammenhang zwischen der idealen Rangreihe und jeweils der echten Rangreihe der Ärzte aus Schleswig-Holstein und Bayern ($r_s = 0$) versus
>
> H_1: Es besteht ein Zusammenhang zwischen der idealen Rangreihe und jeweils der echten Rangreihe der Ärzte aus Schleswig-Holstein und Bayern ($r_s \neq 0$).

Für die Überprüfung der Übereinstimmung zweier Rangreihen kann die Rangkorrelation nach Spearmans r_s (rho) oder Kendall`s τ_b (tau) berechnet werden. Vorraussetzung ist, dass die Variablen mindestens auf Ordinalniveau gemessen werden können.

r_s basiert auf den zugeordneten Rängen, nicht auf den direkten Merkmalsausprägungen. Die einzelnen Rangdifferenzen (d_i) dienen als Indikator für die Verschiedenheit der beiden Rangreihen. Bei vollkommener Übereinstimmung wäre jedes $d_i = 0$. Je größer die d_i sind, desto geringer ist der Zusammenhang zwischen beiden Variablen. Letztendlich ist der Rangkorrelationskoeffizient nach Spearman nichts anderes als der auf diese Rangnummern angewandte Pearson-Korrelationskoeffizient.

Es gilt: $-1 \leq r_s \leq +1$; das bedeutet, bei völliger Übereinstimmung wird der Extremwert +1 und bei völliger Gegensinnigkeit der Wert -1 eingenommen (50).

Auf r_s basierende statistische Tests finden sich in Standard-Lehrbüchern der Statistik, zum Beispiel in (51). Die Berechnung von r_s erfolgte mit Hilfe von SPSS.

Eine weitere Möglichkeit zur Ermittlung der Stärke eines Zusammenhanges zwischen zwei Rangreihen ist die Berechnung des Rangkorrelationskoeffizienten nach Kendall, genauer Kendall`s tau τ_b (47). Dabei wird die erste Rangreihe natürlich angeordnet (Ankerreihe) und die Vergleichsreihe hinsichtlich der Proversionen (+) bzw. Inversionen (-) geprüft. Somit basiert die Überprüfung auf einer Fehlordnung der Ränge innerhalb der Vergleichsreihe (52). Der Grad des Zusammenhanges ist daher das Verhältnis der tatsächlich berechneten Summe der Pro- und Inversionen zur maximal möglichen Anzahl.

Als Formel ausgedrückt (46) ist: $\tau_b = \frac{S}{1/2\ n(n-1)}$.

Dabei bezeichnet S die Summe der Pro- und Inversionen (+ und – Zeichen) und n die Anzahl der Objekte (hier: n = 20).

Auch für Kendall`s τ_b gilt: $-1 \leq \tau_b \leq +1$. Sofern $\tau_b = 0$ ergibt, heißt das jedoch nicht, dass kein Zusammenhang besteht, sondern nur, dass eine monotone Komponente fehlt.

Zum statistischen Testen können z.B. die Tabellen aus *Siegel* (46) verwendet werden. Unter H_0 sind alle n!-Zuordnungen der Vergleichsreihe zur Ankerreihe somit gleich wahrscheinlich, das heißt es gibt keine Korrelation. Die Berechnung von Kendall`s tau erfolgte mit Hilfe von SPSS.

Zu beachten ist, dass beide Rangkorrelationskoeffizienten nicht numerisch direkt miteinander verglichen werden können, da die Korrelation auf unterschiedliche Weise

gemessen wird. Kendall`s tau wurde hier zum Vergleich der ermittelten Rangreihen mit der idealen Rangreihe lediglich als deskriptives Maß verwendet.

2.5.3.3 Permutationstest nach Diaconis

Mit Hilfe der Rangkoeffizienten kann die Stärke der Zusammenhänge zwischen den Rangreihen berechnet werden. Zur Signifikanzprüfung hinsichtlich der Nullhypothese, ob die Begriffe tatsächlich nur zufällig übereinstimmen, wurde der Permutationstest nach Diaconis zur Hilfe genommen (53). Der Vorteil besteht darin, dass nur eine Reihe zufällig beobachtet werden muss, wohingegen die andere Rangreihe fest vorgegeben sein kann. Dieser Test ist nichtparametrisch, und es wird vorausgesetzt, dass eine paarige Beobachtung für jedes n vorgenommen wird; für Details siehe Referenz (53).

Zur Berechnung dient die Spearmansche Faustregel $D = \sum_{i=1}^{n} \left| R(X_i) - R(Y_i) \right|$, wobei $R(X_i)$ den Rang in der Reihe X des Begriffes i bezeichnet und $R(Y_i)$ den Rang in der Reihe Y des Begriffes i.

Für den Mittelwert (AV), die Standartabweichung (SD) und der Varianz (VAR) unter der Nullhypothese der Spearmanschen Faustregel gilt (53):

$$AV(D) = \frac{1}{3}(n^2-1),\ SD(D) = \sqrt{VAR(D)},\ VAR(D) = \frac{1}{45}(n+1)(2n^2+7)\ .$$

Zum Testen wird die standardisierte Teststatistik $T = \frac{D - AV}{SD}$ verwendet.

In der vorliegenden Arbeit war n = 20, das heißt es wurde die Anordnung der 20 Wahrscheinlichkeitsbegriffe in den verschiedenen Rangreihen miteinander verglichen.

2.5.4 Vergleich der Angaben zu Begriffen in der Liste und im Kontext

2.5.4.1 Wilcoxon-Test

Getestet wird die Hypothese, ob sich die vergebenen Prozentzahlen für die Begriffe

„häufig – gelegentlich – selten“

sowie

„möglicherweise – oft – wahrscheinlich“

aus der Liste in Abschnitt A signifikant von den vergebenen Prozentzahlen in den Fallbeispielen aus Abschnitt B („häufig – gelegentlich – selten“) oder Abschnitt D („möglicherweise – oft – wahrscheinlich“) unterscheiden. Die Hypothesen wurden bezüglich des Median der Paardifferenzen formuliert.

Die hierbei zu prüfende Hypothese 3 lautet:

Es gibt Unterschiede in der prozentualen Bewertung von verbalen Wahrscheinlichkeitsbegriffen insbesondere hinsichtlich des Median der Paardifferenzen, abhängig davon, ob diese Begriffe einzeln in einer Liste aufgeführt werden oder in einen bestimmten Sachverhalt eingebettet sind.

H_0: Es gibt keine Unterschiede in der Bewertung der Mediandifferenz der Begriffe

- $A_{\text{häufig}} = B_{\text{häufig}}$ und $A_{\text{möglicherweise}} = D_{\text{möglicherweise}}$
- $A_{\text{gelegentlich}} = B_{\text{gelegentlich}}$ und $A_{\text{oft}} = D_{\text{oft}}$
- $A_{\text{selten}} = B_{\text{selten}}$ und $A_{\text{wahrscheinlich}} = D_{\text{wahrscheinlich}}$

H_1: Es gibt Unterschiede in der Bewertung der Mediandifferenz der Begriffe

- $A_{\text{häufig}} \neq B_{\text{häufig}}$ und $A_{\text{möglicherweise}} \neq D_{\text{möglicherweise}}$
- $A_{\text{gelegentlich}} \neq B_{\text{gelegentlich}}$ und $A_{\text{oft}} \neq D_{\text{oft}}$
- $A_{\text{selten}} \neq B_{\text{selten}}$ und $A_{\text{wahrscheinlich}} \neq D_{\text{wahrscheinlich}}$

Die Nullhypothese kann zugunsten der Alternativhypothese verworfen werden, wenn der berechnete p-Wert kleiner ist als das Signifikanzniveau von $\alpha = 0{,}05$.

Diese Hypothese wird mit Hilfe des Wilcoxon-Rangsummen-Tests überprüft. Mit dem Wilcoxon-Rangsummen-Test kann sowohl hinsichtlich der Richtung als auch der Größe der Unterschiede innerhalb der Paare eine Aussage gemacht werden (46).

Unter der Nullhypothese sind die Differenzen der paarig angeordneten Werte mit dem Median gleich Null. Sie wird verworfen, wenn entweder die Rangsumme der negativen oder die der positiven Differenzen deutlich unterschiedlich ist. Die Berechnung des Wilcoxon-Rangsummen-Tests wurde mit Hilfe von SPSS/Windows (Version 12.8/18) durchgeführt.

Es wurde außerdem der Hodges-Lehmann-Punktschätzer für den Median der Mediandifferenzen und das dazugehörige 95%-Konfidenzintervall mit Hilfe von SAS/StatXact 8 PROCs berechnet.

2.5.4.2 Mann-Whitney-Test

Zur Prüfung der Hypothese, ob sich die Ärzte aus Schleswig-Holstein gegenüber den Ärzten aus Bayern bezogen auf die Vergabe der Prozentzahlen für die Wahrscheinlichkeitsbegriffe „häufig – gelegentlich – selten“ aus der Liste in Abschnitt A und dem Fallbeispiel in Abschnitt B unterscheiden, wurde der Mann-Whitney-Test verwendet. Der Test wurde auch zum Vergleich der Bundesländer-spezifischen Prozentzahlen für die Wahrscheinlichkeitsbegriffe „möglicherweise – oft – wahrscheinlich“ aus Abschnitt A sowie dem Fallbeispiel in Abschnitt D eingesetzt.

Diese Fragestellungen führten zur Hypothese 4:

Ärzte in Schleswig-Holstein und Bayern unterscheiden sich bei der Bewertung in der Vergabe von Prozentzahlen bezüglich der sechs vorgegebenen Wahrscheinlichkeitsbegriffe.

H_0: Es gibt keine Unterschiede in der prozentualen Bewertung der Begriffe durch die Ärzte in Schleswig-Holstein und Bayern

- $A_{\text{häufig}} = B_{\text{häufig}}$ und $A_{\text{möglicherweise}} = D_{\text{möglicherweise}}$
- $A_{\text{gelegentlich}} = B_{\text{gelegentlich}}$ und $A_{\text{oft}} = D_{\text{oft}}$
- $A_{\text{selten}} = B_{\text{selten}}$ und $A_{\text{wahrscheinlich}} = D_{\text{wahrscheinlich}}$

H_1: Es gibt Unterschiede in der Bewertung der Begriffe durch die Ärzte in Schleswig-Holstein und Bayern

- $A_{\text{häufig}} \neq B_{\text{häufig}}$ und $A_{\text{möglicherweise}} \neq D_{\text{möglicherweise}}$
- $A_{\text{gelegentlich}} \neq B_{\text{gelegentlich}}$ und $A_{\text{oft}} \neq D_{\text{oft}}$
- $A_{\text{selten}} \neq B_{\text{selten}}$ und $A_{\text{wahrscheinlich}} \neq D_{\text{wahrscheinlich}}$

Als Signifikanzniveau diente $\alpha = 0{,}05$.

Im Unterschied zum Wilcoxon-Rangsummen-Test werden beim Mann-Whitney-Test zwei unabhängige Gruppen miteinander verglichen (54).

Der Mann-Whitney-Test wurde mit Hilfe von SPSS durchgeführt, so dass über die Option „Exakt“ genaue Überschreitungswahrscheinlichkeiten vorgelegt werden konnten.

Des Weiteren folgte die Berechnung des Mann-Whitney-Test für die Differenzvariable, das heißt für die Variablen „häufig“ (ohne Kontext) aus Abschnitt A und „häufig“ (Kontext) aus Abschnitt B, den beiden Variablen „gelegentlich“ (ohne Kontext) und „gelegentlich“ (Kontext) sowie den beiden abhängigen Variablen „selten“ (ohne Kontext) und „selten“ (Kontext) wurden die Differenzen gebildet. Die Berechnungen wurden mit SAS/StatXact durchgeführt.

Das gleiche Vorgehen wurde für die drei Variablen aus Abschnitt D „möglicherweise – oft – wahrscheinlich“ verwendet.

2.5.5 Auswertung der Angaben zu den vorgegebenen Wahrscheinlichkeitsquoten

2.5.5.1 Auswertung der Begriffe der Ärzte aus Schleswig-Holstein und Bayern

Im zweiten Abschnitt des Fragebogens waren für den zweiten Sachverhalt (Abschnitt C) fünf quantitative Wahrscheinlichkeitsausdrücke vorgegeben worden. Diese Ausdrücke wurden in Anlehnung an die EU-Richtlinie Summary of Product Characteristics (SPC-Richtlinie) zu „Häufigkeitsangaben von Nebenwirkungen“ von 1999 gewählt (25). Das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) hatte 2002 diese Richtlinie

erstmals umgesetzt und die Produktinformation für das Auftreten von Nebenwirkungen von Arzneimitteln an den Stand der wissenschaftlichen Erkenntnis angepasst und genauer umschrieben (2). Seitdem erfolgt die Angabe des Auftretens von Nebenwirkungen in Form von Zahlenangaben in Prozent und in Form des entsprechenden Häufigkeitsbegriffes. Nochmals aufgearbeitet und verbessert wurden diese Angaben 2007 und 2009. Tabelle 2 gibt einen Überblick über Häufigkeitsangaben im Sinne der EU-Richtlinie (SPC-Guideline) (25).

Tabelle 2: SPC-Richtlinie „Häufigkeitsangaben von Nebenwirkungen“

SPC-Guideline	Deutsch	Prozent-Angaben
very common	sehr häufig	≥ 10
common	häufig	≥1 – < 10
uncommon	gelegentlich	≥0,1 – < 1
rare	selten	≥0,01 – < 0,1
very rare	sehr selten	< 0,01

(SPC = Summary of Product Characteristics)

Zunächst wurden die eigenen verbalen Häufigkeitsangaben der Ärzte aus Schleswig-Holstein und Bayern direkt miteinander verglichen.

2.5.5.2 Vergleich der ausgewerteten Begriffe mit der Richtlinie

In diesem Abschnitt wurden die von den Teilnehmern der Studie vergebenen numerischen Prozentwerte mit denen verglichen, die vom Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) in der Nebenwirkungstabelle (2) vorgegeben werden. Die interessierende Frage war, ob die von den Teilnehmern der Befragung genannten numerischen Wahrscheinlichkeiten für die verbalen Wahrscheinlichkeitsbegriffe „häufig“, „gelegentlich“ und „selten“, den numerischen Wahrscheinlichkeiten für diese verbalen Wahrscheinlichkeiten nach der offiziellen Klassifikation für Nebenwirkungen entsprachen.

Kruse und *Sandner* geben die Häufigkeiten von Nebenwirkungen nach der offiziellen Nebenwirkungstabelle für Nebenwirkungen bei pharmazeutischen Produkten in Tabelle 3 wie folgt an (55):

Tabelle 3: Klassifikation der Nebenwirkungen

Wahrscheinlichkeit	Definition
sehr häufig	Bei mehr als 10% der Patienten tritt die UAW auf.
häufig	Bei 1-10% der Patienten tritt die Nebenwirkung auf.
gelegentlich	Die Wahrscheinlichkeit für die UAW liegt zwischen 0,1% und 1%.
selten	Die Wahrscheinlichkeit liege bei 0,01-0,1% (ein bis zehn Patienten von 10.000)
sehr selten	Die Wahrscheinlichkeit liegt unter 0,01%, UAW treten also bei weniger als einem Patienten pro 10.000 Patienten auf.

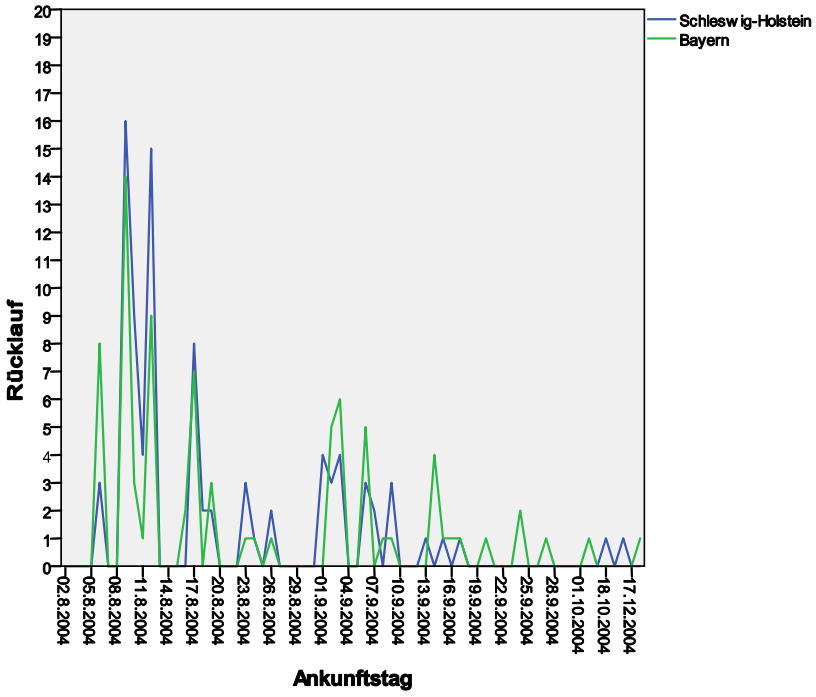
3. Ergebnisse

3.1 Ergebnisse zum Fragebogenrücklauf

Im Rahmen dieser Studie waren 600 Fragebögen an Ärzte in Schleswig-Holstein und Bayern verschickt worden. Nach erfolgtem Pre-Test gab es zwei Wellen mit jeweils 300 Fragebögen, sowie nachfolgend ein Erinnerungsschreiben mit beiliegendem Fragebogen. Bei der Studie konnten insgesamt sechs Briefe nicht zugestellt werden und fünf Ärzte verweigerten ihre Teilnahme.

Die erste Welle mit 300 Fragebögen wurde am 02.08.2004 gestartet. Es wurden jeweils 150 Fragebögen an Ärzte in Schleswig-Holstein und Bayern verschickt. Am 25.08.2004 wurde derselbe Fragebogen mit einem Erinnerungsschreiben an diejenigen Ärzte geschickt, die bis zu diesem Zeitpunkt noch nicht geantwortet hatten. Insgesamt wurden 175 Fragebögen von den angeschriebenen Ärzten zurückgeschickt. Eine Übersicht über den Rücklauf zeigt Abbildung 1.

Abbildung 1: Rücklauf 1. Welle – Schleswig-Holstein und Bayern



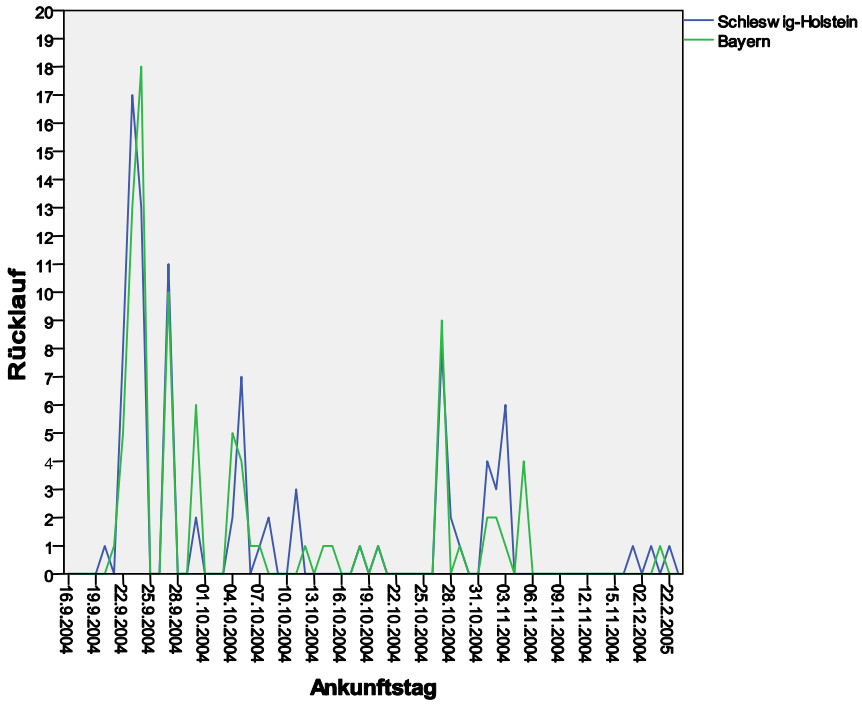
Die Ausschöpfung belief sich auf 170 Fragebögen, entsprechend einer Quote von 57,63%. Eine kurze Übersicht bietet Tabelle 4.

Tabelle 4: Übersicht Gesamtausschöpfung 1. Welle

	Anzahl	Prozent
Gesamtstichprobe	**300**	**100,00**
Stichprobenneutrale Ausfälle	5	1,70
Bereinigte Stichprobe	**295**	**100,00**
Verweigerungen	0	0,00
Berentung	0	0,00
Realisierte Fragebögen		
nach Erstversand	121	41,00
nach Erinnerung	49	16,60
Gesamtausschöpfung	**170**	**57,60**

Der detaillierte Rücklauf sowie die Berechnung für die einzelnen Untergruppen der Ärzte sind im Anhang (Anlage 9.1.6) abgebildet.

Abbildung 2: Rücklauf 2. Welle – Schleswig-Holstein und Bayern



Die zweite Welle wurde am 16.09.2004, sechs Wochen nach der ersten Welle, gestartet. Es wurden erneut 300 Fragebögen an Ärzte in Schleswig-Holstein und Bayern verschickt, jeweils 150 Fragebögen pro Bundesland. Diesmal wurde das Erinnerungsschreiben mit

dem Fragebogen circa vier Wochen später, am 20.10.2004, an die Ärzte verschickt, deren Fragebogen noch nicht zurückgekommen war. Eine Übersicht über den Rücklauf der zweiten Welle zeigt Abbildung 2.

Tabelle 5: Übersicht Gesamtausschöpfung 2. Welle

	Anzahl	Prozent
Gesamtstichprobe	**300**	**100,00**
Stichprobenneutrale Ausfälle	1	0,30
Bereinigte Stichprobe	**299**	**100,00**
Verweigerungen	3	1,00
Berentung	2	0,70
Realisierte Fragebögen		
nach Erstversand	135	45,20
nach Erinnerung	45	15,00
Gesamtausschöpfung	**180**	**61,20**

In dieser zweiten Welle kamen insgesamt 186 Fragebögen zurück. Bezogen auf die Bruttozahl entspricht das 62,00 %. Die Ausschöpfung der zweiten Welle betrug 180 Fragebögen. Das entspricht einer Ausschöpfungsquote von 61,20%. Veranschaulicht sind der Rücklauf und die Ausschöpfung in Tabelle 5.

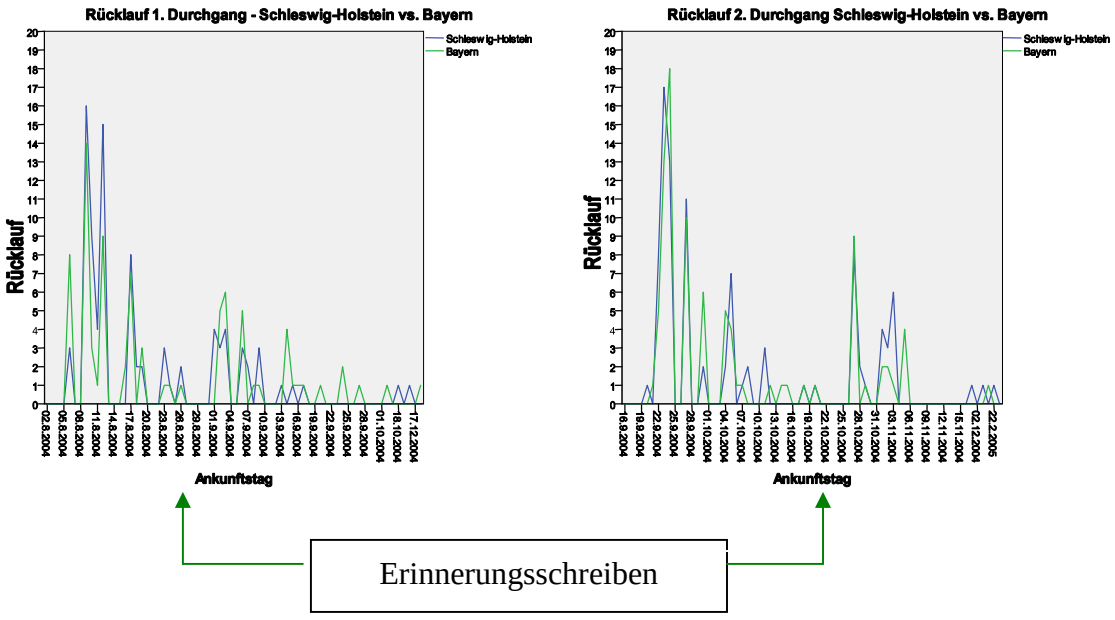
Tabelle 6: Übersicht Gesamtausschöpfung der beiden Wellen

	Anzahl	Prozent
Gesamtstichprobe	**600**	**100,00**
Stichprobenneutrale Ausfälle	6	1,00
Bereinigte Stichprobe	**594**	**100,00**
Verweigerungen	3	0,05
Berentung	2	0,03
Realisierte Fragebögen		
nach Erstversand	256	43,10
nach Erinnerung	94	15,80
Gesamtausschöpfung	**350**	**58,90**

Insgesamt waren an 600 Ärzte in Schleswig-Holstein und Bayern der Fragebogen verschickt worden. Die Zahl der zurückgeschickten Fragebögen betrug im Brutto insgesamt 361 (60,17%). Die eigentliche Ausschöpfung lag bei 350 Fragebögen mit einer Ausschöpfungsquote von 58,92%. Eine kurze Übersicht bietet Tabelle 6.

Die Rücklaufdauer für die Auswertung war auf 50 Werktage festgesetzt. Einzelne Fragebögen, die nach diesem Tag eingingen, wurden so behandelt, als ob sie innerhalb dieser Zeitspanne eingegangen wären. Die mittlere Rücklaufdauer der versandten Fragebögen für die erste Welle betrug 14 Tage, für die zweite Welle 10 Tage. Für die erste Welle wurde ein Erinnerungsschreiben nach circa drei Wochen versandt. Bis zu diesem Zeitpunkt hatten bereits 40,33% der Befragten geantwortet. Für die zweite Welle erfolgte das Erinnerungsschreiben nach circa vier Wochen. Bis dahin waren 46,67% der versandten Fragebögen wieder eingegangen. Erwartungsgemäß stieg die Rücklaufquote nach dieser Erinnerung nochmals an (Abbildung 3).

Abbildung 3: Gegenüberstellung der Rückläufe



Der detaillierte Rücklauf sowie die Berechnung für die einzelnen Untergruppen der Ärzte sind im Anhang zu finden (Anlage 9.1.6).

3.2 Soziodemographische Daten

Von den 350 Studienteilnehmern stammen 168 Probanden aus Bayern und 180 Probanden aus Schleswig-Holstein. Zwei Teilnehmer haben keine Angaben zu ihrem aktuellen

Wohnsitz gemacht und wurden deshalb in der Studie nicht weiter berücksichtigt, da sie die für die Zielstellung wichtige Frage nach dem aktuellen Wohnsitz (Bayern oder Schleswig-Holstein) nicht beantwortet haben.

In Tabelle 7 ist die medizinische Fachrichtung der Ärztinnen und Ärzte unterteilt nach Bundesland dargestellt. Von den 180 aus Schleswig-Holstein stammenden Ärzten gehören 109 der medizinischen Fachrichtung „Innere Medizin“ und 71 Ärzte der Fachrichtung „Anästhesie“ an.

Tabelle 7: Medizinische Fachrichtung der Teilnehmer unterteilt nach Bundesland

	Fachrichtung	
Bundesland	Innere Medizin	Anästhesie
Bayern	81 (43%)	86 (55%)
Schleswig-Holstein	109 (57%)	71 (45%)
Gesamt	190 (100%)	157 (100%)

Die bayerischen Mediziner unterteilen sich in 81 Personen der Fachrichtung „Innere Medizin“ und 86 Anästhesisten.

Von den 348 Probanden haben 345 Personen Angaben zu ihrem Geschlecht gemacht. 294 Studienteilnehmer sind männlich (85.2%) und 51 Teilnehmerinnen sind weiblich (14.8%).

Tabelle 8 stellt das Geschlecht der an Studie teilnehmenden Ärztinnen und Ärzte getrennt nach Bundesland dar. Von den 168 in Bayern tätigen Medizinern sind 147 Ärzte männlich (88.6%) und 19 Ärzte weiblich (11.4%).

Tabelle 8: Geschlecht der Befragten unterteilt nach Bundesland

	Geschlecht		
Bundesland	Weiblich	Männlich	Gesamt
Bayern	19 (11.4%)	147 (88.6%)	166 (100%)
Schleswig-Holstein	32 (17.9%)	147 (82.1%)	179 (100%)

Von den 180 in Schleswig-Holstein tätigen Ärzten sind ebenfalls 147 Teilnehmer männlich (82.1%) und 32 weiblich (17.9%).

Tabelle 9: Geschlecht der Befragten unterteilt nach Fachrichtung

	Geschlecht		
Fachrichtung	Weiblich	Männlich	Gesamt
Innere Medizin	17 (8.9%)	173 (91.1%)	190 (100%)
Anästhesie	34 (21.9%)	121 (78.1%)	155 (100%)

Von den 190 teilnehmenden Internisten an der Befragung sind 17 Personen weiblich (8.9%) und 173 Personen männlich (91.1%). Die 155 Internisten unterteilen sich in 121 Anästhesisten (78.1%) und 34 Anästhesistinnen (21.9%) (Tabelle 9).
343 Studienteilnehmer haben Angaben zu ihrem Alter gemacht. Der jüngste Proband ist 33 Jahre und der älteste 86 Jahre alt. Im Mittel sind die Teilnehmer 50 Jahre alt (unteres Quartil 46 Jahren, oberes Quartil 57 Jahre).
Im Mittel sind die Ärzte aus Schleswig mit 50 Jahren vier Jahre jünger als ihre Kollegen aus Bayern mit im Mittel 54 Jahren (Tabelle 10).

Tabelle 10: Alter unterteilt nach Bundesland

	n = Anzahl der Studienteilnehmer	Mittelwert	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
Bayern	165	54	54	36	86	49	59
Schleswig-Holstein	176	50	50	33	69	45	56

3.3 Ergebnisse zum Rangreihenvergleich – Hypothese 1

In diesem Abschnitt wird Hypothese 1 untersucht:

> Es gibt einen Zusammenhang in der prozentualen Bewertung der 20 Wahrscheinlichkeitsbegriffe aus der Liste im Sinne einer Reihenfolge von Rängen (Rangreihenfolge) unter den Ärzten aus Schleswig-Holstein und den Ärzten aus Bayern.
>
> H_0: Die Bewertung der Begriffe stimmt zufällig überein ($W = 0$)
>
> H_1: Die Bewertung der Begriffe ist konkordant ($W > 0$)

Zunächst wurde für die Auswertung der 20 Wahrscheinlichkeitsbegriffe aus der Liste für jeden Begriff der Mittelwert gebildet und im Vergleich die Ergebnisse der Ärzte aus Schleswig-Holstein und Bayern einander gegenübergestellt. Zur Verdeutlichung der Spannbreite der prozentualen Angaben wurden zusätzlich Minimum und Maximum der Werte angegeben. Eine kurze Übersicht bietet Tabelle 11, eine ausführliche tabellarische und graphische Auswertung in Form von Histogrammen ist im Anhang (Anlage 9.1.7) zu finden.
Beim Vergleich der absoluten Häufigkeiten fällt auf, dass die Ärzte in Bayern die verbalen Wahrscheinlichkeitsbegriffe etwas niedriger hinsichtlich des Mittelwertes bewerten. Lediglich die Begriffe „unklar“, „fraglos“, „stets“, „nie“, „immer“ und „fraglich“ fallen in

der Bewertung höher aus. Insgesamt gesehen gibt es aber eine große Übereinstimmung im Trend der Werte.

Tabelle 11: Ausgewählte Lage- und Streuungsmaße für die Ärzte in Schleswig-Holstein und Bayern

	n = Anzahl der Befragten		Mittelwert		Minimum		Maximum	
	Ärzte SH	Ärzte BA	Ärzte SH	Ärzte BA	Ärzte SH	Ärzte BA	Ärzte SH	Ärzte BA
selten	177	166	10.30	10.03	0.05	0.001	100.00	80.00
niemals	176	164	7.16	6.67	0.00	0.00	100.00	100.00
sicher	173	167	92.65	92.47	0.00	5.00	100.00	100.00
unsicher	166	163	39.74	35.99	0.00	0.00	100.00	100.00
oft	176	167	67.57	65.62	10.00	1.00	100.00	95.00
möglicherweise	170	163	34.34	31.73	0.55	0.10	100.00	90.00
gelegentlich	176	166	19.98	17.46	1.00	0.10	90.00	90.00
überwiegend	175	166	74.37	72.65	20.00	5.00	100.00	95.00
meist	176	167	81.29	80.37	20.00	10.00	99.00	100.00
denkbar	171	163	24.07	22.02	0.00	0.005	90.00	80.00
häufig	176	167	70.22	65.32	5.00	1.00	100.00	95.00
unklar	147	136	33.90	35.37	0.00	0.00	100.00	100.00
normalerweise	176	163	82.29	82.03	10.00	10.00	100.00	100.00
fraglos	160	155	81.39	85.50	0.00	0.00	100.00	100.00
stets	176	167	89.81	90.59	0.00	0.00	100.00	100.00
nie	176	164	10.87	13.34	0.00	0.00	100.00	100.00
wahrscheinlich	177	166	75.20	72.86	5.00	5.00	98.00	99.00
immer	176	167	93.41	93.76	0.00	0.00	100.00	100.00
fraglich	157	154	29.04	31.26	0.00	0.00	80.00	100.00
typischerweise	177	162	78.49	76.86	0.00	0.00	100.00	100.00

Tabelle 11 zeigt außerdem sehr deutlich die Unsicherheit bei der Bewertung der 20 Wahrscheinlichkeitsbegriffe vor allem im Hinblick auf Minimum und Maximum der Werte. Ein Studienteilnehmer bewertet den Begriff „selten“ mit 100%, ein anderer mit 0,05%. Es fällt auf, dass diese Werte nicht nur bei einzelnen Teilnehmern variieren, sondern durchweg bei Ärzten in Schleswig-Holstein und Bayern.

Ob es tatsächlich eine Übereinstimmung bei der Bewertung der unterschiedlichen Begriffe unter den Ärzten gibt, sollte mit Hilfe von Kendall`s W ermittelt werden.

Bei der Berechnung des Konkordanz-Koeffizenten Kendall`s W mit Hilfe von SPSS wird zusätzlich auf Basis der verwendeten Daten eine echte Rangreihe für den jeweiligen Datenpool ermittelt. Die sich ergebenden echten Rangreihen für die Ärzte aus Schleswig-Holstein und Bayern sind in Tabelle 12 dargestellt.

Tabelle 12: Echte Rangreihen nach Bundesland

Echte Rangreihe Ärzte Schleswig-Holstein		Echte Rangreihe Ärzte Bayern	
Rang	**Begriff**	**Rang**	**Begriff**
1	niemals	1	niemals
2	nie	2	selten
3	selten	3	nie
4	gelegentlich	4	gelegentlich
5	denkbar	5	denkbar
6	fraglich	6	fraglich
7	unklar	7	unklar
8	möglicherweise	8	möglicherweise
9	unsicher	9	unsicher
10	oft	10	oft
11	häufig	11	häufig
12	überwiegend	12	überwiegend
13	wahrscheinlich	13	wahrscheinlich
14	typischerweise	14	typischerweise
15	meist	15	meist
16	normalerweise	16	normalerweise
17	fraglos	17	fraglos
18	stets	18	stets
19	sicher	19	sicher
20	immer	20	immer

Zum Vergleich der Rangreihenfolgen bezüglich ihrer Konkordanz werden die Berechnung von Kendall`s W, der χ^2- Test zur Prüfung der Signifikanz und zur Ermittlung des Konfidenzintervalles das Bootstrapverfahren durchgeführt (Tabelle 13).

Tabelle 13: Ergebnisse aus dem Vergleich der Rangreihenfolgen

	Kendall`s W	**Kendall`s** W_b	χ^2	**95% KI**
Ärzte SH	0,681**	0,679**	1772,23	(0,593; 0,746)
Ärzte BA	0,669**	0,671**	1550,49	(0,672; 0,743)

** bedeutet: $p \leq 0{,}005$; SH = Schleswig-Holstein, BA = Bayern, W_b = Kendall`s W nach Bootstrap, χ^2 = Chi-Quadrat, KI = Konfidenzintervall,

Die Konkordanz der Angaben ist in beiden Ländern groß, die Bewertung der Begriffe stimmt somit nicht nur zufällig überein.

3.4 Ergebnisse im Vergleich mit einer idealen Rangreihe – Hypothese 2

Im folgenden Abschnitt soll Hypothese 2 überprüft werden:

> Die errechneten echten Rangreihenfolgen der Ärzte aus Schleswig-Holstein sowie der Ärzte aus Bayern stimmen mit einer vorgegebenen ‚idealen' Rangreihenfolge bezüglich der 20 Wahrscheinlichkeitsbegriffe aus der Liste überein.
>
> H_0: Es besteht kein Zusammenhang zwischen der idealen Rangreihe und jeweils der echten Rangreihe der Ärzte aus Schleswig-Holstein und Bayern $(r_s = 0)$ versus
>
> H_1: Es besteht ein Zusammenhang zwischen der idealen Rangreihe und jeweils der echten Rangreihe der Ärzte aus Schleswig-Holstein und Bayern $(r_s \neq 0)$.

Die Ergebnisse der idealen und der echten (beobachteten) Rangreihen der Ärzte aus Schleswig-Holstein und Bayern sind in Tabelle 14 gegenübergestellt.

Beim Vergleich der echten Rangreihe der Ärzte aus Schleswig-Holstein und Bayern mit der idealen Rangreihe fallen einige Unterschiede auf. Die Positionen 2 (= selten und nie), 3 (= nie und selten), 5 (= unsicher und denkbar), 6 (= denkbar und fraglich), 7 (= fraglich und unklar), 9 (= unklar und unsicher), 11 (= wahrscheinlich und häufig), 12 (= häufig und überwiegend), 13 (= überwiegend und wahrscheinlich), 14 (= meist und typischerweise), 15 (= normalerweise und meist) und 16 (= typischerweise und normalerweise) unterscheiden sich zwischen den Rangreihen. Bei den Begriffen „selten", „nie", „denkbar", „fraglich", „häufig", „überwiegend", „meist" und „normalerweise" beträgt der Unterschied in der Rangreihenfolge jeweils nur eine Position nach oben beziehungsweise unten. Die Begriffe „unklar", „wahrscheinlich" und „typischerweise" unterscheiden sich um jeweils zwei Positionen. Der Begriff „unklar" liegt an Position 9 in der idealen Rangreihe und an Position 7 in den echten Rangreihen, der Begriff „wahrscheinlich" liegt an Position 11 beziehungsweise 13 und der Begriff „typischerweise" an Position 16 und 14. Lediglich der Begriff „unsicher" unterscheidet sich um vier Positionen. In der idealen Rangreihe liegt der Begriff an Position 5, während er bei den echten Rangreihenfolgen an Position 9 liegt. Die übrigen acht Begriffe stimmen vollständig überein. Vergleicht man dagegen die echten Rangreihenfolgen untereinander, so gibt es eine deutlich größere Übereinstimmung als im direkten Vergleich mit der idealen Rangreihe. Lediglich die Begriffe „nie" und „selten" variieren um eine Position.

Tabelle 14: Gegenüberstellung der idealen Rangreihe und der beobachteten Rangreihen der Ärzte Schleswig-Holsteins (SH) und Bayerns (BA)

Ideale Rangreihe		Echte Rangreihe Ärzte SH		Echte Rangreihe Ärzte BA	
Rang	**Begriff**	**Rang**	**Begriff**	**Rang**	**Begriff**
1	niemals	1	niemals	1	niemals
2	selten	2	nie	2	selten
3	nie	3	selten	3	nie
4	gelegentlich	4	gelegentlich	4	gelegentlich
5	unsicher	5	denkbar	5	denkbar
6	denkbar	6	fraglich	6	fraglich
7	fraglich	7	unklar	7	unklar
8	möglicherweise	8	möglicherweise	8	möglicherweise
9	unklar	9	unsicher	9	unsicher
10	oft	10	oft	10	oft
11	wahrscheinlich	11	häufig	11	häufig
12	häufig	12	überwiegend	12	überwiegend
13	überwiegend	13	wahrscheinlich	13	wahrscheinlich
14	meist	14	typischerweise	14	typischerweise
15	normalerweise	15	meist	15	meist
16	typischerweise	16	normalerweise	16	normalerweise
17	fraglos	17	fraglos	17	fraglos
18	stets	18	stets	18	stets
19	sicher	19	sicher	19	sicher
20	immer	20	immer	20	immer

Während die Ärzte aus Schleswig-Holstein den Begriff „nie“ an Position 2 gewählt haben und den Begriff „selten“ an Position 3, ist es bei den Ärzten aus Bayern genau anders herum. Damit stimmen die Ärzte aus Bayern in diesem Begriff mit der idealen Rangreihe überein. Insgesamt stimmen die Ärzte in Schleswig-Holstein und Bayern untereinander somit in neunzehn Begriffen überein.

Die Stärke der Korrelation zwischen den einzelnen Rangreihen wurde mit Spearman`s $r_s$ untersucht. Für die ideale Rangreihe im Vergleich mit der echten Rangreihe ergibt sich bei beiden Bundesländern eine große Korrelation. Auch die Untersuchung der Korrelation mit Kendall`s τ_b ergibt hohe Werte für beide Bundesländer. Die Werte sind in Tabelle 15 dargestellt.

Tabelle 15: Rangkorrelation nach Bundesland

Ärzte nach Bundesland	Spearmans r_s	Kendall`s τ_b
Schleswig-Holstein	0,973**	0,821**
Bayern	0,974**	0,832**

** bedeutet: $p \leq 0{,}01$

Die Korrelation der beobachteten Rangreihen der Ärzte aus Schleswig-Holstein und Bayern im Vergleich mit der idealen Rangreihe ist in beiden Ländern groß, die Nullhypothese wird zurückgewiesen. Der Permutationstest von Diaconis bestätigt das Ergebnis hinsichtlich der Signifikanz, da $|T| > 1{,}96$. Die Ergebnisse des Permutationstests sind in Tabelle 16 dargestellt.

Tabelle 16: Ergebnisse des Permutationstest

Rangreihenvergleich	D	SD	AV	T	p
ideale RR vs. RR Ärzte-SH	18	19,41	133	-5,92	0,32
ideale RR vs. RR Ärzte-BA	16	19,41	133	-6,03	0,88
RR Ärzte-SH vs. RR Ärzte-BA	2	19,41	133	-4,86	1,18

(RR = Rangreihe; SH = Schleswig-Holstein; BA = Bayern; D = Spearmansche Faustregel; SD = Standardabweichung; AV = Mittelwert; T = Teststatistik der Standardnormalverteilung, p = Überschreitungswahrscheinlichkeit)

3.5 Ergebnisse zur Bewertung der Begriffe im Kontext – Hypothese 3

Im folgenden Abschnitt wird Hypothese 3 überprüft:

> Es gibt Unterschiede in der prozentualen Bewertung von verbalen Wahrscheinlichkeitsbegriffen insbesondere hinsichtlich des Median der Paardifferenzen, abhängig davon, ob diese Begriffe einzeln in einer Liste aufgeführt werden oder in einen bestimmten Sachverhalt eingebettet sind.

H_0: Es gibt keine Unterschiede in der Bewertung der Mediandifferenz der Begriffe

- $A_{\text{häufig}} = B_{\text{häufig}}$ und $A_{\text{möglicherweise}} = D_{\text{möglicherweise}}$
- $A_{\text{gelegentlich}} = B_{\text{gelegentlich}}$ und $A_{\text{oft}} = D_{\text{oft}}$
- $A_{\text{selten}} = B_{\text{selten}}$ und $A_{\text{wahrscheinlich}} = D_{\text{wahrscheinlich}}$

H_1: Es gibt Unterschiede in der Bewertung der Mediandifferenz der Begriffe

- $A_{\text{häufig}} \neq B_{\text{häufig}}$ und $A_{\text{möglicherweise}} \neq D_{\text{möglicherweise}}$
- $A_{\text{gelegentlich}} \neq B_{\text{gelegentlich}}$ und $A_{\text{oft}} \neq D_{\text{oft}}$
- $A_{\text{selten}} \neq B_{\text{selten}}$ und $A_{\text{wahrscheinlich}} \neq D_{\text{wahrscheinlich}}$

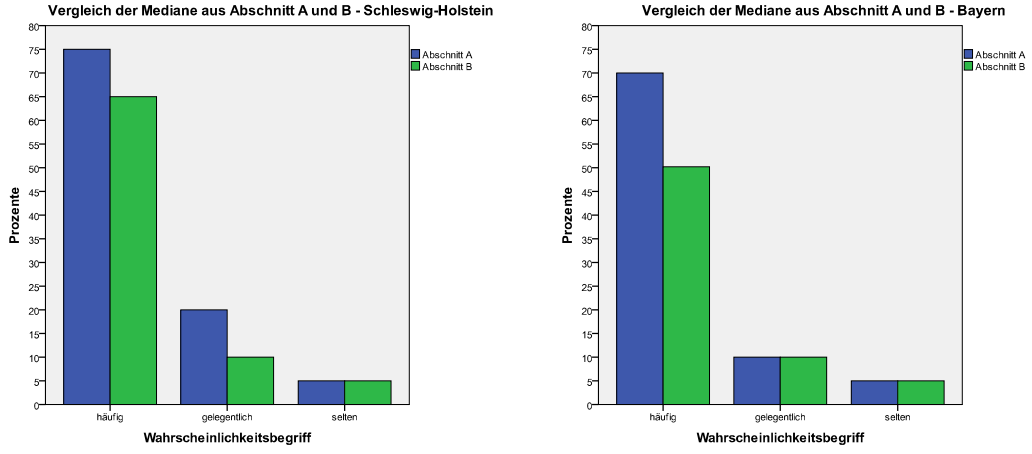
Die Ergebnisse bezüglich des Median im Vergleich der drei Begriffe aus der Liste (Abschnitt A) und dem Kontext in Abschnitt B des Fragebogens sind in Tabelle 17 dargestellt:

Tabelle 17: Ergebnisse aus der Berechnung des Median (Liste und Kontext B)

		N	Median	25%-Quartil	75%-Quartil
Schleswig-Holstein	häufig	178	65.00	30.00	75.00
	gelegentlich	179	10.00	5.00	20.00
	selten	177	5.00	1.00	5.00
Bayern	häufig	164	50.25	25.00	70.00
	gelegentlich	162	10.00	5.00	20.00
	selten	163	5.00	1.00	5.00

Beim Vergleich der Mediane für die beiden Bundesländer Bayern und Schleswig-Holstein zeigt diese Tabelle, dass die Teilnehmer aus Schleswig-Holstein mit 65.00% einen höheren Median für den Begriff „häufig“ verwenden, als die Probanden aus Bayern mit 50.25%. Für die Wahrscheinlichkeitsbegriffe „gelegentlich“ und „selten“ ist der Median für die beide Bundesländer gleich.
Eine übersichtlichere Darstellung zum direkten Vergleich der Begriffe bietet Abbildung 4:

Abbildung 4: Vergleich der Mediane aus Abschnitt A und B



Die Begriffe „möglicherweise – oft – wahrscheinlich“ aus der Liste (Abschnitt A) und dem dritten Sachverhalt (Abschnitt D) werden ebenfalls überprüft. Die Ergebnisse bezüglich des Median sind in Tabelle 18 dargestellt:

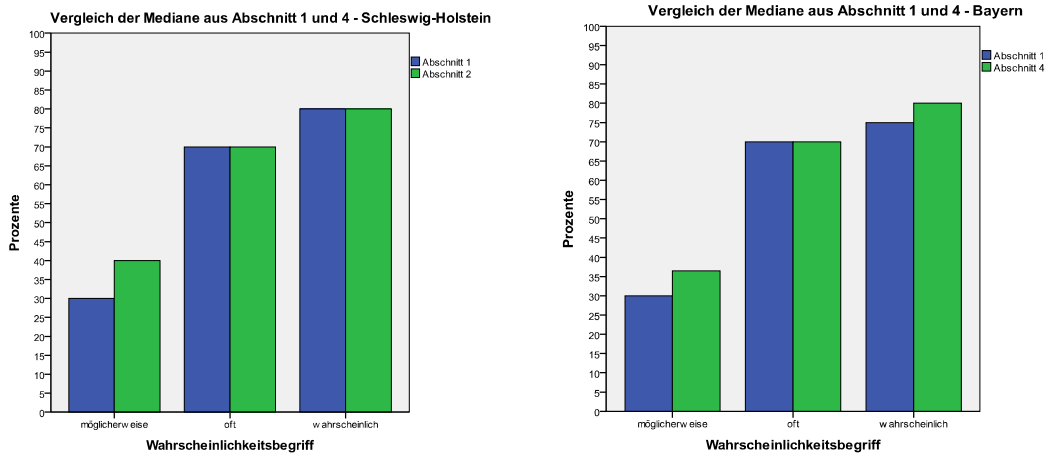
Tabelle 18: Ergebnisse aus der Berechnung des Median (Liste und Kontext D)

		N	Median	25%-Quartil	75%-Quartil
Schleswig-Holstein	möglicherweise	167	40.00	20.00	50.00
	oft	169	70.00	60.00	80.00
	wahrscheinlich	174	80.00	70.00	90.00
Bayern	möglicherweise	160	36.50	20.00	50.00
	oft	160	70.00	50.00	80.00
	wahrscheinlich	162	80.00	70.00	90.00

Beim Vergleich der Mediane für die beiden Bundesländer Bayern und Schleswig-Holstein zeigt diese Tabelle, dass die Teilnehmer aus Schleswig-Holstein mit 40.00% einen höheren Median für den Begriff „möglicherweise“ verwenden, als die Probanden aus Bayern mit 36.50%. Für die Wahrscheinlichkeitsbegriffe „oft“ und „wahrscheinlich“ ist der Median für die beide Bundesländer gleich.

Eine übersichtlichere Darstellung zum direkten Vergleich der Begriffe zeigt Abbildung 5:

Abbildung 5: Vergleich der Mediane aus Abschnitt A und D



Die entscheidende Frage ist jedoch, ob sich die vergebenen Prozentzahlen in Abschnitt A auch signifikant von den vergebenen Prozentzahlen im Fallbeispiel aus Abschnitt B unterscheiden und wird mit Hilfe des Wilcoxon-Rangsummen-Tests überprüft.

Im ersten Beispiel stellen die Wahrscheinlichkeitsbegriffe „häufig“ (ohne Kontext) aus Abschnitt A und der Begriff „häufig“ (Kontext) aus Abschnitt B die abhängigen Zufallsvariablen dar. Analog kann dies auf die Wahrscheinlichkeitsbegriffe „gelegentlich“ und „selten“ übertragen werden.

In Tabelle 19 sind die Ergebnisse des Wilcoxon-X-Tests für die Mediziner bezogen auf die Bundesländer erfasst. Es wird der Hodges-Lehmann-Punktschätzer für den Median der Mediandifferenzen und das dazugehörige 95%-Konfidenzintervall angegeben. Für den Hodges-Lehmann Schätzer werden alle möglichen Wertepaare der von den Studienteilnehmern aus Abschnitt A und B vergebenen numerischen Wahrscheinlichkeiten für die zwei abhängigen Verteilungen gebildet. Zwischen diesen Paaren wird die Differenz gebildet. Von diesen Differenzen wird der Median über alle Paare gebildet.

Tabelle 19: Ergebnisse des Wilcoxon-X-Tests für die Mediziner unterteilt nach Bundesland

	häufig			häufig	
	ohne Kontext – Kontext			ohne Kontext	Kontext
	Mediandifferenz	95% CI	p-Wert	Median	Median
Schleswig-Holstein	22.50	15.50;27.50	$5.52 * 10^{-12}$	75.00	65.00
Bayern	20.00	15.00;27.50	$7.56 * 10^{-12}$	70.00	50.25

	gelegentlich			gelegentlich	
	ohne Kontext – Kontext			ohne Kontext	Kontext
	Mediandifferenz	95% CI	p-Wert	Median	Median
Schleswig-Holstein	6.25	2.50;9.73	$5.93 * 10^{-5}$	20.00	10.00
Bayern	5.00	2.50;7.50	$1.53 * 10^{-4}$	10.00	10.00

	selten			selten	
	ohne Kontext – Kontext			ohne Kontext	Kontext
	Mediandifferenz	95% CI	p-Wert	Median	Median
Schleswig-Holstein	3.25	1.75;4.75	$2.21 * 10^{-5}$	5.00	5.00
Bayern	3.00	1.50;4.75	$5.77 * 10^{-5}$	5.00	5.00

Für alle Vergleiche ergeben sich signifikante p-Werte, das heißt es gibt einen Unterschied in der Vergabe der Prozentzahlen aus beiden Fragebogenteilen.

Im zweiten Beispiel stellen die Wahrscheinlichkeitsbegriffe „möglicherweise“ (ohne Kontext) aus Abschnitt A und „möglicherweise“ (Kontext) aus Abschnitt D die abhängigen Zufallsvariablen dar. Analog kann dies auf die Wahrscheinlichkeitsbegriffe „oft“ und „wahrscheinlich“ übertragen werden.

In Tabelle 20 sind die Ergebnisse dieses Wilcoxon-X-Tests für die Mediziner bezogen auf die Bundesländer erfasst. Für den Begriff „möglicherweise“ ergeben sich signifikante p-Werte, das heißt es gibt einen Unterschied in der Vergabe der Prozentzahlen aus beiden Fragebogenteilen. Das Gleiche gilt für den Begriff „wahrscheinlich“ bei der Vergabe der Prozentzahlen durch die Ärzte in Bayern.

Tabelle 20: Ergebnisse des Wilcoxon-X-Tests für die Mediziner unterteilt nach Bundesland

	möglicherweise ohne Kontext – Kontext			möglicherweise ohne Kontext	möglicherweise Kontext
	Mediandifferenz	95% CI	p-Wert	Median	Median
Schleswig-Holstein	-7.50	-12.50;-2.50	$6.26 * 10^{-3}$	30.00	40.00
Bayern	-9.50	-15.00;-3.50	$1.93 * 10^{-3}$	30.00	36.50

	oft ohne Kontext – Kontext			oft ohne Kontext	oft Kontext
	Mediandifferenz	95% CI	p-Wert	Median	Median
Schleswig-Holstein	5.00	0.00;7.50	$9.36 * 10^{-2}$	70.00	70.00
Bayern	0.00	-3.00;7.50	$5.42 * 10^{-1}$	70.00	70.00

	wahrscheinlich ohne Kontext – Kontext			wahrscheinlich ohne Kontext	wahrscheinlich Kontext
	Mediandifferenz	95% CI	p-Wert	Median	Median
Schleswig-Holstein	0.00	-5.00;2.50	$7.84 * 10^{-1}$	80.00	80.00
Bayern	-5.00	-8.75;0.00	$1.94 * 10^{-2}$	75.00	80.00

Bei der Bewertung des Begriffes „wahrscheinlich“ durch die Ärzte in Schleswig-Holstein ergeben sich keine signifikanten Unterschiede. Der Begriff „oft“ wird sowohl durch die Ärzte in Schleswig-Holstein als auch in Bayern so bewertet, dass kein signifikanter Unterschied vorliegt; in diesem Fall gilt die Nullhypothese.

3.6 Ergebnisse zur Bewertung der Begriffe durch die Ärzte in Schleswig-Holstein und Bayern – Hypothese 4

Im folgenden Abschnitt soll mit Hilfe des Mann-Whitney-Tests Hypothese 4 überprüft werden:

Ärzte in Schleswig-Holstein und Bayern unterscheiden sich bei der Bewertung in der Vergabe von Prozentzahlen bezüglich der sechs vorgegebenen Wahrscheinlichkeitsbegriffe.

H_0: Es gibt keine Unterschiede in der prozentualen Bewertung der Begriffe durch die Ärzte in Schleswig-Holstein und Bayern

- $A_{\text{häufig}} = B_{\text{häufig}}$ und $A_{\text{möglicherweise}} = D_{\text{möglicherweise}}$
- $A_{\text{gelegentlich}} = B_{\text{gelegentlich}}$ und $A_{\text{oft}} = D_{\text{oft}}$
- $A_{\text{selten}} = B_{\text{selten}}$ und $A_{\text{wahrscheinlich}} = D_{\text{wahrscheinlich}}$

H_1: Es gibt Unterschiede in der Bewertung der Begriffe durch die Ärzte in Schleswig-Holstein und Bayern

- $A_{häufig} \neq B_{häufig}$ und $A_{möglicherweise} \neq D_{möglicherweise}$
- $A_{gelegentlich} \neq B_{gelegentlich}$ und $A_{oft} \neq D_{oft}$
- $A_{selten} \neq B_{selten}$ und $A_{wahrscheinlich} \neq D_{wahrscheinlich}$

Die Ergebnisse für die Begriffe „häufig“, „gelegentlich“ und „selten“ aus der Liste (Abschnitt A) und dem Kontext B sind in Tabelle 21 dargestellt:

Tabelle 21: Mann-Whitney-Test / Bayern versus Schleswig-Holstein

		Bayern – Schleswig-Holstein		
Abschnitt A	(ohne Kontext)	Mediandifferenz	95%-CI	p-Wert
häufig		0.00	-5.00;0.00	$6.62 * 10^{-2}$
gelegentlich		0.00	-5.00;0.00	$2.25 * 10^{-2}$
selten		0.00	0.00;0.00	$7.98 * 10^{-1}$
Abschnitt B	(Kontext)			
häufig		-5,00	-10.00;0.00	$2.47 * 10^{-2}$
gelegentlich		0.00	-4.00;0.00	$9.79 * 10^{-.2}$
selten		0.00	0.00;0.00	$7.65 * 10^{-1}$

Für die verbalen Wahrscheinlichkeiten „häufig“ und „selten“ aus Abschnitt A liegt zwischen den bayerischen Ärzten und schleswig-holsteinischen Ärzten kein signifikanter Unterschied vor, aber für den Begriff „gelegentlich“. Für den Wahrscheinlichkeitsausdruck „häufig“ aus Abschnitt B liegt ein signifikanter p-Wert vor. Für die beiden verbalen Wahrscheinlichkeiten „gelegentlich“ (Kontext“) und „selten“ (Kontext) liegt zwischen Ärzten aus Bayern und Schleswig-Holstein kein signifikanter Unterschied vor.

Tabelle 22 stellt die Ergebnisse des Mann-Whitney-Tests für die aus den Variablen „häufig“ (ohne Kontext) aus Abschnitt A und „häufig“ (Kontext) aus Abschnitt B, den beiden Variablen „gelegentlich“ (ohne Kontext) und „gelegentlich“ (Kontext) sowie den beiden abhängigen Variablen „selten“ (ohne Kontext) und „selten“ (Kontext) jeweilig gebildete Differenzvariable dar. Bei allen Ergebnissen liegt kein signifikanter Unterschied vor.

Tabelle 22: Differenzvariable – Bayern vs. Schleswig-Holstein / Mann-Whitney-Test

	Bayern – Schleswig-Holstein		
	Mediandifferenz	95%-CI	p-Wert
häufig	0.00	-5.00;0.00	$5.89 * 10^{-1}$
gelegentlich	0.00	0.00;0.00	$8.30 * 10^{-1}$
selten	0.00	0.00;0.00	$7.67 * 10^{-1}$

Die Ergebnisse für die Begriffe „möglicherweise“, „oft“ und „wahrscheinlich“ aus der Liste (Abschnitt A) und dem Kontext D sind in Tabelle 23 dargestellt:

Tabelle 23: Mann-Whitney-Test / Bayern versus Schleswig-Holstein

		Bayern – Schleswig-Holstein		
Abschnitt A	(ohne Kontext)	Mediandifferenz	95%-CI	p-Wert
möglicherweise		0.00	-5.00;0.00	$2.73 * 10^{-1}$
oft		0.00	-5.00;0.00	$6.07 * 10^{-1}$
wahrscheinlich		-3.00	-5.00;0.00	$5.56 * 10^{-2}$
Abschnitt D	(Kontext)			
möglicherweise		0.00	-5.00;0.00	$5.64 * 10^{-1}$
oft		0.00	-5.00;0.00	$5.76 * 10^{-1}$
wahrscheinlich		0.00	0.00;0.00	$8.80 * 10^{-1}$

Die Tabelle zeigt, dass keine signifikanten p-Werte vorliegen und damit kein signifikanter Unterschied in der Vergabe der Prozentzahlen zwischen den Medizinern in Bayern und Schleswig-Holstein zu beobachten ist.

Tabelle 24 stellt die Ergebnisse des Mann-Whitney-Tests für die aus den Variablen „möglicherweise“ (ohne Kontext) aus Abschnitt A und „möglicherweise (Kontext) aus Abschnitt D, den beiden Variablen „oft (ohne Kontext) und „oft“ (Kontext) sowie den beiden abhängigen Variablen wahrscheinlich (ohne Kontext) und wahrscheinlich (Kontext) jeweilig gebildete Differenzvariable dar.

Tabelle 24: Differenzvariable – Bayern vs. Schleswig-Holstein / Mann-Whitney-Test

	Bayern – Schleswig-Holstein		
	Mediandifferenz	95%-CI	p-Wert
möglicherweise	0.00	0.00;2.50	$5.69 * 10^{-1}$
oft	0.00	0.00;5.00	$3.28 * 10^{-1}$
wahrscheinlich	0.00	0.00;5.00	$1.42 * 10^{-1}$

Bei allen Ergebnissen liegt auch hier kein signifikanter Unterschied vor, das heißt, es gibt keine Unterschiede in der prozentualen Bewertung der Begriffe durch die Ärzte beider Bundesländer.

3.7 Ergebnisse der Auswertung zu den Wahrscheinlichkeitsquoten

3.7.1 Angaben der Ärzte aus Schleswig-Holstein und Bayern

In Anlehnung an die SPC-Richtlinie „Häufigkeitsangaben von Nebenwirkungen“ wurden den Studienteilnehmern in Abschnitt C des Fragebogens fünf Wahrscheinlichkeitsquoten vorgegeben. In den nachfolgenden Tabellen ist der durch die Ärzte jeweils am häufigsten verwendete Wahrscheinlichkeitsbegriff zur entsprechenden Quote aufgeführt. Tabelle 25 stellt die Ergebnisse der Ärzte aus Schleswig-Holstein (n = 170) dar, Tabelle 26 die Antworten der Ärzte aus Bayern (n = 158).

Tabelle 25: Am häufigsten genannter Begriff zur vorgegebenen Quote der Ärzte Schleswig-Holsteins

Quoten-Vorgaben	Meistgenannter Begriff	Anzahl Teilnehmer	Häufigkeit in Prozent
1 : 1000	sehr selten/ selten	62/ 62 (124)	36.5/ 36.5 (73.0)
1 : 100	selten	76	44.7
1 : 10	gelegentlich	55	32.4
1 : 5	häufig	52	30.8
1 : 2	häufig	46	27.4

Tabelle 26: Am häufigsten genannter Begriff zur vorgegebenen Quote der Ärzte Bayerns

Quoten-Vorgaben	Meistgenannter Begriff	Anzahl Teilnehmer	Häufigkeit in Prozent
1 : 1000	sehr selten	64	40.5
1 : 100	selten	73	45.9
1 : 10	gelegentlich	52	33.1
1 : 5	häufig	41	26.1
1 : 2	sehr häufig	32	20.3

Dabei ist hinzuzufügen, dass die Ärzte aus Schleswig-Holstein zur Quote 1:1000 sowohl den Begriff „sehr selten“ als auch den Begriff „selten“ gleich häufig verwendet haben. Außerdem haben die Ärzte in Bayern zur Quote „1:2“ den Begriff „sehr häufig“ benutzt im Gegensatz zu den Ärzten aus Schleswig-Holstein mit dem Begriff „häufig“.

Es ist zu erkennen, dass sowohl die Ärzte aus Schleswig-Holstein als auch die Ärzte aus Bayern eine typische Reihenfolge der Begriffe eingehalten haben. Auch der prozentuale Anteil der Antworten durch die Teilnehmer beider Datenpools stimmt annähernd überein.

Im Anhang (Anlage 9.1.8) finden sich ausführliche Übersichtstabellen mit den häufigsten Begriffen, die zu den jeweiligen Wahrscheinlichkeitsquoten aufgeführt wurden.

3.7.2 Vergleich der ausgewerteten Begriffe mit der Richtlinie

Die von den Teilnehmern der Studie vergebenen numerischen Prozentwerte werden mit der vom Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) herausgegebenen Nebenwirkungstabelle (2) verglichen. Es wird geprüft, ob die von den Teilnehmern der Befragung genannten numerischen Wahrscheinlichkeiten für die verbalen Wahrscheinlichkeitsbegriffe „häufig“, „gelegentlich“ und „selten“, den numerischen Wahrscheinlichkeiten für die verbalen Wahrscheinlichkeiten „häufig“, „gelegentlich“ und „selten“ nach der offiziellen Klassifikation für Nebenwirkungen entsprechen. Den Wahrscheinlichkeitsangaben aus Abschnitt C des Fragebogens lassen sich die verbalen Wahrscheinlichkeitsbegriffe für Nebenwirkungen wie folgt zuordnen (Tabelle 27).

Tabelle 27: Zuordnung der verbalen Wahrscheinlichkeitsbegriffe zu den numerischen Wahrscheinlichkeiten

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

Tabelle 28 zeigt nur die Anzahl der Antworten aller Ärzte, die nach den Vorgaben laut Tabelle 27 als korrekt zugeordnet zur entsprechend vorgegebenen Quote gelten.

Tabelle 28: Korrekt zugeordneter Begriff zur vorgegebenen Quote Ärzte Schleswig-Holstein und Ärzte Bayern

Quoten	Definierte verbale Wahrscheinlichkeit	Angaben Ärzte Schleswig-Holstein		Angaben Ärzte Bayern	
		Anzahl	Prozent	Anzahl	Prozent
1 : 1000	gelegentlich	4	2.4	5	3.2
1 : 100	häufig	7	4.1	3	1.9
1 : 10	sehr häufig	4	2.4	1	0.6
1 : 5	sehr häufig	7	4.1	13	8.3
1 : 2	sehr häufig	27	16.1	32	20.3

Tabelle 29 stellt die Anzahl der Antworten aller Ärzte zu den vorgegebenen Quoten dar. Zur besseren Verdeutlichung der korrekt und inkorrekt genannten numerischen Wahrscheinlichkeiten dient die Gegenüberstellung der Angaben durch die Ärzte in Schleswig-Holstein (n = 170) und Bayern (n = 158).

Tabelle 29: Korrekte und inkorrekte Antworten der Studienteilnehmer

		1:1000		1:100		1:10		1:5		1:2	
Schleswig-	korrekt	4	2.4%	7	4.1%	4	2.4%	7	4.1%	27	16.1%
Holstein	inkorrekt	166	97.6%	163	95.9%	166	97.6%	162	95.9%	141	83.9%
Bayern	korrekt	5	3.2%	3	1.9%	1	0.6%	13	8.3%	32	20.3%
	inkorrekt	153	96.8%	155	98.1%	156	99.4%	144	91.7%	126	79.7%

Nicht alle Ärzte haben zu den fünf Wahrscheinlichkeitsquoten vollständige Angaben gemacht. Einige Ärzte haben einzelne Begriffe ausgelassen, so dass die unterschiedlichen Gesamtzahlen in Tabelle 29 zustande kommen.

4. Diskussion

Die vorliegende Studie untersuchte, ob der Gebrauch von umgangssprachlichen Wahrscheinlichkeitsbegriffen in der ärztlichen Tätigkeit innerhalb Deutschlands einheitlich ist. Um dieses exemplarisch zu überprüfen, wurden Ärzte aus zwei großen, von der Sprache her sehr unterschiedlichen Bundesländen ausgewählt – Schleswig-Holstein und Bayern.

Bei der Auswahl der Wahrscheinlichkeitsbegriffe wurde auf eine Liste von 20 Begriffen zurückgegriffen, die zu den am häufigsten genutzten Wahrscheinlichkeitsbegriffen in der Literatur und dem tägliche Sprachgebrauch in Deutschland zählen, und die bereits in einzelnen Studien zu diesem Thema untersucht worden waren. Von großem Interesse war in dieser Studie auch, inwiefern diese allgemein gebräuchlichen Begriffe verstanden und interpretiert werden. Als ein Beispiel sollten die bekannten Häufigkeitsbegriffe zu Nebenwirkungen von Medikamenten (zu finden in jedem Beipackzettel) „sehr häufig – häufig – gelegentlich – selten – sehr selten“ dienen. Diese werden anhand entsprechender Zahlenangaben im Beipackzettel erläutert. Zu erwarten war, dass diese somit von den Ärzten, unabhängig davon in welchem Bundesland sie tätig sind, einheitlich bewertet werden.

4.1 Aufbau und Ablauf der Studie

Als Vorlage dieser Studie war ein Fragebogen aus einer 1997 veröffentlichte Doktorarbeit von *Suß* mit dem Thema „Ist die Rangordnung von umgangssprachlichen Wahrscheinlichkeitsausdrücken in der umgangssprachlichen Kommunikation unter Ärzten konsistent?“ (39) vorgegeben worden. In dieser Doktorarbeit wurden verschiedene, sehr statistisch formulierte Fragebogenvarianten an angiologisch tätige Ärzte verschickt. Es waren die 20 Wahrscheinlichkeitsbegriffe zunächst in einer Liste und anschließend in mehreren Sachverhalten vorgegeben worden. Diese sollten von den Ärzten prozentual bewertet werden. Die Ergebnisse wurden hinsichtlich einer Rangreihenfolge untersucht, sowie der Vergleich mit einer ‚idealen’ Rangreihenfolge vorgenommen. Bei dieser Studie fiel die sehr niedrige Rücklaufquote von insgesamt 14,55% auf.

Aus diesem Grund musste der von *Suß* vorgegebene Fragebogen, der sehr statistisch formuliert worden war, geändert werden. An erste Stelle stand deshalb, dass der Fragebogen dieser Studie nicht zu abstrakt und zu statistisch gestaltet wurde. Mit alltäglichen Beispielen und wenig statistisch klingenden Formulierungen sollten die

Teilnehmer motiviert werden, bei dieser Studie mitzumachen. Es fällt erfahrungsgemäß leichter, zu einem Sachverhalt spontan zu antworten, der einem im täglichen Leben so oder so ähnlich bereits begegnet ist, oder in den man sich mühelos hineinversetzten kann. Das Ergebnis dieser Formulierungen war eine Rücklaufquote von insgesamt 60,17%.

Für diese Studie sollten Ärzte aus unterschiedlichen Bundesländern, exemplarisch für ganz Deutschland, befragt werden. Anhand der vier formulierten Hypothesen sollte untersucht werden, wie groß Unterschiede und Gemeinsamkeiten bezüglich der Bewertung von Wahrscheinlichkeitsbegriffen innerhalb Deutschlands sind. Die Entscheidung fiel auf Ärzte aus dem nördlichsten und aus dem südlichsten Bundesland, da die sprachlichen Unterschiede hier sehr deutlich sind. Da mein Arbeitsort derzeit in Schleswig-Holstein liegt, war es aufgrund organisatorischer Gegebenheiten einfacher, Ärzte in Schleswig-Holstein zu befragen und nicht zum Beispiel Ärzte in Mecklenburg-Vorpommern. Es ist jedoch anzunehmen, dass auch jene Auswahl an Ärzten im Antwortverhalten keinen Unterschied gemacht hätte. Bayern wurde als flächenmäßig größtes und südlichstes Bundesland gewählt. Ein interessanter Aspekt waren insbesondere die unterschiedliche sprachliche Entwicklung sowie die Sprachvariationen innerhalb dieser beiden Bundesländer und damit der unterschiedlichen Studienteilnehmer. Zu prüfen war, ob diese Sprachentwicklung einen großen Einfluss auf die Antworten der Studienteilnehmer hat, obwohl Hoch- oder Standarddeutsch in ganz Deutschland bereits ab dem Kindergarten gelehrt wird, und somit unabhängig von dem Dialekt überall in Deutschland verstanden wird. Vor allem in so speziellen Bereichen wie der Medizin gibt es eine feste Standard- und Fachsprache, die sich insbesondere in der Literatur widerspiegelt und in der Medizin gelehrt wird.

Das Verhältnis zwischen den Ärzten, die innerklinisch tätig sind und somit eher untereinander kommunizieren und solchen, die niedergelassen sind beziehungsweise einen ausgedehnten Patientenkontakt haben, wurde mit der Wahl der Fachrichtungen Innere Medizin und Anästhesie umgesetzt. Auch hier war zu erwarten, dass Ärzte aus anderen Fachrichtungen ähnlich geantwortet hätten.

Zu Beginn der Studie wurde ein verkürzter Pre-Test durchgeführt. Der inhaltliche Kontext und vor allem die Verständlichkeit des Fragebogens sowie der entsprechende Zeitaufwand zum Beantworten der Fragen waren wichtig. Um eine möglichst große Streuung hinsichtlich des medizinischen und alltäglichen Umfeldes und der demographischen Spannbreite zu erhalten, wurden verschiedenen Personen aus einem breiten persönlichen und beruflichen Umfeld befragt. Bezüglich des organisatorischen Ablaufes war die erste

Welle ausschlaggebend und wurde diesbezüglich wie ein Pre-Test gewertet. Dieses Vorgehen hatte verschiedene Gründe. Ein Grund für das Nicht-Versenden von Pre-Test-Fragebögen war die Aktualität der Adressen. Wäre es trotzdem zu großen Ausfällen in der ersten Welle gekommen, so hätte sicherlich in der zweiten Welle etwas geändert werden müssen. Wie vermutet, hielt sich die Anzahl der Ausfälle in Grenzen. Letztendlich soll der Pre-Test zu erwartende Probleme aufdecken, kann diese zu einem späteren Zeitpunkt aber nicht völlig ausschließen oder im Sinne von *Koschnik* „Was später passiert, kann man vorher nicht wissen, sondern im günstigsten Fall abschätzen“ (57).

Rücklauf und Ausschöpfung lagen, gemessen an der Arbeit von *Suß* (39), weit über den Erwartungen. Gerade weil bekannt ist, dass Ärzte im Allgemeinen wenig Zeit und Interesse an freiwilligen Studien haben, überraschte der gute und relativ zügige Rücklauf. Ein entscheidender Grund hierfür war vermutlich auch das unkomplizierte Anschreiben. Der Teilnehmer wurde nicht – wie leider oft üblich – mit Fachbegriffen überschüttet, sondern es wurde auf einfache und nette Art versucht, das Interesse für diese Studie zu wecken. Das ist in dieser Studie gelungen, da auch das Angebot per Telefon oder Internet mit mir Kontakt aufzunehmen, in vielen Fällen in Anspruch genommen wurde. Ein anderer Grund wird der Fragebogen an sich sein, sowie das Erinnerungsschreiben, welches nach jeweils circa drei und vier Wochen nach dem Fragebogen verschickt worden war. Der Rücklauf stieg nach dem Erinnerungsschreiben nochmals deutlich an (Abschnitt 3.1, Abbildung 3). Der Rücklauf war in beiden Wellen in Schleswig-Holstein insgesamt etwas höher als in Bayern. In der zweiten Welle war der Rücklauf sogar größer als in der ersten Welle. Das trifft sowohl für die Ärzte aus Schleswig-Holstein als auch für die Ärzte in Bayern zu. Ob dieses am späteren Versenden des Erinnerungsschreibens liegt, bleibt nur zu vermuten. Insgesamt war die Ausschöpfung in meiner Studie mit 170 Fragebögen in der ersten Welle (57,63%) und mit 180 Fragebögen in der zweiten Welle (61,20%) erfreulich hoch. Vergleicht man die Ergebnisse der Ausschöpfung mit ähnlichen Studien zu diesem Thema, so darf man mit den Ergebnissen zufrieden sein. Die Ausschöpfung belief sich zum Beispiel in der Studie von *Suß* auf 15,20% im ersten Hauptfeld und 11,60% im zweiten Hauptfeld (39). Vergleicht man aber die Ergebnisse allgemein mit den geforderten Ausschöpfungsergebnissen in der Literatur, so ist das Ergebnis ernüchternd. So beschreibt *Dillmann*, dass die Gesamtausschöpfung einer Studie, sofern man die Total-Design-Methode (TDM) verwendet, nie unterhalb von 60% liegen sollte (58). Insgesamt findet man Ausschöpfungsquoten von 10 % bis 90 % (38).

4.2 Ergebnisse

Zu Beginn des Fragebogens war eine Liste mit 20 Wahrscheinlichkeitsbegriffen vorgegeben worden, die prozentual zu bewerten waren. Beeindruckend war die unglaublich große Spannbreite in der Bewertung dieser Begriffe sowohl in Schleswig-Holstein als auch in Bayern. So waren bei 9 der 20 Begriffe Spannbreiten von 0 – 100% aufgefallen. Dies galt bei allen Ärzten für die Begriffe „niemals – unsicher – unklar – fraglos – stets – nie – immer – typischerweise“, den Begriff „sicher“ der Ärzte in Schleswig-Holstein und den Begriff „fraglich“ bei den Ärzten in Bayern. Insgesamt wurden 11 der 20 Begriffe bei den Ärzten in Schleswig-Holstein mit mindestens 0% bewertet und 15 Begriffe mit maximal 100%. Bei den Ärzten in Bayern konnten 9 der 20 Begriffe 0% bedeuten und 12 Begriffe 100%. Die Frage bei diesen Ergebnissen ist, ob die Fragestellung im Fragebogen nicht eindeutig klar formuliert worden war oder ob die individuellen Bewertungen tatsächlich so unterschiedlich waren. Diese Zahlen regen auf alle Fälle im Hinblick auf eine einheitliche Bewertung von Wahrscheinlichkeitsbegriffen zum Nachdenken an. Obwohl es individuell massive Unterschiede in der Bewertung der 20 Begriffe gab, so stellten sich die Mittelwerte in der Gesamtheit der beiden Gruppen – Ärzte aus Schleswig-Holstein sowie Ärzte aus Bayern – relativ einheitlich dar. Bei der Untersuchung der Zusammenhänge zwischen diesen beiden Gruppen wurden echte Rangreihen ermittelt, das heißt abhängig vom Mittelwert wurden die 20 Begriffe in eine Reihenfolge mit bestimmter Rangordnung gebracht. Anhand verschiedener Testverfahren wurde gezeigt, wie groß die Übereinstimmung in der Verwendung der Wahrscheinlichkeitsbegriffe bezüglich einer Rangreihenfolge unter den Ärzten der beiden Bundesländer untereinander tatsächlich ist. Schaut man sich die Rangreihen im Vergleich der Ärzte untereinander an, so fallen als Unterschiede lediglich die Positionen 2 und 3 auf. Die Ärzte aus Schleswig-Holstein hatten auf Rangplatz 2 den Begriff „nie“ genannt, auf Rangplatz 3 den Begriff „selten“. Bei den Ärzten aus Bayern lagen diese Nennungen genau anders herum. Die große Übereinstimmung in 18 Rangpositionen machte deutlich, dass die Wahrscheinlichkeitsbegriffe und Prozentangaben relativ einheitlich in ihrer Stellung untereinander benutzt wurden. Wie die Begriffe „nie“ und „selten“ deutschlandweit eingeordnet werden, bleibt zumindest in dieser Studie ungeklärt.

Ein weiterer Vergleich bezüglich der beiden echten Rangreihen wurde mit einer idealen Rangreihe durchgeführt. In den Testverfahren wurde eine deutliche Korrelation der Reihen zu dieser idealen Reihe festgestellt (z.B. r_s Ärzte-SH = 0,973; r_s Ärzte-BA = 0,974). Vergleicht man jedoch die eigenen Rangreihen in der Gegenüberstellung mit der idealen

Rangreihe von 1997, so fallen erhebliche Diskrepanzen auf. Hat sich seitdem das Verständnis der Ärzte zu Wahrscheinlichkeitsbegriffen geändert? Dieses ist zu vermuten, da sich insgesamt 12 Begriffe in ihren Positionen im Vergleich mit der idealen Rangreihe unterscheiden. Lediglich „niemals“ hat immer den Rang 1 und „oft“ den Rangplatz 10. Auch die Positionen 17 bis 20 sind mit den Begriffen „fraglos – stets – sicher – immer“ konsistent geblieben. Letztendlich bleibt aber auch diese Frage offen und zu diskutieren, ob es überhaupt eine ‚ideale' Rangreihe gibt.

Hinsichtlich der Kontextabhängigkeit wurde geprüft, ob bestimmte Wahrscheinlichkeitsbegriffe der Liste in Abhängigkeit von einem Sachverhalt anders bewertet werden. Vorgegeben waren im Sachverhalt B die Begriffe „häufig – gelegentlich – selten“ und im Sachverhalt D die Begriffe „möglicherweise – oft – wahrscheinlich“. In den Testergebnissen wurde deutlich, dass die Unterschiede im Gebrauch der Begriffe in einer Liste oder eingebettet in einen Sachverhalt doch zum Teil signifikant waren. Bei allen Begriffen aus dem Sachverhalt B war das der Fall. Ebenso im Sachverhalt D bei den Begriffen „möglicherweise“ bei allen Ärzten und „wahrscheinlich“ bei den Ärzten in Bayern. Allein der Begriff „oft“ scheint in seiner Verwendung konsistent zu sein. Auch gab es keinen signifikanten Unterschied im Gebrauch des Begriffs „wahrscheinlich“ durch die Ärzte in Schleswig-Holstein. War dieses Ergebnis zu erwarten? Eigentlich sollten alle Begriffe stets die gleiche Bedeutung haben, jedoch weiß man aus der Literatur (z.B. 7, 13, 18, 19, 20, 21) , dass dieses leider nicht der Fall ist. Auch in dieser Studie bestätigte sich die überwiegende Diskrepanz der Bewertung der Begriffe in der Liste und einem Kontext. Lediglich der Begriff „oft“, der auch in den Rangreihen konsistent war, wird auch hier ohne signifikanten Unterschied bewertet. Warum der Begriff „wahrscheinlich“ geographisch so verschieden bewertet wird, bleibt ungeklärt. Ebenso bleibt die Frage offen, ob sich die Ergebnisse bei einer anderen Auswahl an Begriffen geändert hätten.

Anhand der sechs Begriffe aus den Sachverhalten B und D wurde untersucht, wie groß die Unterschiede in der Bewertung dieser Begriffe durch die Ärzte in Bayern und Schleswig-Holstein sind. Sowohl die Bewertung im Kontext als auch die Bewertung in der Liste für diese Begriffe gingen in die Auswertung ein. Signifikante Unterschiede zwischen den Ärzten in Schleswig-Holstein und den Ärzten in Bayern gab es nur bezüglich des Begriffs „gelegentlich“ bei der Bewertung in der Liste und dem Begriff „häufig“ in seiner Verwendung im Sachverhalt. Bei allen übrigen Begriffen gab es keine signifikanten Unterschiede. Stellen diese Ergebnisse zufrieden? Zumindest zeigen diese Testergebnisse, dass die Unterschiede innerhalb der Ärztegruppen zu diesen Begriffen im Nord-Südgefälle

nicht sehr groß sind. Man könnte erwarten, dass dann auch innerhalb Deutschlands die Bewertung so ausfallen könnte und sich somit die Unterschiede in der Bewertung von Wahrscheinlichkeitsbegriffen in Grenzen halten. Jedoch wurden nur sechs Begriffe von vielen getestet und es bleibt offen, wie es sich mit der Bewertung anderer Begriffe verhält.
Unabhängig von den vier Hypothesen wurde das Verständnis der Ärzte zu geläufigen Wahrscheinlichkeitsquoten untersucht. Vorgegeben waren die Wahrscheinlichkeitsquoten „1:1000 – 1:100 – 1:10 – 1:5 – 1:2“. Wahrscheinlichkeitsquoten sind in Deutschland vor allem aus dem Beipackzettel von Medikamenten zur Darstellung der unerwünschten Wirkungen seit langem bekannt. Da Quoten sonst sehr selten verwendet werden, sind diese in den Beipackzetteln entsprechend mit „sehr häufig – häufig – gelegentlich – selten – sehr selten“ vorgegeben und bewertet. Die in dieser Studie vorgegebenen Quoten sollten somit nach dieser Einteilung mit „gelegentlich – häufig – sehr häufig (für entsprechend 1:10 bis 1:2) bewertet werden. Wie im Abschnitt 3.7 und dem Anhang 9.1.8 dargestellt, wurde eine Vielzahl von Begriffen zu den vorgegeben Quoten von den Ärzten angegeben. Es ist zu erkennen, dass sowohl die Ärzte aus Schleswig-Holstein als auch die Ärzte aus Bayern eine typische Reihenfolge in ihren Begriffsangaben eingehalten haben. Auch der prozentuale Anteil der Antworten durch die Teilnehmer beider Datenpools stimmte annähernd überein. Es ist zu vermuten, dass die typische Reihenfolge der Wahrscheinlichkeitsbegriffe „sehr selten – selten – gelegentlich – häufig“ auch bei anderen vorgegebenen Wahrscheinlichkeitsquoten in ähnlich absteigender Form angegeben worden wäre. Einzig der Begriff „sehr häufig“ (in Anlehnung an die SPC-Richtlinie) fehlt in der Auswertung der Ärzte in Schleswig-Holstein. Beeindruckend ist auch die kleine Anzahl der Studienteilnehmer, die die Begriffe anhand der SPC-Richtlinie „richtig“ ausgelegt hatten. So hatten zum Beispiel nur 4 von 170 Ärzten in Schleswig-Holstein die Quote von 1:1000 mit „gelegentlich“ bewertet (Ärzte-BA: 5 von 158). Lediglich der Begriff „sehr häufig“ zur Quote von 1:2 wurde von mehreren Ärzten aus beiden Bundesländern korrekt angegeben (Ärzte SH: 27 von 168 und Ärzte-BA: 32 von 158). Die Ergebnisse machen deutlich, wie wenig einheitlich die Bewertung von Wahrscheinlichkeitsbegriffen und Quoten ist und wie dringend Aufklärungsarbeit und Schulungen von Nöten sind. Kann ich sonst meinem Arzt glauben, wenn er mir erzählt, dass eine Nebenwirkung „selten“ ist?

Anhand sämtlicher Daten ist festzuhalten, dass die Gemeinsamkeiten im Umgang und in der Bewertung von Wahrscheinlichkeitsbegriffen unter den Ärzten in Schleswig-Holstein und Bayern deutlich größer sind als die Unterschiede. Es bleibt aber auch festzuhalten,

dass die Unsicherheit im Umgang und in der „korrekten“ Bewertung von Wahrscheinlichkeiten unabhängig vom Bundesland vermutlich deutschlandweit besteht.

4.3 Diskussion im Überblick

Die Herausforderungen, die während der gesamten Studie auftraten, waren vielfältig. Die erste Herausforderung war die Erstellung eines ansprechenden Fragebogens. Dieser sollte mittels Anschreiben zum Weiterlesen motivieren. Die Umsetzung, eine statistische Studie in einen eher einfach gehaltenen und stellenweise amüsanten Fragebogen zu verpacken, gelang letztendlich nach vielen Stunden und Tagen.

Die nächste Herausforderung waren per se die Studienteilnehmer. Ärzte haben im Allgemeinen wenig Zeit und noch weniger Lust, in ihrer Freizeit „lästige“ Studien zu beantworten. Umso erfreulicher waren der erzielte Rücklauf sowie eine Ausschöpfung von insgesamt fast 60%. Auch der Spagat zwischen den beiden Fachrichtungen „Anästhesie“ und „Innere Medizin“ scheint gelungen. Wichtig und notwendig ist in diesem Zusammenhang auch das Erinnerungsschreiben zu sehen.

Insgesamt gesehen waren die Bewältigung und Umsetzung der oben genannten Herausforderungen erfolgreich.

Diese Studie macht deutlich, dass sich die Probleme zum Verständnis von Wahrscheinlichkeitsbegriffen und ihre Auslegung trotz langjähriger Untersuchungen und diverser Studien nicht geändert haben und auch in Deutschland weiterhin ein interessantes Thema ist und vermutlich auch bleibt.

Seit 1980 war die große Variationsbreite bei der Zuordnung von Zahlen ein Problem (8, 9, 10). Auch in dieser Studie fällt auf, dass individuell zu vielen Begriffen eine sehr große Spannweite in der Bewertung vorherrscht. Beispielsweise reichen die Angaben zu den Begriffen „fraglos, stets und nie“ von 0 – 100%!

Ebenso bestätigt sich die Abhängigkeit in der Bewertung eines Begriffes vom Kontext. Dieses hatte bereits *Pepper* 1974 (17) erstmals festgestellt. Bis in die Gegenwart war die Kontextabhängigkeit Thema vieler Studien (7, 13, 18, 19, 20, 21) und konnte auch mit dieser Studie eindrücklich belegt werden. Eine weitere Übereinstimmung dieser Studie mit der Literatur besteht in der Unsicherheit zu Häufigkeitsangaben von Nebenwirkungen im Rahmen der Umsetzung der SPC-Richtlinie. Diese Studie zeigt sehr deutlich, dass die Probleme zum Verständnis der Begrifflichkeiten auch in Deutschland trotz langjähriger Erfahrung und Arbeit damit weiterhin bestehen. Studien zu diesem Thema und der Umsetzung in anderen Ländern zeigten ähnliche Probleme (26, 27, 29).

Anhand dieser Ergebnisse wird deutlich, dass die Forschung auf dem Gebiet der medizinischen Risikokommunikation weiter verstärkt werden muss. Ärzte sollten wenigstens untereinander einheitlich kommunizieren. Studien zu diesem Thema fordern, dass weitere Risikoformate benutzt werden müssen, da bisher weder Begriffe, noch Zahlen oder Graphiken alleine eine befriedigende Lösung sind (31, 32, 33).

Die Forschung und Aufklärung auf diesem Gebiet muss weitergehen, damit die Auslegung von Wahrscheinlichkeitsbegriffen standardisiert wird. Wahrscheinlichkeiten sollten nicht so ausgelegt werden, wie *Dolezal* sie 1976 (59) in seinem Artikel `Translating medical idiom´ ironischerweise beschrieb:

`In my experience….´	I have seen ONE such patient.
`In my series….´	I have seen TWO such patients.
`In case, after Case, after CASE….´	I have seen THREE such patients.
`I haven`t had a whole lot experience with this….´	Well, ah, actually, ah, never seen one before.

5. Zusammenfassung

In unserem Beruf als Ärzte gehen wir täglich ganz selbstverständlich mit Begriffen um, die eine Wahrscheinlichkeit ausdrücken sollen. Wir benutzen Ausdrücke wir „häufig“, „wahrscheinlich“ und „selten“ und haben dabei einen ganz bestimmten Prozentwert im Kopf. Aber interpretiert unser Gegenüber – sei es nun Arzt, Klinikapotheker oder Patient – diesen Begriff genauso? Und macht es einen Unterschied, ob ich in Schleswig-Holstein oder in Bayern tätig bin? Da aus der Literatur bekannt ist, dass es eine große Unsicherheit in der Bewertung von Wahrscheinlichkeitsbegriffen gibt, sollte diese Studie klären, ob sich diese Unsicherheit auch bei den Ärzten in Deutschland widerspiegelt und ob Unterschiede im Gebrauch dieser Begriffe bezüglich Nord- und Süddeutschland bestehen.

Im Rahmen dieser Studie wurden insgesamt 600 Fragebögen an Ärzte in Schleswig-Holstein und Bayern verschickt. Befragt wurden Ärzte der Fachrichtungen Anästhesie und Innere Medizin. Die Fragebögen wurden in zwei Wellen verschickt, gefolgt von einem Erinnerungsschreiben im Abstand von drei beziehungsweise vier Wochen. Die Gesamtausschöpfung belief sich insgesamt auf 350 Fragebögen (58.9%).

Grundlage für verschiedene Fragestellungen und Hypothesen waren die Daten zur Bewertung von 20 in einer Liste vorgegebenen Wahrscheinlichkeitsbegriffen. Bei der Bewertung dieser Begriffe fiel bei allen Ärzten eine sehr große Spannweite auf. So sind nach den Ergebnissen dieser Liste zu 9 dieser 20 Begriffe Angaben von 0 – 100% möglich! Anhand verschiedener Testverfahren konnten diese Begriffe letztendlich aber in feste Rangreihenfolgen gebracht werden. Die Rangreihenfolgen der Ärzte aus Schleswig-Holstein und Bayern zeigten insgesamt große Gemeinsamkeiten (Hypothese 1). Diese Rangreihen mussten dem Vergleich mit einer idealen Rangreihe standhalten (Hypothese 2), wobei auch hier mehr Gemeinsamkeiten als Unterschiede zu finden waren. Eine weitere Fragestellung war, inwiefern sich sechs Begriffe in ihrer Bewertung verändern, sofern sie nicht mehr nur in einer Liste aufgezählt werden, sondern in einen Kontext eingebettet sind (Hypothese 3). Von diesen Begriffen waren beinahe alle kontextabhängig, nur der Begriff „oft“ wurde sowohl in der Liste als auch im Kontext mit 70% und somit gleich bewertet. Nicht nur die Kontextabhängigkeit sondern auch die Bewertung innerhalb der beiden Bundesländer wurde näher untersucht (Hypothese 4). Alle statistischen Testverfahren haben insgesamt eine sehr große Übereinstimmung in der Bewertung der Wahrscheinlichkeitsbegriffe unter den Ärzten in Schleswig-Holstein und Bayern gezeigt. Deutlich wurde aber auch die Unsicherheit, die weiterhin in einer eindeutigen Darstellung

der einzelnen Wahrscheinlichkeitsbegriffe besteht. Der Begriff „häufig“ beispielsweise entspricht im Mittel bei den Ärzten in Schleswig-Holstein 75%, sofern er in einer Liste aufgeführt ist, oder 65% in einem vorgegebenen Kontext (Ärzte Bayern 70% in der Liste versus 50,25% im Kontext).

Im Rahmen dieser Studie wurde auch das Verständnis der Ärzte zu den Häufigkeitsangaben von Nebenwirkungen (Umsetzung der SPC-Richtlinie zu Beipackzetteln von Medikamenten) anhand fünf vorgegebener Wahrscheinlichkeitsquoten überprüft. Hierbei zeigten sich in den Ergebnisse eine Vielzahl an Antwortmöglichkeiten sowie eine große Unsicherheit in der „korrekten“ (im Sinne des Beipackzettels) Zuordnung von Begriffen zu den Quoten.

Aufgrund dieser deutlichen Ergebnisse sollte die Forschung auf dem Gebiet der Wahrscheinlichkeitsbegriffe und ihrer Darstellung in Worten, Zahlen oder beispielsweise auch Graphiken weiter vorangetrieben werden!

6. Literaturverzeichnis

1. Dolan JG, Iadarola S: Risk communication formats for low probability events: An exploratory study of patient preferences. BMC Med Inform Decis Mak 8, 114-20 (2008)

2. BfArM Medizinprodukte: Darstellung des aktuellen wissenschaftlichen Erkenntnisstandes über die Umsetzung der SPC-Guideline "Häufigkeitsangaben zu Nebenwirkungen" (2002), http://www.bfarm.de (Tag des Zugriffs: 11.03.2010)

3. Simpson RH: Stability in Meanings for Quantitative Terms: A Comparison over 29 Years. Quarterly Journal of Speech 49, 146-51, (1963)

4. Girone JA: Further elucidation of medical jargon. N Engl J Med 286, 382 (1972)

5. Nakao MA, Axelrod S: Numbers are better than words: Verbal specifications of frequency have no place in medicine. Am J Med 74, 1061-65 (1983)

6. Damrosch SP, Soeken K: Communicating probability in clinical reports: Nurses' numerical associations to verbal expressions. Res Nurs Health 6, 85-87 (1983)

7. Kong A, Barnett GO, Mosteller F, Youtz C: How medical professionals evaluate expressions of probability. N Engl J Med 315, 740-44 (1986)

8. O'Brien BJ: Words or numbers? The evaluation of probability expressions in general practice. J R Coll Gen Pract 39, 98-100 (1989)

9. Reagan RT, Mosteller F, Youtz C: Quantitative meanings of verbal probability expressions. J Appl Psychol 74, 433-42 (1989)

10. Shaw NJ, Dear PR: How do parents of babies interpret qualitative expressions of probability? Arch Dis Child 65, 520-23 (1990)

11. Mazur DJ, Hickam DH: Patients' interpretations of probability terms. J Gen Intern Med 6, 237-40 (1991)

12. Mazur DJ, Merz JF: Patients' interpretations of verbal expressions of probability: implications for securing informed consent to medical interventions. Behav Sci Law 12, 417-26 (1994)

13. Chesher T: Communicating with patients: Do your patients understand you? Are you speaking the same language? Med J Aust 142, 584-85 (1985)

14. Mazur DJ, Hickam DH: Patients' preferences for risk disclosure and role in decision making for invasive medical procedures. J Gen Intern Med 12, 114-17 (1997)

15. Edwards A, Prior L: Communication about risk: Dilemmas for general practitioners. The Department of General Practice Working Group, University of Wales College of Medicine. Br J Gen Pract 47, 739-42 (1997)

16. Ferguson WJ, Candib LM: Culture, language, and the doctor-patient relationship. Fam Med 34, 353-61 (2002)

17. Pepper S, Prytulak LS: Sometimes Frequently Means Seldom: Context Effects in the Interpretation of Quantitative Expressions. Journal of Research in Personality 8, 95-101 (1974)

18. Hammerton M: How much is a large part? Appl Ergon 7, 10-12 (1976)

19. Marteau TM, Saidi G, Goodburn S, Lawton J, Michie S, Bobrow M: Numbers or words? A randomized controlled trial of presenting screen negative results to pregnant women. Prenat Diagn 20, 714-18 (2000)

20. Daniels TW: Communicating risk: We as doctors are not alone. BMJ 327, 1403 (2003)

21. Karelitz TM, Budescu DV: You say "probable" and I say "likely": Improving interpersonal communication with verbal probability phrases. J Exp Psychol Appl 10, 25-41 (2004)

22. Rohrmann B: Empirische Studien zur Entwicklung von Antwortskalen für die Sozialwissenschaftliche Forschung. Zeitschrift für Sozialpsychologie 9, 222-45 (1978)

23. Misselbrook D, Armstrong D: Thinking about risk. Can doctors and patients talk the same language? Fam Pract 19, 1-2 (2002)

24. Edwards A, Elwyn G, Mulley A: Explaining risks: Turning numerical data into meaningful pictures. BMJ 324, 827-30 (2002)

25. EU-Guideline: A guideline of Summery of Product Characteristics (1999), http://mednet3.who.int/prequal/ (Tag des Zugriffs: 23.03.2009)

26. Paling J: Strategies to help patients understand risks. BMJ 327, 745-48 (2003)

27. Berry DC, Raynor DK, Knapp P, Bersellini E: Patients' understanding of risk associated with medication use: Impact of European Commission guidelines and other risk scales. Drug Saf 26, 1-11 (2003)

28. Burkiewicz JS, Vesta KS, Hume AL: Improving effectiveness in communicating risk to patients. Consult Pharm 23, 37-43 (2008)

29. Moore RA, Derry S, McQuay HJ, 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 Res Ther 10, R20 (2008)

30. Epstein RM, Alper BS, Quill TE: Communicating evidence for participatory decision making. JAMA 291, 2359-66 (2004)

31. Timmermans D, Molewijk B, Stiggelbout A, Kievit J: Different formats for communicating surgical risks to patients and the effect on choice of treatment. Patient Educ Couns 54, 255-63 (2004)

32. Timmermans DR, Ockhuysen-Vermey CF, Henneman L: Presenting health risk information in different formats: The effect on participants' cognitive and emotional evaluation and decisions. Patient Educ Couns 73, 443-47 (2008)

33. Lipkus IM: Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. Med Decis Making 27, 696-713 (2007)

34. Visschers VH, Meertens RM, Passchier WW, de Vries NN: Probability information in risk communication: A review of the research literature. Risk Anal 29, 267-87 (2009)

35. Czarnecki A: Intelligent risk communication: Can it be improved? Drug Saf 31, 1-6 (2008)

36. Sobel R: Language and the physician. Public Health Rev 31, 27-32 (2003)

37. Böcker FH: Marketing. 7. Auflage, S. 235ff., Lucius & Lucius Verlagsgesellschaft mbH, Stuttgart (2003)

38. Porst R: Ausschöpfungen bei sozialwissenschaftlichen Umfragen: Die Sicht der Institute. ZUMA-Arbeitsbericht (1996)

39. Suß J: Ist die Rangordnung von umgangssprachlichen Wahrscheinlichkeitsausdrücken in der umgangssprachlichen Kommunikation unter Ärzten konsistent? Med. Diss. Fachbereich Humanmedizin, Philipps-Universität Marburg (1997)

40. Müller U: Umgangssprachliche Kommunikation unter Ärzten. Medizinisches Zentrum für Methodenwissenschaften und Gesundheitsforschung, Philipps-Universität Marburg (1994)

41. Barbour S, Stevenson P: Variation im Deutschen: Soziolinguistische Perspektiven. S. 10ff, 45, 88, 86, Walter de Gruyter GmbH & Co Berlin (1998)

42. König W: dtv-Atlas Deutsche Sprache. 13. Auflage, S. 77, 102f, 135ff, 233, Deutscher Taschenbuchverlag GmbH & Co. KG München (2001)

43. Schleswig-Holstein Arztfindex (2004), http://www.arztfindex.de (Tag des Zugriffs: 02.03.2004)

44. Barth S: Die schriftliche Befragung (1998), http://www.stephan-barth.de (Tag des Zugriffs: 23.03.2005)

45. Diekmann A: Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen. 11. Auflage, S. 439ff, Rowohlt Verlag Reinbek bei Hamburg (2004)

46. Siegel S: Nichtparametrische statistische Methoden. 4. Auflage, S. 217ff, 234f, 270f, Dietmar Klotz GmbH Frankfurt am Main (1997)

47. Kendall MG, Gibbons JD: Rank Correlation Methods. 5. ed., S. 117ff, 121ff, Oxford University New York (1990)

48. Efron B: The Jackknife, the Bootstrap and Other Resampling Plans. 6. ed., S. 29ff, Society for Industrial and Applied Mathematics Philadelphia, CCP Montpelier, Vermont (1994)

49. Efron B, Tibshirani R. J.: An Introduction to the Bootstrap. S. 45ff, 60ff, Chapman & Hall/CRC, New York (1998)

50. Bamberg GB, F: Statistik. Oldenbourgs Lehr- und Handbücher der Wirtschafts- und Sozialwissenschaften. 4. Auflage, S. 39f, Oldenbourg Verlag München (2001)

51. Sachs LH, J: Angewandte Statistik. 13. Auflage, S. 104f, 646, Springer Verlag, Berlin (2009)

52. Bortz JL, Lienert GA, Boehnke K: Verteilungsfreie Methoden in der Biostatistik. 2. Auflage, S. 423ff, Springer Verlag, Berlin (2000)

53. Diaconis P: Group Representations in Probability and Statistics. Lecture Notes-Monograph Series. Gupta S S, Series Editor, Vol. 11, S. 11ff, Institiute of Mathematical Statistics Hayward, Harvard (1988)

54. Zöfel P: Ein Begleitbuch zur computergestützten Anwendung. Eine Datenmenge - was tun? Ein Leitfaden zu statistischen Lösungen. S. 103ff, Addison-Wesley Verlag, München (2001)

55. Kruse MS, F: BASICS Allgemeine Pharmakologie. S. 36, Elsevier, Urban & Fischer, München (2009)

56. Porst R: Praxis der Umfrageforschung. Studienscripte zur Soziologie. S. 91ff, Teubner Studienscripte (1985)

57. Koschnik WJ: Werbeplanung, Mediaplanung, Marktforschung, Kommunikationsforschung, Mediaforschung. Focus Lexikon (2005), http://medialine.@focus.de, (Tag des Zugriffs: 17.03.2005)

58. Dillmann D: Mail and Telephone Surveys. The Total Design Method. S. 360, Wiley, New York (1983)

59. Dolezal J: Translating Medical Idiom. New England Journal of Medicine 291, 176, (1976)

60. Bausinger H: Deutsch für Deutsche. Dialekte, Sprachbarrieren, Sondersprachen. S. 24, (c) Fischer Taschenbuch Verlag GmbH, Frankfurt am Main (1984)

7. Abbildungsverzeichnis

Abbildung 1: Rücklauf 1. Welle – Schleswig-Holstein und Bayern 21
Abbildung 2: Rücklauf 2. Welle – Schleswig-Holstein und Bayern 22
Abbildung 3: Gegenüberstellung der Rückläufe 24
Abbildung 4: Vergleich der Mediane aus Abschnitt A und B 32
Abbildung 5: Vergleich der Mediane aus Abschnitt A und D 33

8. Tabellenverzeichnis

Tabelle 1: Die ideale Rangreihe der 20 Wahrscheinlichkeitsbegriffe 14
Tabelle 2: SPC-Richtlinie „Häufigkeitsangaben von Nebenwirkungen“ 19
Tabelle 3: Klassifikation der Nebenwirkungen 20
Tabelle 4: Übersicht Gesamtausschöpfung 1. Welle 22
Tabelle 5: Übersicht Gesamtausschöpfung 2. Welle 23
Tabelle 6: Übersicht Gesamtausschöpfung der beiden Wellen 23
Tabelle 7: Medizinische Fachrichtung der Teilnehmer unterteilt nach Bundesland 25
Tabelle 8: Geschlecht der Befragten unterteilt nach Bundesland 25
Tabelle 9: Geschlecht der Befragten unterteilt nach Fachrichtung 25
Tabelle 10: Alter unterteilt nach Bundesland 26
Tabelle 11: Ausgewählte Lage- und Streuungsmaße für die Ärzte in Schleswig-Holstein und Bayern 27
Tabelle 12: Echte Rangreihen nach Bundesland 28
Tabelle 13: Ergebnisse aus dem Vergleich der Rangreihenfolgen 28
Tabelle 14: Gegenüberstellung der idealen Rangreihe und der beobachteten Rangreihen der Ärzte Schleswig-Holsteins (SH) und Bayerns (BA) 30
Tabelle 15: Rangkorrelation nach Bundesland 31
Tabelle 16: Ergebnisse des Permutationstest 31
Tabelle 17: Ergebnisse aus der Berechnung des Median (Liste und Kontext B) 32
Tabelle 18: Ergebnisse aus der Berechnung des Median (Liste und Kontext D) 33
Tabelle 19: Ergebnisse des Wilcoxon-X-Tests für die Mediziner unterteilt nach Bundesland 34
Tabelle 20: Ergebnisse des Wilcoxon-X-Tests für die Mediziner unterteilt nach Bundesland 35
Tabelle 21: Mann-Whitney-Test / Bayern versus Schleswig-Holstein 36
Tabelle 22: Differenzvariable – Bayern vs. Schleswig-Holstein / Mann-Whitney-Test 36
Tabelle 23: Mann-Whitney-Test / Bayern versus Schleswig-Holstein 37
Tabelle 24: Differenzvariable – Bayern vs. Schleswig-Holstein / Mann-Whitney-Test 37
Tabelle 25: Am häufigsten genannter Begriff zur vorgegebenen Quote der Ärzte Schleswig-Holsteins 38
Tabelle 26: Am häufigsten genannter Begriff zur vorgegebenen Quote der Ärzte Bayerns 38
Tabelle 27: Zuordnung der verbalen Wahrscheinlichkeitsbegriffe zu den numerischen Wahrscheinlichkeiten 39
Tabelle 28: Korrekt zugeordneter Begriff zur vorgegebenen Quote Ärzte Schleswig-Holstein und Ärzte Bayern 39
Tabelle 29: Korrekte und inkorrekte Antworten der Studienteilnehmer 40

9. Anhang

9.1 Material zum Vertiefen des Themas

9.1.1 Anschreiben

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 Ärztekammern 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: konrad@imbs.uni-luebeck.de; Tel.: 0451-5002781

Ich danke Ihnen sehr im Voraus für Ihre Teilnahme!

9.1.2 Fragebogen

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!

Montagmorgen, 7.30 Uhr, eine Arztpraxis irgendwo in Schleswig-Holstein...
Sie sind der Hausarzt von Herrn Meier, 58 Jahre, von Beruf Mathemathiklehrer. 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 den Patienten 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 Ein-vernehmen 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?

_____

In welchem Bundesland haben Sie überwiegend studiert?

_____

In welchem Bundesland haben Sie Ihren Abschluss gemacht?

_____

In welchem Bundesland arbeiten Sie zur Zeit?

_____

Nochmals vielen Dank für Ihre Teilnahme!

9.1.3 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.

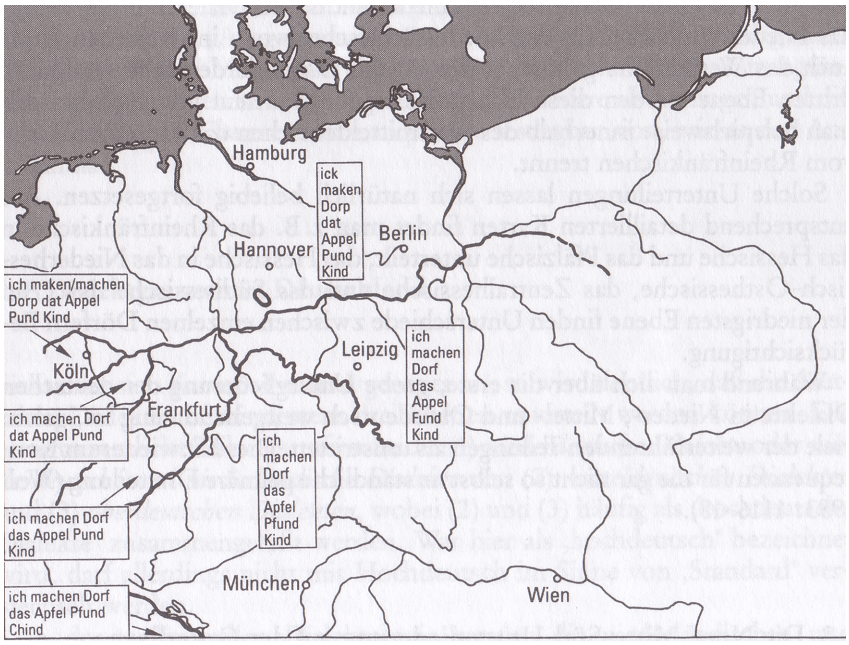
Sollten Sie den Fragebogen bereits ausgefüllt und an mich zurückgeschickt haben, bedanke ich mich sehr für Ihre Teilnahme.
Falls nicht, möchte ich Sie hiermit nochmals daran erinnern. Nehmen Sie sich bitte diese Zeit – es dauert nur ca. 5 Minuten! Bedenken Sie, dass jede Rückantwort die Aussagekraft 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: konrad@imbs.uni-luebeck.de; Tel.: 0451-5002781

Ich danke Ihnen nochmals sehr für Ihre Teilnahme!

9.1.4 Was bedeutet die Zweite Lautverschiebung?

Regionaltypische mundartliche Formen gebrauchter Wörter zur Realisierung der Zweiten (hochdeutschen) Lautverschiebung;



entnommen aus (41), mit freundlicher Genehmigung des Fischer Taschenbuch Verlags GmbH © (60)

9.1.5 SPSS-Syntax zum Bootstrap

Syntax zum Bootstrap am Beispiel der Daten der Ärzte aus Schleswig-Holstein (n=180); in Anlehnung an *Suß* (39)

```
GET FILE = 'c:\matrixshneu.sav'.
      COMPUTE ID=$CASENUM.
      SAVE outfile = 'c:\bootshneu.sav'.
      INPUT PROGRAM.
      LOOP SAMP = 1 to 1000.
      LOOP V = 1 to 180.
COMPUTE ID = TRUNC(UNIFORM(180)) + 1.
      END CASE.
      LEAVE SAMP.
      END LOOP.
      END LOOP.
      END FILE.
      END INPUT PROGRAM.
      SORT CASES BY ID.
      MATCH FILES /FILE* /TABLE 'c:\bootshneu.sav' /BY ID.
      SORT CASES BY SAMP.
      SPLIT FILE BY SAMP.

NPAR TESTS /KENDALL=sel niem si unsi oft moe gel ueb mei den hau unkl nor
flos ste nie wahr im frag typ/MISSING LISTWISE.
```

9.1.6 Rücklauf

Bruttorücklauf

Im Bruttorücklauf sind zufällige und systematische Ausfälle mit eingerechnet

Brutto 1.Rücklauf					Brutto 2.Rücklauf				Bruttorücklauf insgesamt			
	insgesamt raus		**insgesamt zurück**		**insgesamt raus**		**insgesamt zurück**		**insgesamt raus**		**insgesamt zurück**	
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Ärzte insg.	**300**	**100**	**175**	**58,33**	**300**	**100**	**186**	**62,00**	**600**	**100**	**361**	**60,17**
INSH	75	100	47	62,67	98	100	64	65,31	173	100	111	64,16
ANSH	75	100	43	57,33	52	100	32	61,54	127	100	75	59,06
Ärzte SH	150	100	90	60,00	150	100	96	64,00	300	100	186	62,00
INBA	75	100	40	53,33	75	100	46	61,33	150	100	86	57,33
ANBA	75	100	45	60,00	75	100	44	58,67	150	100	89	59,33
Ärzte BA	150	100	85	56,67	150	100	90	60,00	300	100	175	58,33

- Stichprobenneutrale Ausfälle = unbekannt verzogen; das waren insgesamt 6 (ANSH1=1,INBA1=3, ANBA1=1, INBA2=1)
- Systematischen Ausfälle; das waren insgesamt 5 (INSH2=1Rente, ANSH2=2verweigert, INBA2=1v, ANBA2=1R)

Bereinigte Stichprobe (= Brutto - neutrale Ausfälle)

Rücklauf 1. Durchgang					Rücklauf 2. Durchgang				Rücklauf insgesamt			
	insgesamt raus		**insgesamt zurück**		**insgesamt raus**		**insgesamt zurück**		**insgesamt raus**		**insgesamt zurück**	
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Ärzte insg.	**295**	**100**	**170**	**57,63**	**299**	**100**	**185**	**61,87**	**594**	**100**	**355**	**59,76**
INSH	75	100	47	62,67	98	100	64	65,31	173	100	111	64,16
ANSH	74	100	42	56,76	52	100	32	61,54	126	100	74	58,73
Ärzte SH	149	100	89	59,73	150	100	96	64,00	299	100	185	61,87
INBA	72	100	37	51,39	74	100	45	60,81	146	100	82	56,16
ANBA	74	100	44	59,46	75	100	44	58,67	149	100	88	59,06
Ärzte BA	146	100	81	55,48	149	100	89	59,73	295	100	170	57,63

Ausschöpfung/Nettostichprobe (= bereinigte Stichprobe - systematische Ausfälle)

Rücklauf 1. Durchgang					Rücklauf 2. Durchgang				Rücklauf insgesamt			
	insgesamt raus		**insgesamt zurück**		**insgesamt raus**		**insgesamt zurück**		**insgesamt raus**		**insgesamt zurück**	
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Ärzte insg.	**295**	**100**	**170**	**57,63**	**294**	**100**	**180**	**61,22**	**589**	**100**	**350**	**59,42**
INSH	75	100	47	62,67	97	100	63	64,95	172	100	110	63,95
ANSH	74	100	42	56,76	50	100	30	60,00	124	100	72	58,06
Ärzte SH	149	100	89	59,73	147	100	93	63,26	296	100	182	61,49
INBA	72	100	37	51,39	73	100	44	60,27	145	100	81	55,86
ANBA	74	100	44	59,46	74	100	43	58,12	148	100	87	58,78
Ärzte BA	146	100	81	55,48	147	100	87	59,18	293	100	168	57,34

Auschöpfungsquote (= Ausschöpfung,zurück ⁄ bereinigte Stichprobe,raus x 100)

Rücklauf 1. Durchgang				Rücklauf 2. Durchgang			Rücklauf insgesamt		
	Aus-schöpfung	Bereinigte Stichprobe	Quote (%)	Aus-schöpfung	Bereinigte Stichprobe	Quote (%)	Aus-schöpfung	Bereinigte Stichprobe	Quote (%)
Ärzte insg.	**170**	**295**	**57,63**	**180**	**299**	**60,20**	**350**	**594**	**58,92**
INSH	47	75	62,67	63	98	64,29	110	173	63,58
ANSH	42	74	56,76	30	52	57,69	72	126	57,14
Ärzte SH	89	149	59,73	93	150	62,00	182	299	60,87
INBA	37	72	51,39	44	74	59,46	81	146	55,48
ANBA	44	74	59,46	43	75	57,33	87	149	58,39
Ärzte BA	81	146	55,48	87	149	58,39	168	295	56,95

Übersicht Erinnerungsschreiben (bereinigte Stichprobe)

1. Durchgang

	insgesamt raus		**insgesamt zurück**		**ohne Erinnerung zurück**		**zur Erinnerung raus**		**nach Erinnerung zurück**	
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Ärzte insg.	**300**	**100**	**170**	**56,67**	**121**	**40,33**	**179**	**100**	**49**	**27,37**
INSH	75	100	47	62,67	39	52,00	36	100	8	22,22
ANSH	75	100	42	56,00	29	38,67	46	100	13	28,26
Ärzte SH	150	100	89	59,33	68	45,33	82	100	21	25,61
INBA	75	100	37	49,33	25	33,33	50	100	12	24,00
ANBA	75	100	44	58,67	28	37,33	47	100	16	34,04
Ärzte BA	150	100	81	54,00	53	35,33	97	100	28	28,87

2. Durchgang

	insgesamt raus		**insgesamt zurück**		**ohne Erinnerung zurück**		**zur Erinnerung raus**		**nach Erinnerung zurück**	
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Ärzte insg.	**300**	**100**	**185**	**61,67**	**140**	**46,67**	**160**	**100**	**45**	**28,13**
INSH	98	100	64	65,31	46	46,94	52	100	18	34,61
ANSH	52	100	32	61,54	24	46,15	28	100	8	28,57
Ärzte SH	150	100	96	64,00	70	46,67	80	100	26	32,50
INBA	75	100	45	60,00	33	44,00	42	100	12	28,57
ANBA	75	100	44	58,67	37	49,33	38	100	7	18,42
Ärzte BA	150	100	89	59,33	70	46,67	80	100	19	23,75

Insgesamt										
	insgesamt raus		**insgesamt zurück**		**ohne Erinnerung zurück**		**zur Erinnerung raus**		**nach Erinnerung zurück**	
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Ärzte insg.	**600**	**100**	**355**	**59,17**	**261**	**43,50**	**339**	**100**	**94**	**27,73**
INSH	173	100	111	64,16	85	49,13	88	100	26	29,54
ANSH	127	100	74	58,27	53	41,73	74	100	21	28,38
Ärzte SH	300	100	185	61,67	138	46,00	162	100	47	29,01
INBA	150	100	82	54,67	58	38,67	92	100	24	26,09
ANBA	150	100	88	58,67	65	43,33	85	100	23	27,06
Ärzte BA	300	100	170	56,67	123	41,00	177	100	47	26,55

9.1.7 Übersicht Lage- und Streuungsmaße für die 20 Wahrscheinlichkeitsbegriffe

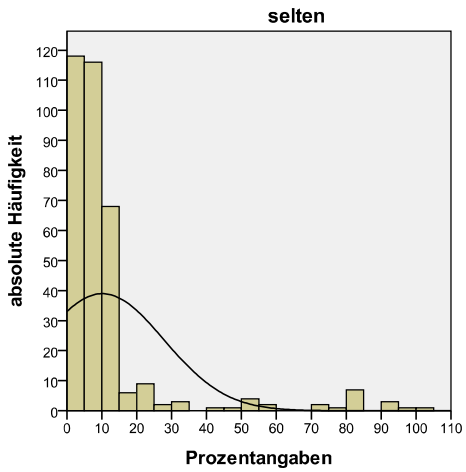
Lage- und Streuungsmaße für die Ärzte aus Schleswig-Holstein (n = 180)

	n = Anzahl der Befragten	Mittelwert	Standardabweichung	Modalwert	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
selten	177	10.30	18.80	5.00	5.00	0.05	100.00	2.50	10.00
niemals	176	7.16	24.37	0.00	0.00	0.00	100.00	0.00	0.00
sicher	173	92.65	17.70	100.00	99.00	0.00	100.00	95.00	100.00
unsicher	166	39.74	21.56	50.00	50.00	0.00	100.00	23.75	50.00
oft	176	67.57	18.34	80.00	70.00	10.00	100.00	60.00	80.00
möglicherweise	170	34.34	19.70	50.00	30.00	0.55	100.00	20.00	50.00
gelegentlich	176	19.98	15.16	20.00	20.00	1.00	90.00	10.00	25.00
überwiegend	175	74.37	12.63	70.00	75.00	20.00	100.00	70.00	80.00
meist	176	81.29	12.24	90.00	80.00	20.00	99.00	75.00	90.00
denkbar	171	24.07	19.98	10.00	20.00	0.00	90.00	10.00	40.00
häufig	176	70.22	18.30	80.00	75.00	5.00	100.00	61.25	80.00
unklar	147	33.90	21.81	50.00	40.00	0.00	100.00	15.00	50.00
normalerweise	176	82.29	14.79	90.00	90.00	10.00	100.00	80.00	90.00
fraglos	160	81.39	29.95	100.00	95.00	0.00	100.00	80.00	100.00
stets	176	89.81	23.21	100.00	99.00	0.00	100.00	90.00	100.00
nie	176	10.87	29.21	0.00	0.00	0.00	100.00	0.00	0.10
wahrscheinlich	177	75.20	17.47	80.00	80.00	5.00	98.00	70.00	90.00
immer	176	93.41	20.95	100.00	100.00	0.00	100.00	99.00	100.00
fraglich	157	29.04	20.41	10.00	20.00	0.00	80.00	10.00	50.00
typischerweise	177	78.49	17.95	80.00	80.00	0.00	100.00	75.00	90.00

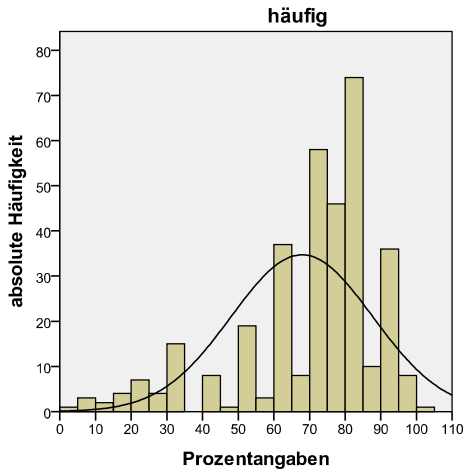
Lage- und Streuungsmaße für die Ärzte aus Bayern (n = 168)

	n = Anzahl der Befragten	Mittelwert	Standardabweichung	Modalwert	Median	Minimum	Maximum	25%-Quantil	75%-Quantil
selten	166	10.03	16.53	5.00	5.00	0.001	80.00	2.00	10.00
niemals	164	6.67	23.19	0.00	0.00	0.00	100.00	0.00	0.00
sicher	167	92.47	18.53	100.00	99.50	5.00	100.00	95.00	100.00
unsicher	163	35.99	21.38	50.00	40.00	0.00	100.00	20.00	50.00
oft	167	65.62	20.21	80.00	70.00	1.00	95.00	54.00	80.00
möglicherweise	163	31.73	19.35	30.00	30.00	0.10	90.00	15.00	50.00
gelegentlich	166	17.46	15.71	10.00	10.00	0.10	90.00	10.00	20.00
überwiegend	166	72.65	15.94	80.00	75.00	5.00	95.00	66.00	80.00
meist	167	80.37	14.91	90.00	80.00	10.00	100.00	75.00	90.00
denkbar	163	22.02	20.32	10.00	10.00	0.005	80.00	5.00	35.00
häufig	167	65.32	21.13	80.00	70.00	1.00	95.00	50.00	80.00
unklar	136	35.37	21.40	50.00	40.00	0.00	100.00	16.25	50.00
normalerweise	163	82.03	13.40	90.00	85.00	10.00	100.00	75.00	90.00
fraglos	155	85.50	26.35	100.00	95.00	0.00	100.00	90.00	100.00
stets	167	90.59	22.02	100.00	99.80	0.00	100.00	95.00	100.00
nie	164	13.34	32.57	0.00	0.00	0.00	100.00	0.00	0.00
wahrscheinlich	166	72.86	17.44	70.00	75.00	5.00	99.00	66.00	85.00
immer	167	93.76	19.45	100.00	100.00	0.00	100.00	99.00	100.00
fraglich	154	31.26	19.04	50.00	30.00	0.00	100.00	11.88	50.00
typischerweise	162	76.86	18.44	80.00	80.00	0.00	100.00	70.00	90.00

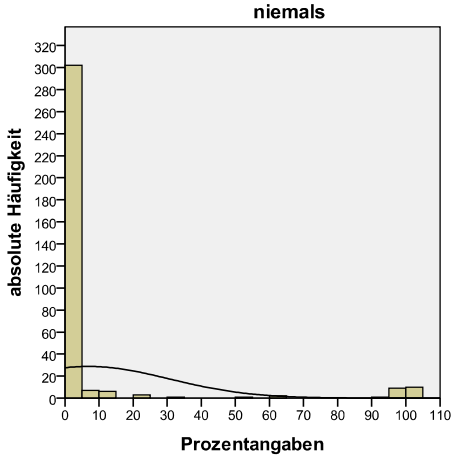
Darstellung in Form von Histogrammen (Ärzte insgesamt)



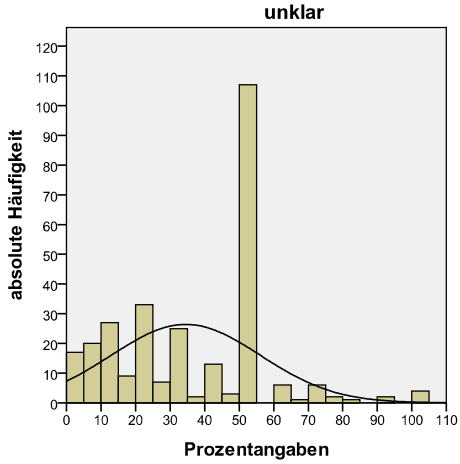
Begriff „selten“: N=345 Angaben, N=5 Angaben fehlend



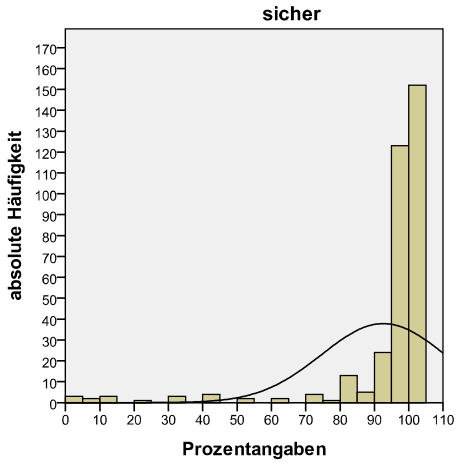
Begriff „häufig“: N=345 Angaben; n=5 Angaben fehlend



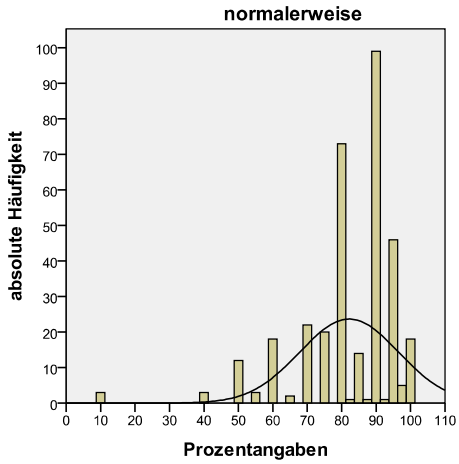
Begriff „niemals“: N=342 Angaben, N=8 Angaben fehlend



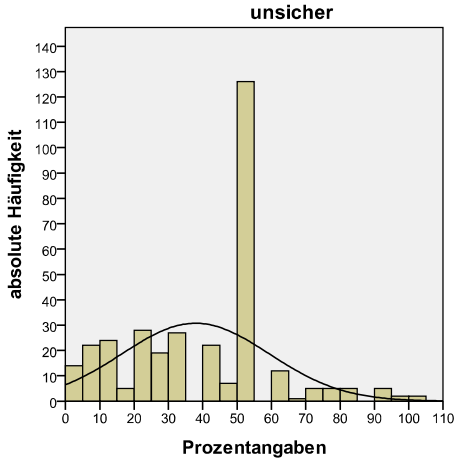
Begriff „unklar“: N=285 Angaben, N=65 Angaben fehlend



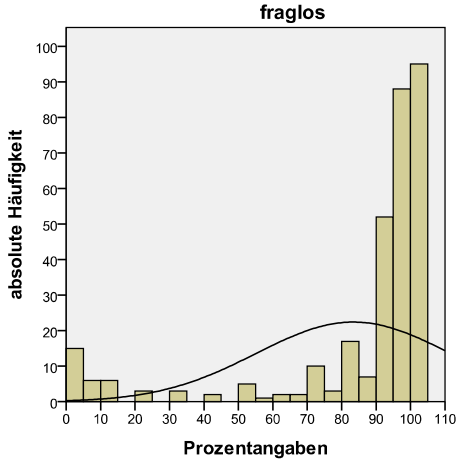
Begriff sicher“: N=342 Angaben; N=8 Angaben fehlend



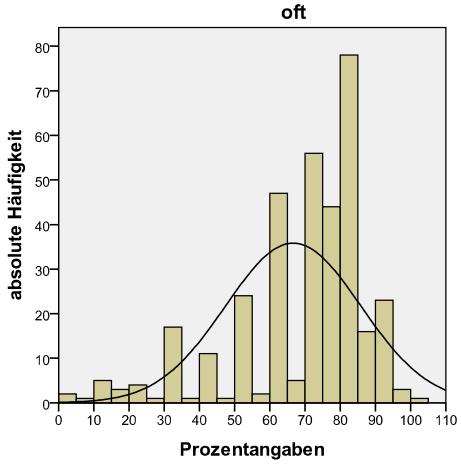
Begriff „normalerweise“: N=341 Angaben; N=9 Angaben fehlend



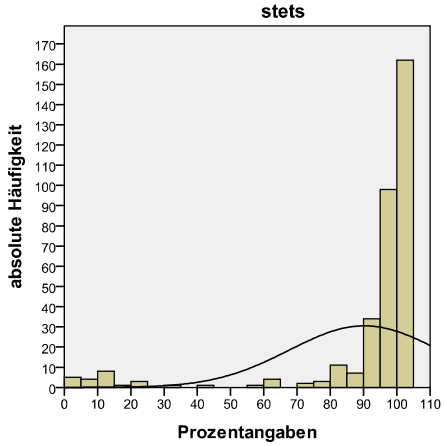
Begriff „unsicher“: N=331 Angaben; N=19 Angaben fehlend



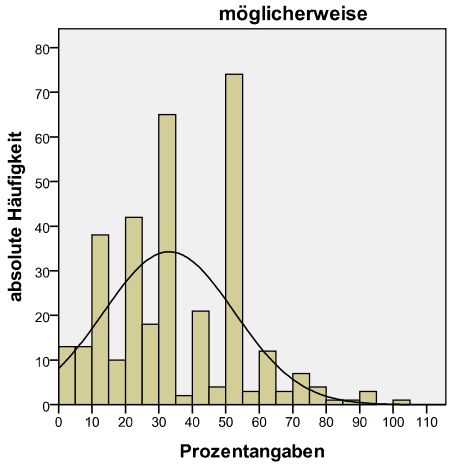
Begriff „fraglos“: N=317 Angaben; n=13 Angaben fehlend



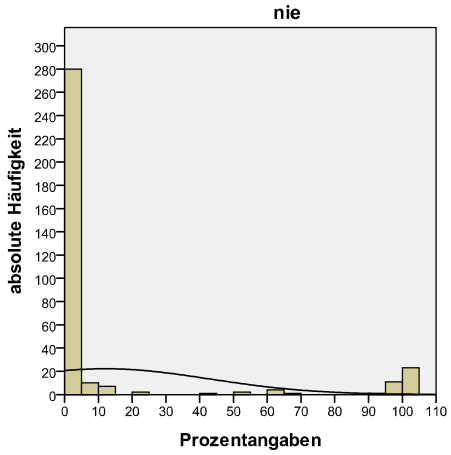
Begriff „oft“: N=345 Angaben; N=5 Angaben fehlend



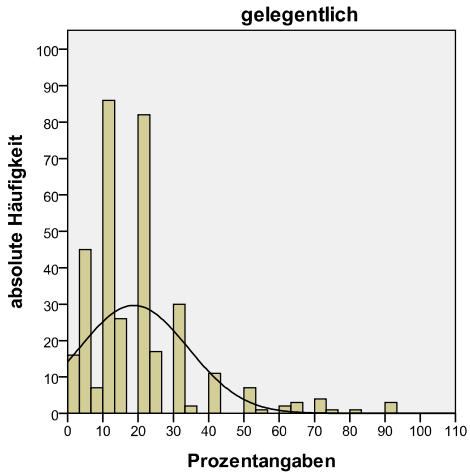
Begriff „stets“: N=345 Angaben; n=5 Angaben fehlend



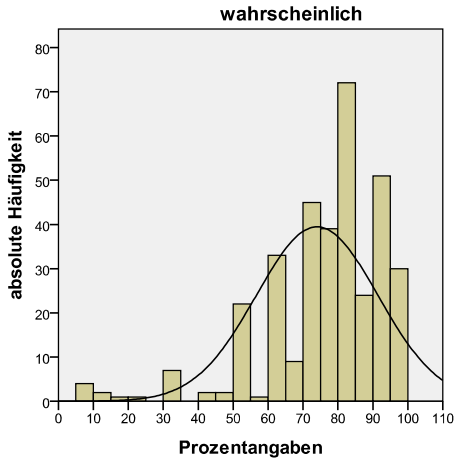
Begriff „möglicherweise“: N=335 Angaben; N=15 Angaben fehlend



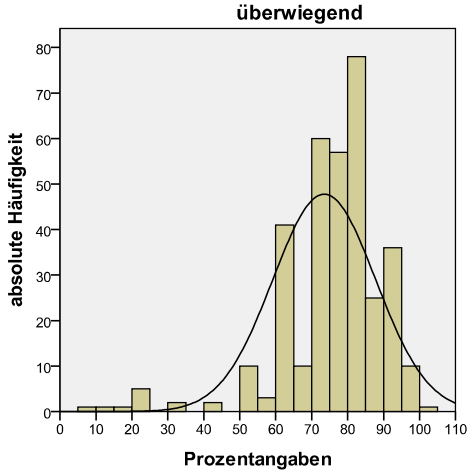
Begriff „nie“: N=342 Angaben; n=8 Angaben fehlend



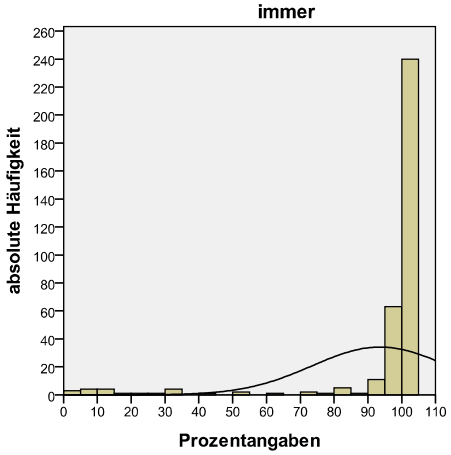
Begriff „gelegentlich“: N=344 Angaben; N=6 Angaben fehlend



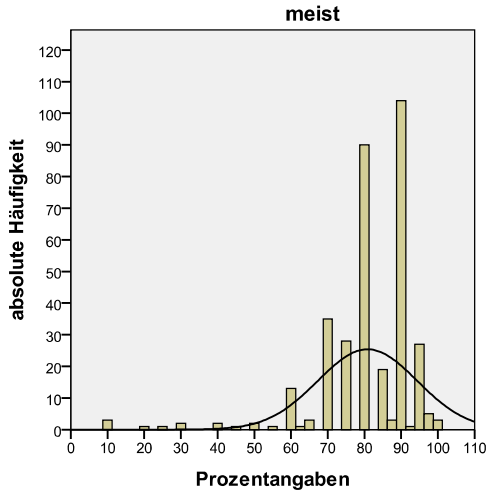
Begriff „wahrscheinlich“: N=345 Angaben; n=5 Angaben fehlend



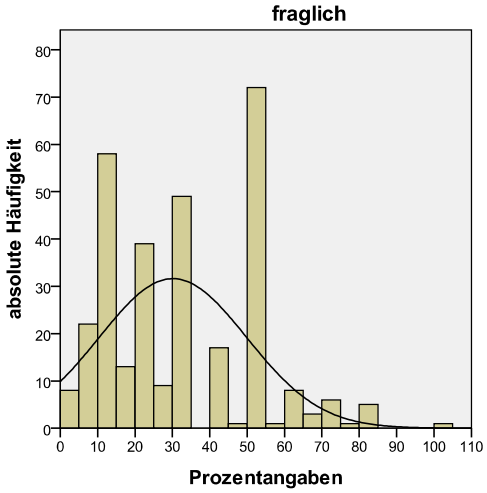
Begriff „überwiegend“: N=343 Angaben; n=7 Angaben fehlend



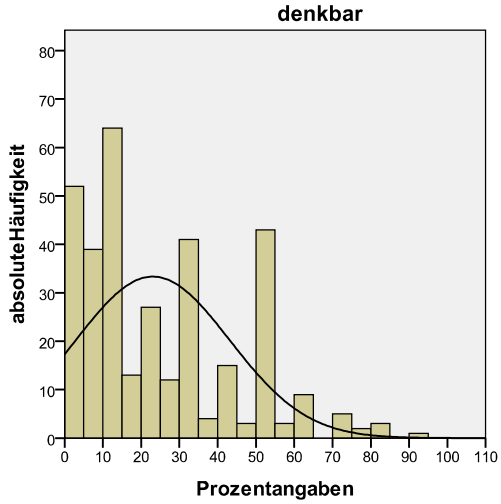
Begriff „immer“: N=345 Angaben; N=5 An Angaben fehlend



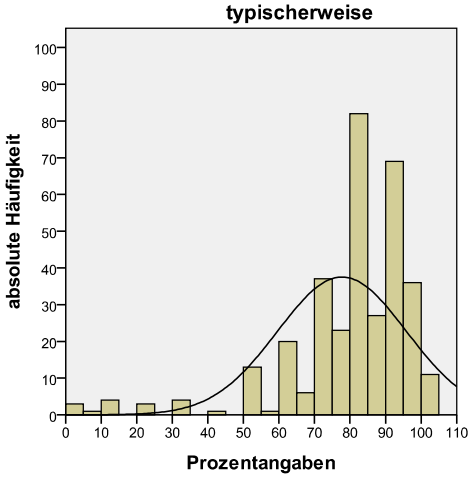
Begriff „meist“: N=345 Angaben; N=5 Angaben fehlend



Begriff „fraglich“: N=313 Angaben; n=17 Angaben fehlend



Begriff „denkbar“: N=336 Angaben; N=14 Angaben fehlend



Begriff „typischerweise“: N=341 Angaben; N=9 Angaben gemacht

9.1.8 Angaben zu Wahrscheinlichkeitsquoten

Wahrscheinlichkeit 1:1000

Ärzte Schleswig-Holstein

1:1000		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
sehr selten	62	36.5%
selten	62	36.5%
extrem selten	10	5.9%

Weitere Nennungen: 4-mal: „gelegentlich“, „fast nie“, „einer von 1000“, „äußerst selten“. Drei Nennungen: „unwahrscheinlich“. Zwei Nennungen: „ganz selten“, „denkbar“, „kaum“. Eine Nennung: „“oft“, „möglicherweise“, „vereinzelt“, „nicht selten“, „sehr gering“, „fraglich“, „nicht so häufig“, „regelmäßig“, „1 Promille“, „Promillbereich“, „so gut wie nie“.

Ärzte Bayern

1:1000		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
sehr selten	64	40.5%
selten	50	31.6%
extrem selten	9	5.7%
gelegentlich	5	3.2%
äußerst selten	5	3.2%

Vier Nennungen entfielen auf die Begriffe „fast nie“ und „unwahrscheinlich“, zwei Nennungen auf „häufig“, „manchmal“, „einer von 1000“ und „Promillbereich“. Eine Nennung gab es zu den verbalen Wahrscheinlichkeiten „ganz selten“, „möglicherweise“, „denkbar“, „ausgesprochen selten“, „nur in Einzelfällen“, „kaum“, „sehr geringe Wahrscheinlichkeit, „ zu häufig“ und in „in Einzelfällen“.

Wahrscheinlichkeit 1:100

Ärzte Schleswig-Holstein

1:100		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
selten	76	44.7%
gelegentlich	42	24.7%
sehr selten	9	5.3%
häufig	7	4.1%
einer von 100	5	2.9%

Viermal wurden genannt: „manchmal“ und „oft“. Zwei Nennungen: „möglicherweise“, „vereinzelt“, „häufiger“, „kommt vor“. Eine Nennung: „sehr oft“, „sehr sehr selten“, „öfter“, „wahrscheinlich“, „wenig wahrscheinlich“, „gehäuft“, „unwahrscheinlich“, „typischerweise“, „nicht auszuschließen“, „relevant“, „ab und zu“, „regelmäßig“, „“1%“, „Prozentbereich“ und „“eher selten“.

Ärzte Bayern

1:100		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
selten	73	45.9%
gelegentlich	35	22.0%
sehr selten	7	4.4%
manchmal	7	4.4%
möglicherweise	6	3.8%

Weitere Nennungen: 4-mal: „oft“. 3-mal: „häufig“. Zwei Nennungen: „möglich“, „einer von 100“, „denkbar“, „kommt vor“, „geringe Wahrscheinlichkeit“, „Prozentbereich“. Eine Nennung: „sehr häufig“, „ganz selten“, „wahrscheinlich“, „vereinzelt“, „häufiger“, „immer“, „ab und zu“, „kaum“, „eher selten“, „gewisses Risiko“, „relativ selten“, „verhältnismäßig selten“.

Wahrscheinlichkeit 1:10

Ärzte Schleswig-Holstein

1:10		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
gelegentlich	55	32.4%
häufig	35	20.6%
oft	18	10.6%
möglicherweise	12	7.1%
manchmal	7	4.1%
selten	6	3.5%

Die Begriffe „sehr häufig“ und „jeder 10.“ wurden viermal genannt. Drei Nennungen entfielen auf die verbalen Wahrscheinlichkeiten „häufiger“ und „nicht selten“. „Öfters“, „möglich“ und „fraglich“ wurden jeweils zweimal genannt. Eine Nennung gab es zu den Begriffen „öfter“, „wenig wahrscheinlich“, „einer von 10“, „denkbar“, „gehäuft“, „überwiegend“, „ziemlich häufig“, „relativ häufig“, „zum Teil“, „mit hoher Wahrscheinlichkeit“, „sehr bedeutsam“, „ab und zu“, „“kann auftreten“, „10%“, „unter Umständen“, „nicht auszuschließen“ und „teilweise“.

Ärzte Bayern

1:10		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
gelegentlich	52	33.1%
häufig	34	21.7%
oft	10	6.4%
jeder 10.	8	5.1%
möglicherweise	7	4.5%
selten	6	3.8%
öfter	5	3.2%
manchmal	5	3.2%

Der Begriff „denkbar“ wurde viermal genannt. Drei Nennungen entfielen auf den verbalen Wahrscheinlichkeitsbegriff „wahrscheinlich“. Zwei Nennungen: „öfters“ und „nicht selten“. Eine Nennung: „sehr häufig“, „sehr oft“, „einer von 10“, „relativ häufig“, „kommt vor“, „regelmäßig“, „geringer Prozentsatz“, „immer wieder (mal)“, „hoher Prozentsatz“, „durchaus möglich“, „damit ist zu rechnen“, „hin und

wieder“, „kritisch“, „eher selten“, „nicht sehr häufig“, „m. Kompl. ist zu rechnen“, nicht ganz selten“, nahezu regelmäßig“ und „bisweilen“.

Wahrscheinlichkeit 1:5

Ärzte Schleswig-Holstein

1:5		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
häufig	52	30.8%
oft	27	16.0%
gelegentlich	16	9.5%
sehr häufig	7	4.1%
möglicherweise	7	4.1%
sehr oft	6	3.6%
jeder 5./ einer von 5	6	3.6%
öfters	5	3.0%
manchmal	5	3.0%
relativ häufig	5	3.0%

Drei Nennungen gab es zu den verbalen Wahrscheinlichkeitsbegriffen „öfter“ und „wahrscheinlich“. Zwei Probanden nannten jeweils die Begriffe „selten“, „recht häufig“, „sehr wahrscheinlich“, „überwiegend“ und „regelmäßig“. Einmal gab es Nennungen zu den Begriffen „denkbar“, „gehäuft“, „typischerweise“, „häufiger“, „fast immer“, „oft möglich“, „unvertretbar hoch“, „großes Risiko“, „extrem häufig“, „oftmals“, „kommt vor“, „mit großer Wahrscheinlichkeit“, „20%“, „immer wieder (mal)“, „viele“, „wiederholt“ und „zu oft“.

Ärzte Bayern

1:5		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
häufig	41	26.1%
oft	33	21.0%
sehr häufig	13	8.3%
jeder 5./ einer von 5	12	7.6%
gelegentlich	11	7.0%
möglicherweise	10	6.4%
öfters	6	3.8%
manchmal	5	3.2%

Der Begriff „wahrscheinlich“ wurde viermal genannt. „Sehr oft“ haben drei Studienteilnehmer der Wahrscheinlichkeit „1:5“ zugeordnet. Zwei Nennungen gab es zu den Begriffen „selten“, „denkbar“ und „gehäuft“. Einmal wurden „recht häufig“, „typischerweise“, „ziemlich oft“, „relativ häufig“, „fraglich“, „großes Risiko“, „regelmäßig“, „hoher Prozentsatz“, „praktisch nie“, „sehr kritisch“, „in relevanter Anzahl“, „etwas häufiger“ und „unsicher genannt.

Wahrscheinlichkeit 1:2

Ärzte Schleswig-Holstein

1:2		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
häufig	46	27.4%
sehr häufig	27	16.1%
oft	18	10.7%
jeder 2. / einer von 2	12	7.1%
sehr oft	10	6.0%
typischerweise	6	3.6%
fast immer	6	3.6%
meist	5	3.0%

Viermal wurde der Begriff „möglich“ genannt. Dreimal gab es jeweils Nennungen zu den Begriffen „wahrscheinlich“, „in der Hälfte der Fälle“, „(50:50)-Chance“ und „regelmäßig“. Zwei Nennungen entfielen jeweils auf „sehr wahrscheinlich“, „“überwiegend“, „meistens“, „höchst wahrscheinlich“ und „immer wieder (mal)“. Die Begriffe „sehr selten“, „recht häufig“, „denkbar“, „unvertretbar hoch“, „sehr großes Risiko“, „lassen sie es sein“, „fast regelhaft“, „immer“, „extrem häufig“, „stets“, „50%“ und „unvertretbar oft“ wurden einmal genannt.

Ärzte Bayern

1:2		
Wahrscheinlichkeitsbegriff	Häufigkeit	Prozent
sehr häufig	32	20.3%
häufig	28	17.7%
sehr oft	26	16.5%
jeder 2./ einer von 2	13	8.2%
oft	8	5.1%
in der Hälfte der Fälle	8	5.1%

Vier Nennungen gab es zu den Begriffen „wahrscheinlich“, „fast immer“ und „regelmäßig“. Zweimal wurden die verbalen Wahrscheinlichkeitsbegriffe „möglicherweise“, „möglich“, „(50:50)-Chance“ und „überwiegend“ genannt. Eine Nennung gab es jeweils zu den Begriffen „meist“, „sehr wahrscheinlich“, „denkbar“, „gehäuft“, „typischerweise“, „relativ häufig“, „meistens“, „sehr großes Risiko“, „immer“, „stets“, „mit großer Wahrscheinlichkeit“, „ungewöhnlich häufig“, „immer wieder (mal)“, „nicht verantwortbar“, „damit ist zu rechnen“, „zu 99% nicht“, „sehr sehr kritisch“, „unvertretbar oft“, „bedeutsam“, „relatives Risiko“, „zu häufig“, „relativ oft“ und „ziemlich sicher“.

9.2 Danksagungen

An dieser Stelle möchte ich allen Personen danken, die zum Gelingen dieser Dissertation beigetragen haben.

Mein besonderer Dank gilt Herrn Prof. Dr. Andreas Ziegler (Direktor am Institut für Medizinische Biometrie und Statistik) für die Überlassung des Themas, die stetig fördernde und treibende wissenschaftliche Unterstützung sowie die verständnisvolle und ausgesprochen geduldige Betreuung.

Bedanken möchte ich mich auch bei allen Mitarbeitern am Institut für Medizinische Biometrie und Statistik, die durch ihr fachliches Wissen, die allzeit hilfsbereite und freundliche Unterstützung sowie die motivierenden Worte in allen Phasen der Fertigstellung dieser Arbeit entscheidend zu ihrem Gelingen beigetragen haben. Namentlich nennen möchte ich unter anderem Frau Prof. Dr. Inken König, Frau Dr. Claudia Hemmelmann, Frau Christiane Wichmann sowie Frau Gabriele Schatton.

Herzlich bedanken möchte ich mich bei Herrn Mischa Marchlewski, der als Dokumentationsassistent am Institut für Medizinische Biometrie und Statistik mit fachlicher Unterstützung und uneingeschränkter Hilfsbereitschaft einen entscheidenden Beitrag zur endgültigen Fertigstellung der Dissertation geleistet hat.

Bedanken möchte ich mich auch bei Frau Steffi Mehlbeer für die gute Zusammenarbeit und die hilfreichen Anregungen zu Beginn dieser Arbeit.

Meinem Mann Marco und meiner Familie danke ich für die engagierte Unterstützung meiner Arbeit, das Verständnis und die große Geduld sowie den uneingeschränkten Glauben an mich, diese Dissertation sowie darüber hinaus.

9.3 Lebenslauf

Persönliche Angaben:

- Familienstand: verheiratet
- Nationalität: deutsch
- Geburtsdatum: 26.05.1977
- Geburtsort: Wismar

Ausbildung:

09/83 – 07/92	Grund- und Gesamtschule in Bad Kleinen
09/92 – 07/95	Gymnasium Fridericianum in Schwerin, Abschluss mit dem Abitur
10/95 – 09/99	Studiengang „Wirtschaftsrecht“ an der Fachhochschule Wismar, Abschluss als Diplom-Wirtschaftsjuristin (FH)
01/00	Eintritt in die Bundeswehr als Sanitätsoffizieranwärter (SaZ 17)
10/00 – 10/06	Studium der Humanmedizin an der Universität zu Lübeck, Abschluss mit der Approbation als Ärztin
10/06 – 10/08	Tätigkeit als Assistenzärztin im Bundeswehrkrankenhaus in Hamburg
10/08	Austritt aus der Bundeswehr und Anstellung am Universitäts-klinikum Schleswig-Holstein, Campus Lübeck
Seit 11/08	Tätigkeit als Assistenzärztin in der Klinik für Anästhesiologie und Intensivmedizin am Universitätsklinikum Schleswig-Holstein, Campus Lübeck