

# **Helfen uns Leitlinien im Individualfall von ihren Empfehlungen abzuweichen? Eine systematische Untersuchung von Leitlinien**

*Allgemeinmedizinisches Institut  
des Universitätsklinikums Erlangen  
Direktor: Prof. Dr. med. T. Kühlein*

Der Medizinischen Fakultät  
der Friedrich-Alexander-Universität  
Erlangen-Nürnberg

zur

Erlangung des Doktorgrades Dr. med.

vorgelegt von

Marc Florian Morgott

**Als Dissertation genehmigt von der  
Medizinischen Fakultät der Friedrich-Alexander-Universität  
Erlangen-Nürnberg**

**Vorsitzender des Promotionsorgans:** Prof. Dr. med. Markus Neurath

**Gutachter/in:** Prof. Dr. Thomas Kühlein  
Prof. Dr. Christoph Alexiou

**Tag der mündlichen Prüfung:** 1. Februar 2022

## Inhalt

1. Zusammenfassung in Deutsch mit deutschem Titel .....	1
2. Eingliederung in den wissenschaftlichen Kontext.....	3
2.1 Einleitung.....	3
2.2 Was sind Effektstärken?.....	5
2.2.1 Pearsons Korrelationskoeffizient $r$ .....	5
2.2.2 Die absolute und relative Risikoreduktion .....	6
2.2.3 Die Number needed to treat (und Number needed to harm).....	8
2.2.4 Die Odds Ratio.....	8
2.2.5 Cohen's $d$ .....	9
2.2.6 Zusammenfassende Bewertung der Effektstärken für unser Projekt ....	10
2.3 Treibende Faktoren für Überversorgung .....	10
2.4 Das Konzept der Quartären Prävention zum Schutz vor Überversorgung ..	12
2.4.1 Primärprävention.....	12
2.4.2 Sekundärprävention.....	12
2.4.3 Tertiärprävention.....	12
2.4.4 Quartärprävention .....	13
2.5 Voruntersuchung: Untersuchung von LL hinsichtlich Effektstärkenberichten	14
2.5.1 Methodik .....	14
2.5.2 Ergebnisse .....	16
2.6 Schlusswort und Überleitung.....	17
3. Originalpublikation .....	19
4. Anhänge .....	20
4.1 Literaturverzeichnis der Einordnung in den wiss. Kontext.....	20
4.2 Abbildungsverzeichnis der Einordnung in den wiss. Kontext .....	23
4.3 Abkürzungsverzeichnis der Einordnung in den wiss. Kontext.....	23
5. Publikationsliste.....	24

6. Danksagung .....	24
7. Lebenslauf.....	25

# Helfen uns Leitlinien im Individualfall von ihren Empfehlungen abzuweichen? Eine systematische Untersuchung von Leitlinien

**Hintergrund:** Leitlinien (LL) wurden eingeführt, um im Sinne von evidenzbasierter Medizin (EbM) auf Grundlage hochwertigster und aktuellster wissenschaftlicher Erkenntnisse Empfehlungen für die klinische Praxis zu erarbeiten und breit zugänglich zu machen. EbM legt aber nicht nur wissenschaftliche Evidenz zugrunde, sondern räumt patienteneigenen Vorzügen und ärztlicher Erfahrung ebenfalls einen wichtigen Stellenwert ein. In Folge dessen ist es regelmäßig notwendig, von LL-Empfehlungen abzuweichen. Das Ziel unserer Untersuchung war es, ob LL Abweichungen von ihren Empfehlungen im Einzelfall unterstützen. EBM fordert auf Basis von Effektstärken für Nutzen und Schaden in individuellen Fall eine Relevanzentscheidung zu treffen. Deshalb suchten wir, um unsere Fragestellung zu operationalisieren nach Angaben von absoluten Effektstärken (AES) für Nutzen und Schaden der in den LL gegebenen Empfehlungen.

**Methoden:** Zur Beantwortung der Studienfrage führten wir eine systematische Untersuchung durch. Diese umfasste die Analyse der wichtigsten LL zu chronischer koronarer Herzkrankheit und Diabetes Mellitus Typ 2 im deutschen und englischen Sprachraum anhand eines vorab entwickelten Fragebogens. Dabei wurden nur LL-Empfehlungen zu pharmakologischen Therapien berücksichtigt. Wenn AES berichtet wurden, beurteilten wir den nötigen Aufwand zum Auffinden dieser in der LL. Zudem erhoben wir, ob die LL eine Patienteninformation bereitstellt und wenn ja, ob sich auch in dieser Angaben zu AES finden lassen.

**Ergebnisse:** Es konnten 13 LL eingeschlossen werden. Dabei wurden 114 Empfehlungen zur Pharmakotherapie identifiziert, zu denen sich in 75% keinerlei Angaben zu Nutzen oder Schaden fanden. Für 31 Empfehlungen (22%) wurden von einer oder mehreren AES bezüglich entweder Nutzen oder Schaden angegeben. Bei

fünf Empfehlungen (3%) wurden eine oder mehrere AES sowohl für Nutzen als auch für Schaden berichtet. Die angegebenen AES waren in drei der 36 Empfehlungen (8%) einfach zu finden. Patienteninformationsmaterial wurde durch drei der 13 LL (23) bereitgestellt. Nur in einer von diesen wurden AES benannt.

Schlussfolgerung: Aktuelle LL zu Diabetes Mellitus Typ 2 und chronischer koronarer Herzkrankheit beinhalten unzureichende Angaben zu AES von Nutzen und Schaden der in ihren Empfehlungen vorgeschlagenen Pharmakotherapie. Dadurch entbehren sie zufriedenstellender Informationen, um im Individualfall von ihren Empfehlungen abzuweichen. LL unterstützen in ihrer gegenwärtigen Form nicht die zentrale Forderung von EBM, medizinische Entscheidungen individuell zu treffen.

## 2. Eingliederung in den wissenschaftlichen Kontext

### 2.1 Einleitung

„Evidenzbasierte Medizin (EbM) ist der gewissenhafte, ausdrückliche und vernünftige Gebrauch der gegenwärtig besten externen, wissenschaftlichen Evidenz für Entscheidungen in der medizinischen Versorgung individueller Patienten. Die Praxis der EbM bedeutet die Integration individueller klinischer Expertise mit der bestverfügbaren externen Evidenz aus systematischer Forschung“ [1, 2]. Leitlinien (LL) sind ein fester Bestandteil der Integration von wissenschaftlicher Evidenz in den klinischen Alltag. Dabei bemühen sich Vereinigungen aller Art, die unüberschaubare Flut von neuen und alten Erkenntnissen zusammenzufassen und daraus - unter Hinzuziehung von Expertenwissen - Empfehlungen zu generieren, um die Ärzteschaft möglichst auf dem aktuellen Stand der Wissenschaft zu halten und ihr ein gewisses Maß an Handlungssicherheit zu geben. Dabei befinden sich diese Leitlinien, durch ihre Natur bedingt, ständig in der Weiterentwicklung und - nicht selten - im Kreuzfeuer der Diskussion, wie an späterer Stelle noch dargestellt werden soll. Boudoulas und Kollegen [3] identifizierten insbesondere folgende Schwächen von LL:

- Einschränkung einer personalisierten Medizin
- Mangel an Beachtung klinischer Erfahrung
- Mangel an Berücksichtigung der Pathophysiologie

Für unsere Überlegungen und Forschungsfrage stand der erste Punkt, die Limitation einer personalisierten Medizin durch LL, im Vordergrund. Insbesondere im hausärztlichen Bereich ist es eine wesentliche Aufgabe, Patienten auch vor unnötiger Medizin oder Übermedikalisierung zu schützen [4]. Dies ist aber nur möglich, wenn dem behandelnden Arzt auch die entsprechenden Kenngrößen vorliegen, die eine bestmögliche Entscheidung ermöglichen, wobei die Entscheidung auch negativ ausfallen kann, also beispielsweise auch auf Medikation zu verzichten. Dies insbesondere vor dem Hintergrund, dass LL vorwiegend ein einzelnes Krankheitsbild in hohem Detail behandeln und Komorbiditäten hingegen oft nicht beachten [5]. Bei unserer alternden Bevölkerung ist Multimorbidität jedoch häufig die Realität im Alltag des Hausarztes. Soll die aus der Anwendung aller entsprechenden LL-Empfehlungen fast unvermeidlich folgende Polypharmazie vermieden werden, kommt man nicht umhin, von manchen Empfehlungen abzuweichen. Daraus ergaben sich folgende

Fragestellungen: Sind Leitlinien gegenwärtig dazu geeignet, (Haus)Ärzte ausreichend zu informieren, damit sie bei der Behandlung individueller Patienten a) sich selbst eine fundierte Meinung bilden und b) den Patienten im Sinne des informed consent umfassend und optimal beraten können? Das eigentliche Ziel wäre dabei, die bestmögliche Entscheidung gemeinsam zu der Frage zu treffen, welche therapeutischen Optionen angewandt werden sollen oder auf welche verzichtet werden kann. Für den Arzt stellt sich dabei vor allem eine Frage: Ist ein gegebenes Medikament/Intervention indiziert oder nicht? Dabei entscheidet sich Indikation grundsätzlich auf Patientenebene. Sie kann auf LL-Ebene nicht getroffen werden. Wie wirksam ist das Medikament / die Intervention / etc., welche ich für den Patienten in Erwägung ziehe? Es zeigt sich immer mehr, dass die Praxis, statistische Signifikanz für wissenschaftliche Ergebnisse als vorrangigen Parameter anzusehen, nicht hinreichend ist. Die Wirksamkeit einer Intervention lässt sich damit nicht ausreichend eruieren. Bortz und Döring bringen dies treffend auf den Punkt: „Die Entwicklung einer Wissenschaft ausschließlich von signifikanten Ergebnissen abhängig zu machen, könnte also bedeuten, dass Theorieentwicklungen weiterverfolgt werden, die auf minimalen, wenngleich statistisch signifikanten Effekten beruhen, deren Erklärungswert für reale Sachverhalte eigentlich zu vernachlässigen ist“ [6] [S. 28]. So kommt es, dass Effektmaße als Ausdruck einer „klinischen Signifikanz“ oder besser „Relevanz“ immer mehr in den Vordergrund rücken [7, 8]. Somit verfolgte unser Forschungsansatz unter dem initialen Arbeitstitel „Effektstärken in Leitlinien“ das *Ziel*, herauszufinden, ob und in welchem Ausmaß in den aktuellen LL Effektstärken für Nutzen und Schaden pharmakologischer Interventionen berichtet werden und darüber zunächst der Arzt über die zu erwartenden Effekte möglicher Therapien informiert wird, um ihm die Abwägung bezüglich eines Für oder Wider dieser Intervention für den individuellen Patienten zu erlauben.

### **Vorarbeiten zum Forschungsprojekt**

Im Folgenden sollen exemplarisch einige wichtige Aspekte der Vorarbeiten zu dem Projekt dargestellt werden. Was sind Effektmaße? Warum besteht die Notwendigkeit zur Abwägung medizinischer Maßnahmen? Zuletzt wird eine kleine Vorstudie, welche ich im Rahmen der Promotionsarbeit durchführte und welche einen initialen Blick auf die Häufigkeit der Nennung verschiedener Effektmaße in Leitlinien gestattete, vorgestellt. *Diese Vorstudie war bei der Auswahl der Kriterien für die*



Hauptuntersuchung hilfreich. Die Vorarbeiten, also der Weg zum Projekt, soll damit im Wesentlichen beschrieben sein. Am Ende wird die Hauptuntersuchung zusammengefasst dargestellt.

## 2.2 Was sind Effektstärken?

Effektstärken oder Effektmaße sind objektive und (typischerweise) standardisierte Maße eines beobachteten Effekts [9] [S. 785], oder nach Cohen [10] [S. 79-80]:

- Das Ausmaß, zu welchem ein Phänomen in der Population vorhanden ist oder
- Das Ausmaß, in welchem die Nullhypothese falsch ist.

Heutzutage wird nicht selten angeführt, dass Effektstärken das Hauptergebnis einer quantitativen Studie seien [11]. Nach Aussage von Cohen [10] sollte ein Effektmaß folgende Kriterien erfüllen:

- Es sollte eine dimensionslose Zahl sein.
- Es sollte nicht von den Einheiten der Maßskala abhängen.
- Es sollte nicht von der Stichprobengröße abhängen und
- Annähernd Null sein, falls die Nullhypothese des Tests nicht verworfen werden soll.

Folglich sind einige statistische Maßzahlen, insbesondere Korrelationskoeffizienten, geeignet. Im Zuge der an späterer Stelle dieses Schriftstückes dargestellten Voruntersuchung sowie praktischen Überlegungen erschienen mir eingangs insbesondere folgende Effektmaße als für uns als relevant:

- Pearsons Korrelationskoeffizient  $r$
- Die absolute Risikoreduktion (ARR)
- Die Number needed to treat (NNT) (und Number needed to harm (NNH))
- Die Odds Ratio (OR)
- Cohen's  $d$

Diese sollen im Folgenden knapp dargestellt werden.

### 2.2.1 Pearsons Korrelationskoeffizient $r$

Der Korrelationskoeffizient  $r$  bemisst die Stärke eines Zusammenhanges basierend auf der Kovarianz zweier annähernd linearen, quantitativen und unabhängigen Variablen [12].

Die Kovarianz  $s_{xy}$  berechnet sich folgendermaßen:

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{n - 1}$$

Dabei sind  $x$  und  $y$  die voneinander unabhängigen Variablen und  $n$  die Stichprobengröße. Der Korrelationskoeffizient  $r$  wird dann anhand der durch das Produkt der Standardabweichungen der Variablen  $x$  und  $y$  dividierten Kovarianz berechnet:

$$r = \frac{s_{xy}}{s_x \cdot s_y}$$

Das Ergebnis ist eine dimensionslose Zahl zwischen  $-1$ , was einen absoluten gegensinnigen Zusammenhang bedeuten würde, und  $+1$ , was einen absoluten gleichsinnigen Zusammenhang beschreibt. Im Falle von  $r=0$  wäre keinerlei Zusammenhang zwischen den betrachteten Variablen evident [11].

Bezüglich der Verwendung von  $r$  als Effektgröße kann weitläufig diskutiert werden, welcher Wert als kleiner oder großer Effekt angenommen werden soll. Insofern finden sich unterschiedliche Empfehlungen. Einige Autoren legen die Einteilung

- $r = 0,1$  oder  $r = 0,2$  geringer Effekt
- $r = 0,3$  mittlerer Effekt
- $r = 0,5$  großer Effekt

nahe (vgl. z.B. [8], [12]), andere bevorzugen eine strengere Maßgabe:

- $r = 0,2$  geringer Effekt
- $r = 0,5$  mittlerer Effekt
- $r = 0,8$  großer Effekt

(vgl. z.B. [11]).

Insgesamt ist Pearsons Korrelationskoeffizient  $r$  eine häufig verwendete und gut verständliche Effektgröße. Zudem lassen sich verschiedene andere Effektstärken in  $r$  umwandeln, was deren Interpretation erleichtern kann [13].

### 2.2.2 Die absolute und relative Risikoreduktion

Die Absolute Risikoreduktion, auch das attributable Risiko genannt, berechnet sich aus der Differenz der Häufigkeit eines Outcomes der Intervention A und der Häufigkeit

eines Outcomes der Intervention B, einschließlich keiner Intervention [12]. Dies geschieht wie folgt:

$f_A$  = Häufigkeit des Outcomes der Intervention A

$f_B$  = Häufigkeit des Outcomes der Intervention B

$ARR = f_A - f_B$

Wenn man in einer Studie beispielsweise die Gabe eines Placebos (=Intervention A), und die Gabe eines Medikamentes (Intervention B) hinsichtlich eines definierten Ereignisses (z.B. Tod) vergleichen möchte, errechnet sich die absolute Risikoreduktion entsprechend

$ARR = \text{Ereignisrate}_{\text{Placebo}} - \text{Ereignisrate}_{\text{Medikament}}$

(vgl. [14]).

Somit gibt die absolute Risikoreduktion an, um wie viele Prozentpunkte das bestehende Risiko durch z.B. eine Intervention gesenkt wird.

Setzt man beispielsweise Ereignisrate Tod<sub>Placebo</sub> = 2% und

Ereignisrate Tod<sub>Medikament</sub> = 1%

So ergibt sich als  $ARR = 2\% - 1\% = 1\%$

Das Risiko einer der Grundgesamtheit angehörigen Person, mit dem Medikament zu versterben, ist also um ein Prozent niedriger als das einer Person unter Placebo-Intervention. Dies grenzt die absolute Risikoreduktion von der relativen Risikoreduktion (RRR) ab, welche die Risiken direkt miteinander vergleicht, und nach der Formel

$RRR = 100 \times (1 - \text{Ereignisrate}_{\text{Medikament}} / \text{Ereignisrate}_{\text{Placebo}})$

berechnet [14]. Das Ergebnis wäre im oben angeführten Beispiel  $RRR = 50\%$ . Da dies irreführend sein kann, rechneten wir in unseren Untersuchungen die relative Risikoreduktion nicht zu den zulässigen Effektmaßen.

Insgesamt lässt sich die absolute Risikoreduktion als solides Maß für den Effekt einer Intervention bezeichnen. Zudem lässt sich aus ihr eine weitere, bildlich gut vorstellbare Effektgröße ableiten:

### 2.2.3 Die Number needed to treat (und Number needed to harm)

Die Number needed to treat lässt sich einfach aus der absoluten Risikoreduktion berechnen, und zwar nach der Formel

$$\text{NNT} = 1/\text{ARR}$$

(vgl. [14, 15]).

Durch sie wird bezeichnet, wie viele Patienten einer Intervention zugeführt werden müssen, damit eine(r) dieser Patienten das angezielte Outcome erfährt bzw. nicht erfährt. Die Number needed to harm errechnet sich wie die Number needed to treat, dabei werden aber negative bzw. nicht erwünschte Outcomes betrachtet.

Die Number needed to treat und Number needed to harm sind sicherlich eine der klinisch intuitivsten Effektgrößen mit direktem klinischen Bezug [15]. Vor dem geistigen Auge ist es wohl für viele Kolleginnen und Kollegen gut machbar, sich Patientengruppen verschiedener Größen vorzustellen und mit Hilfe dieser Vorstellung Therapieoptionen abzuwägen, was mit abstrakten Zahlen typischerweise nicht jedermann intuitiv von der Hand gehen sollte.

### 2.2.4 Die Odds Ratio

Die Odds Ratio bezeichnet den Vergleich zweier „Odds“ (grob als „Chance“ oder „Wahrscheinlichkeit“ übersetzbar), dass nach Exposition gegenüber einem Faktor ein Ereignis eintritt gegenüber der Wahrscheinlichkeit des Eintretens des Ereignisses in einer Referenzsituation [16]. Sie errechnet sich aus dem Kreuzprodukt der Häufigkeiten in der Vierfeldertafel wie folgt:

$$\text{OR} = \frac{\mathbf{a \cdot d}}{\mathbf{b \cdot c}}$$

(vgl. z.B. [12]).

Ein Wert von 1 in der odds Ratio würde keinen Zusammenhang zwischen den Merkmalen bezeichnen. Eine höhere Zahl zeigt eine erhöhte Wahrscheinlichkeit unter dem angelegten Kriterium an, eine kleinere Zahl eine verringerte Wahrscheinlichkeit.

Nehmen wir als Beispiel eine rein hypothetische Untersuchung zur Erkrankungshäufigkeit bei geimpften und nicht geimpften Probanden. Nehmen wir folgende Verteilung, dargestellt in einer Vierfeldertafel, in dieser hypothetischen Untersuchung an:

	<b>Erkrankt</b>	<b>Nicht erkrankt</b>	
<b>Nicht geimpft</b>	a = 37	b = 23	60
<b>Geimpft</b>	c = 9	d = 51	60
	46	74	

Somit wäre die  $OR = (37 \times 51) / (23 \times 9) = 1887 / 207 = 9,116$  und damit die Wahrscheinlichkeit, als nicht geimpfter zu erkranken ca. 9 zu 1.

Wichtig zu bemerken ist, dass die Odds Ratio eine relative Größe ist. Des Weiteren ist sie im Falle von Zahlen über 1 relativ eingängig. Im umgekehrten Fall, falls die Odds Ratio z.B. 0,2 ist, fällt die Interpretation auf den ersten Blick nicht mehr ganz so leicht. Zur Veranschaulichung sei das obige Beispiel in der Vierfeldertafel vertauscht:

	<b>Erkrankt</b>	<b>Nicht erkrankt</b>	
<b>Geimpft</b>	a = 9	b = 51	60
<b>Nicht geimpft</b>	c = 37	d = 23	60
	46	74	

Hier wäre die  $OR = (9 \times 23) / (51 \times 37) = 207 / 1887 = 0,11$ . Für 0,11 geimpfte Personen würde also eine nicht geimpfte Person erkranken - oder einfacher: für elf geimpfte Personen erkranken 100 nicht geimpfte. Schlussendlich stellt die odds Ratio eine Effektgröße dar, welche nach etwas Einarbeitung sehr anschaulich sein kann.

### 2.2.5 Cohen's d

Cohen's d ist eine Maßzahl für die Effektgröße bei Mittelwertunterschieden und somit insbesondere bei der statistischen Auswertung mit t-Test geeignet. Es berechnet sich aus der Differenz der Mittelwerte dividiert durch die Standardabweichung der Kontrollgruppe:

$$d = \frac{m_a - m_b}{\sigma}$$

(vgl. [10]).

Dabei stehen die Variablen für:

d	Effektstärkenindex für t-tests von Mittelwerten in der Standardeinheit
$m_a, m_b$	Mittelwerte der Populationen in der Originaleinheit
$\sigma$	Die Standardabweichung entweder der Population a oder b (da beide als gleich angesehen werden)

Für die Betrachtung von Cohen's d werden gleiche Gruppengrößen und Gruppenvarianzen angenommen. Cohen schlägt dabei folgende Werte zur Interpretation der Größe eines Effekts vor:

- d = 0,2 oder r = 0,2 geringer Effekt
- d = 0,5 mittlerer Effekt
- d = 0,8 großer Effekt

Cohen's d stellt kein für Nicht-Statistiker oder gar Patienten bekanntes oder nachvollziehbares Effektmaß dar.

#### 2.2.6 Zusammenfassende Bewertung der Effektstärken für unser Projekt

Teils bereits zu Beginn, teils im Laufe des Projektes zeigten sich einige der Maßzahlen somit für die LL-Bewertung als wenig relevant. Ein Beispiel ist Cohen's d, da es zum einen für das Zielklientel zu kompliziert ist, zum Anderen in den LL schlicht keine Verwendung finden. Des Weiteren wurden relative Effektmaße aufgrund des nötigen Hintergrundwissens als eher ungeeignet für unsere Zwecke befunden. Absolute Maße, wie die NNT oder ARR, sind dagegen deutlich eingängiger und daher Ärzten und Patienten zugänglicher, was sie für den Gebrauch in LL zweckmäßiger macht. Letztlich wäre ein Ziel, in LL Effektstärken so verständlich zu machen, dass auch Patienten sie verstehen, und mit ihrer Hilfe Entscheidungen treffen können.

### 2.3 Treibende Faktoren für Überversorgung

In der Entscheidung für oder gegen eine Therapiemaßnahme gilt es viele Faktoren zu berücksichtigen.

- Leitlinien: Eine Schwierigkeit stellt die Multimorbidität und aus ihr häufig folgend die Polypharmazie vor allem der alternden Bevölkerung dar. Doch nicht nur ältere Menschen, sondern auch Personen mit geringem sozioökonomischen Status und weiblichem Geschlecht sind von einem höheren Risiko für Multimorbidität betroffen [17]. LL fokussieren im Wesentlichen auf einzelne Erkrankungen. Boyd und Kollegen [5] wendeten konsequent die Empfehlungen von neun einschlägigen LL zu den häufigsten hausärztlich betreuten chronischen Erkrankungen auf eine fiktive Patientin

an. Diese hypothetische 79-jährige Patientin hatte die Diagnosen: Osteoporose, Osteoarthritis, Typ 2 Diabetes Mellitus, Bluthochdruck und chronisch obstruktive Lungenerkrankung. Durch Anwendung aller Empfehlungen würde diese Patientin 19 Medikamentendosen zu fünf Tageszeiten einnehmen, zusätzlich zu 14 nichtpharmakologischen Interventionen. Insgesamt kommen die Autoren zu dem Schluss, dass spätestens die Summe der LL keine angemessene Grundlage zur Behandlung älterer Patienten darstellt. Dabei ist es wahrscheinlich, dass mit der Medikamentenanzahl nicht nur die potentiell dem Patientenwohl zuträglichen Wirkungen entfaltet werden, sondern auch das Risiko von unerwünschten Arzneimittelwirkungen und potentiell schädlichen Medikamenten-Interaktionen steigen [18].

- Defensive Medizin: Ein weiterer Treiber von Polypharmazie und Überversorgung ist die defensive Medizin. Darunter wird verstanden, dass Untersuchungen oder Therapien/Interventionen zur juristischen Absicherung des behandelnden Arztes und weniger zum Wohle des Patienten durchgeführt werden [19]. Insbesondere bildgebende Verfahren oder antiinfektive Therapien können als Beispiele genannt werden, die gerne sicherheitshalber eingesetzt werden. Dass defensive Medizin nicht nur regelmäßig, sondern häufig stattfindet, ist hinreichend belegt (vgl. z.B. [20]). Dies scheint nicht verwunderlich, da juristische Auseinandersetzungen zwischen Patient und Arzt zunehmen [21] [S. 2] und sich dies auch in der Presse widerspiegelt (vgl. z.B. [22]). In Deutschland kann ein Patient klagen, sobald er mit der Behandlung auch nur unzufrieden ist [21] [S. 2]. Die Diskussion zu Grundlagen und Berechtigung für defensive Medizin soll hier nicht vertieft werden, festzustellen ist jedoch, dass defensive Medizin die Arzt-Patienten-Beziehung sowie die medizinische Behandlung der Patienten verschlechtert sowie das Gesundheitssystem belastet [20].
- Ökonomische Faktoren: Auch sozioökonomische Gesichtspunkte spielen häufig eine Rolle bei der Auswahl medizinischer Prozeduren, was nicht zuletzt auch dem in Deutschland eingesetzten DRG-System (DRG steht für „Diagnosis Related Groups“) anzulasten sein dürfte. Dabei werden teils Anreize zur Durchführung von teuren und relativ unaufwändigen Therapiemaßnahmen gesetzt. Auch im niedergelassenen Bereich ist der wirtschaftliche Druck hoch [23].

Von herausragender Bedeutung sollte die individuelle Einstellung und Präferenz der Patienten selbst sein. Wie man nicht zuletzt an der aktuellen Diskussion zur Impfung gegen SARS-CoV-2 erkennen kann [24], gibt es bezüglich medizinischer Maßnahmen unterschiedliche individuelle Meinungen und Präferenzen, welche die Entscheidung für oder gegen eine Intervention bedingen. Aber auch Lifestyle-Entscheidungen können eine Rolle spielen [25]. Insbesondere im Bereich der individuellen Meinungsbildung der Patienten ist eine fundierte, ausführliche und vor Allem neutrale und vertrauensbasierte Aufklärung und Informationsbereitstellung wichtig.

Es gibt also vielfältige Gründe und auch Interessenskonflikte bei der Abwägung für oder gegen eine medizinische Intervention, und insbesondere hinsichtlich des Patientenwohls ist diese Abwägung von herausragender Notwendigkeit.

#### 2.4 Das Konzept der Quartären Prävention zum Schutz vor Überversorgung

Es gibt also einige Anlässe, zu denen Menschen zu viel medizinische Intervention erfahren können. Hier setzt die quartäre Prävention an. Zur Abgrenzung von den klassisch bekannten Formen der Prävention seien diese im Folgenden grob dargestellt.

##### 2.4.1 Primärprävention

Die primäre Prävention umfasst medizinische Interventionen an Personen, die selbst kein Leiden verspüren und bei denen auch ärztlicherseits keine Krankheit diagnostiziert werden kann. Sie verfolgt das Ziel, diese gesund zu erhalten. Meist handelt es sich dabei um populationsbezogene Maßnahmen, darunter fallen z.B.:

- Impfungen
- Früherkennungsuntersuchungen / Check-Ups
- Lebensstilberatungen
- Etc.

(vgl. z.B. [19])

##### 2.4.2 Sekundärprävention

In der Situation der Sekundärprävention fühlt sich der Patient weiterhin gesund, ärztlicherseits lassen sich jedoch Diagnosen stellen, die sich typischerweise aus vermeintlichen oder tatsächlichen pathologischen Messwerten ergeben. Sie entsprechen damit eigentlich eher Risikofaktoren als Krankheiten. In der sekundären Prävention werden Maßnahmen ergriffen, die bei diesen Risikofaktoren ein Ausbrechen einer Erkrankung verhindern sollen. Dies betrifft beispielsweise Diagnosen wie arterielle Hypertonie, Hypercholesterinämie, Diabetes mellitus Typ 2, etc. Es handelt sich um einen Graubereich, in dem unklar ist, ob nun schon Krankheit vorliegt oder nicht (vgl. z.B. [19]).

##### 2.4.3 Tertiärprävention

In der Tertiären Prävention treffen subjektives Leiden und ärztliche Krankheitsdiagnose zusammen. Die Patienten haben das Ereignis erlitten, das in der sekundären Prävention verhindert werden sollte.; z.B. Schlaganfall, Herzinfarkt, osteoporotische

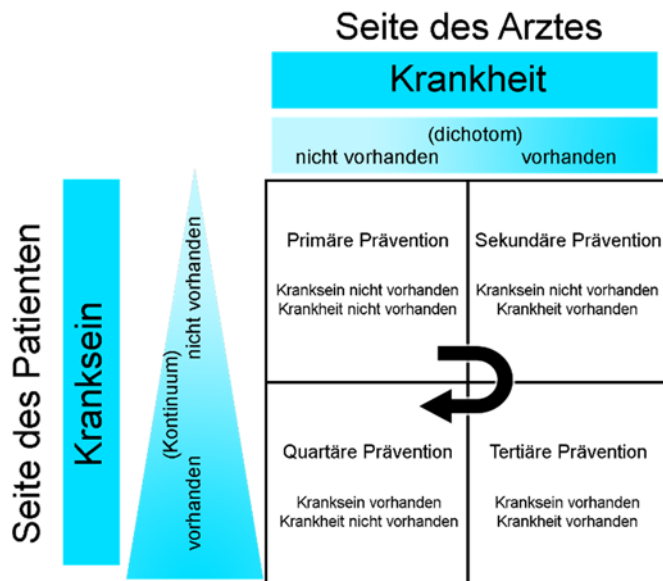


Fraktur, etc. Nun soll ein weiteres Fortschreiten der Erkrankung oder ein Wiederauftreten des Ereignisses nach Möglichkeit verhindert und die Krankheitsfolgen so weit möglich auf ein Minimum begrenzt werden. (vgl. z.B. [19])

#### 2.4.4 Quartärprävention

Die quartäre Prävention beschreibt die Situation, in der Patienten Symptome verspüren, aber ärztlicherseits keine Krankheit diagnostiziert werden kann. Diese Situation ist in der Hausarztpraxis häufig und sollte nicht in Überdiagnostik enden. Stattdessen sollten wesentliche gefährliche Krankheiten, die das Symptom verursachen könnten, ausgeschlossen und dann aufmerksam abgewartet werden. Die meisten Symptome werden von alleine vergehen und werden keineswegs von gefährlichen Krankheiten verursacht. Quartäre Prävention meint aber nicht nur die Prävention der in dieser Situation häufigen Überdiagnostik, sondern Überdiagnostik allgemein und darüber hinaus die Verhinderung aller unnötiger Formen von Medizin. Dies ist insbesondere im Setting des Hausarztes von Belang, wo im Niedrigprävalenzbereich – in dem sich viele Patienten vorstellen, aber nur relativ wenige wirklich ernsthaft krank sind – die Versuchung von zu viel praktizierten Diagnostiken, Überweisungen und Verordnungen besonders hoch ist [19].

An diesen genannten Punkten setzt die quartäre Prävention an. Sie zielt auf „jegliche Form der Verhinderung nicht bedarfsgerechter Medizin“ [19] [S. 597] ab, stellt also hier auch wieder das Patientenwohl vorne an. Wie zuvor genannt, arbeitet der Hausarzt im Niedrigprävalenzbereich. Dort geht es häufig „um das Problem des sich krank Fühlens, ohne dass eine konkrete Krankheit vorliegt“ [4]. Man kann diese Gegebenheit in Relation zu den anderen Präventionsformen in einer Vierfeldertafel, wie in Abb. 1 dargestellt, veranschaulichen.



**Abb.1:**

Krankheit als Dichotomie gegenüber dem Kontinuum des Krankseins.

Nach: [4]

Umso mehr ist es wichtig, dass dem behandelnden Arzt die Möglichkeit gegeben wird, für den individuellen Patienten abzuwägen, welche Therapien sinnvoll erscheinen. Dies wäre insbesondere dann auf dem Boden der Evidence Based Medicine möglich, wenn in den entsprechenden LL diese Informationen enthalten wären.

## 2.5 Voruntersuchung: Untersuchung von LL hinsichtlich Effektstärkenberichten

Um einen groben Überblick zu bekommen, ob und welche Effektgrößen in Leitlinienempfehlungen verwendet werden, wurde im Januar 2016 vorab eine Untersuchung von LL durchgeführt.

### 2.5.1 Methodik

#### 2.5.1.1 Auswahl von Effektmaßen

Für einen groben Überblick wurden folgende Effektmaße ausgewählt:

- d                      Cohen's d
- r                      Korrelationskoeffizient nach Pearson
- WMD                weighted mean difference (gewichtete mittlere Differenz)
- SMD                standardized mean difference (standardisierte Mittelwertdifferenz)

Die Entscheidung basierte dabei auf praktikablen Überlegungen, mit Cohen's d als anerkanntem Effektmaß und Pearson's r wegen seiner breiten Zugänglichkeit. Die WMD und SMD wurden gewählt, da sie als Effektmaße bei Metaanalysen etabliert sind und solche für CPG als hohe Stufe der Evidenz eine entsprechende Rolle spielen sollten.

### 2.5.1.2 *Literaturrecherche*

Die Datenbanken auf <http://www.awmf.org> und [www.leitlinien.de](http://www.leitlinien.de) wurden nach Leitlinien durchsucht, welche folgende Einschlusskriterien erfüllen mussten:

- Deutschsprachige Publikation
- Langversion vorliegend
- Gültigkeit zum Untersuchungszeitpunkt (Januar 2016)
- S3-Standard

Die auf diesen Plattformen identifizierten Leitlinien wurden danach in zwei Gruppen separiert, je nachdem ob sie die Einschlusskriterien erfüllten oder nicht.

### 2.5.1.3 *Literaturanalyse*

Unter Verwendung des Freewareprogrammes Adobe Acrobat Reader DC (Version 2015.010.20056) wurden die pdf-Dateien jeder eingeschlossenen Leitlinie mit Hilfe der Suchfunktion anhand einer Schlagwortsuche durchforstet. Als Schlagworte dienten:

- Effekt
- Effektstärke
- Effektgröße
- Effektgrösse
- d=
- r=
- ES
- WMD
- SMD

Es wurde jeder Treffer, welcher eine tatsächliche Beschreibung einer Effektgröße in der Leitlinie anzeigte, notiert.

Effektstärken wurden gezählt, wenn eine quantifizierbare Angabe vorlag, also wenn eindeutige Zahlen genannt wurden oder Kategorien vordefiniert wurden (wie z.B.  $d=0.3$  beschreibt einen geringen Effekt,  $d=0.5$  beschreibt einen moderaten Effekt,  $d=0.7$  beschreibt einen hohen Effekt) und sich dann auf diese Kategorien bezogen wurde. Zur Wertung reichte eine einzelne Nennung der entsprechenden Effektgröße in der Leitlinie, sofern sich diese auf eine Empfehlung in der jeweiligen Leitlinie bezog.

## 2.5.2 Ergebnisse

### 2.5.2.1 Datenbankrecherche

Die Datenbankrecherche wurde am 07. und 08. Januar 2016 durchgeführt und resultierte in 156 gefundenen S3-Leitlinien in ihrer Langversion. 120 davon erfüllten auch die anderen oben genannten Einschlusskriterien.

### 2.5.2.2 Literaturanalyse

Die Ergebnisse der Literaturanalyse kann **Tab.1** entnommen werden.

Name of ES measurement	Reported in GPG	
	Number	%
Cohen's d	17	14,2
SMD	10	8,3
r	8	6,7
WMD	7	5,8

**Tab.1:** Ergebnis der Literaturanalyse. Dargestellt sind die absolute Nummer der Effektstärkennennungen über die 120 untersuchten Leitlinien hinweg sowie eine prozentuale Angabe, welcher Anteil der untersuchten Leitlinien welches Effektstärkemaß zumindest einmalig berichtete.

Cohen's d wurde in 17 (14,2%) der betrachteten Leitlinien zumindest einmal genannt. Der Korrelationskoeffizient r wurde insgesamt von 8 Leitlinien (entsprechend 6,7%) berichtet. SMD bei 10 (8,3%) und WMD bei 7 (5,8%).

### 2.5.2.3 Diskussion der Vorstudienresultate

Es zeigt sich eine relativ geringe Anzahl an Nennungen der ausgewählten Effektstärkemaße in den Empfehlungen der untersuchten Leitlinien. Nebenbei konnten weitere Zahlen, die nicht Teil dieser Untersuchung waren, identifiziert werden, die sporadisch in den Leitlinien genannt wurden, namentlich Konfidenzintervalle, Odds ratios und/oder hazard ratios und/oder risk reductions, welche ebenso Effektgrößen

darstellen und für die Hauptuntersuchung als potentielle Beobachtungsziele dienen könnten.

Einige Leitlinien bieten weitere Dokumente wie Leitlinienreports oder evidence reports, welche Nennungen von Effektstärkemaßen beinhalten könnten. Diese Dokumente wurden in dieser Voruntersuchung nicht berücksichtigt.

Fehlende Effektstärkenennungen in Leitlinien könnten darauf zurückzuführen sein, dass die zugrundeliegenden Studien ihrerseits keine Nennung dieser Kennzahlen vornehmen und somit den Leitlinienerstellern diese Daten vorenthalten.

## 2.6 Schlusswort und Überleitung

In dieser Einordnung in den wissenschaftlichen Kontext wurden Vorüberlegungen und Voruntersuchungen zu der Arbeit „Helfen uns Leitlinien im Individualfall von ihren Empfehlungen abzuweichen? Eine systematische Untersuchung von Leitlinien“, welche im Folgenden einzusehen ist, dargestellt.

Insgesamt steht nach unserer Auffassung bei der ärztlichen Arbeit das Wohlergehen des individuellen Patienten im Fokus. LL „... zielen auf die Krankheit ab, nicht auf einen bestimmten Patienten.“ [26]. Sie sollten nicht als Versuch gesehen werden, die Patientenversorgung zu standardisieren, sondern individualisierte Medizin ermöglichen. Angesichts der Multimorbidität der alternden Bevölkerung ist es essentiell, im Individualfall von Leitlinienempfehlungen abzuweichen. Hierfür sollten die LL Hilfestellungen bieten, indem sie zur individuellen Abwägung von Risiken und Nutzen der empfohlenen Therapien Nennungen von Effektmaßen bereitstellen. Inwiefern sie dies tun war Gegenstand unserer Untersuchung. Zu dieser entschieden wir uns durch Diskussion in unserer Forschungsgruppe, LL zu zwei hausärztlich hoch relevanten Themenfeldern, der chronischen KHK sowie dem Diabetes Mellitus Typ 2, hinsichtlich der Fragestellung zu untersuchen. Um einen internationalen Überblick zu gewinnen, wurden als Ziel-LL die jeweils hauptsächlich gültigen LL in Deutsch- und englischsprachigen Ländern bzw. der Europäischen Union benannt:

- Deutschland
- Österreich
- Europa
- England
- USA
- Kanada

- Australien

Die Recherche nach Schweizer Leitlinien ergab deren Verweis auf die europäisch gültigen Leitlinien, sodass die Schweiz in unserer Analyse nicht einzeln genannt wird. Hier machten wir die je vorwiegend gültige LL ausfindig und schlossen sie in unsere Analyse ein, sofern sie in deutscher oder englischer Sprache publiziert und aktuell gültig war. Es konnten 13 LL eingeschlossen werden, deren Untersuchung hinsichtlich ihrer Empfehlungen im Bereich Therapie, ob diese Vor- und/oder Nachteile der entsprechenden pharmakologischen Intervention berichteten, erfolgte durch zwei Personen (Hrn. Heinmüller und mich selbst). Dabei wurden 114 Empfehlungen zur Pharmakotherapie identifiziert, von denen 75% keinerlei Angaben zu Nutzen oder Schaden enthielten.

Im Folgenden ist die entsprechende Publikation im Journal of Evaluation in Clinical Practice mit ihren Ergebnissen vollumfänglich einzusehen.

### 3. Originalpublikation

Morgott M, Heinmüller S, Hueber S, Schedlbauer A, Kühlein T. Do guidelines help us to deviate from their recommendations when appropriate for the individual patient? A systematic survey of clinical practice guidelines. *J Eval Clin Pract.* 2019;1–9. <https://doi.org/10.1111/jep.13187>

## 4. Anhänge

### 4.1 Literaturverzeichnis der Einordnung in den wiss. Kontext

- [1] Cochrane Deutschland. <https://www.cochrane.de/de/ebm>; Abgerufen am 2021  
Abgerufen am 2021 Feb 1.
- [2] Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312(7023): 71–2  
[<https://doi.org/10.1136/bmj.312.7023.71>][PMID: 8555924]
- [3] Boudoulas KD, Leier CV, Geleris P, Boudoulas H. The shortcomings of clinical practice guidelines. *Cardiology* 2015; 130(3): 187–200  
[<https://doi.org/10.1159/000371572>][PMID: 25790843]
- [4] Kuehle T, Sghedoni D, Visentin G, Gérvas J, Jamouille M. Quartäre Prävention, eine Aufgabe für Hausärzte. *Primary Care* 2010; 10(18): 350–4.
- [5] Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* 2005; 294(6): 716–24  
[<https://doi.org/10.1001/jama.294.6.716>][PMID: 16091574]
- [6] Bortz J, Döring N. Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler. 4. überarb. Aufl., [Nachdr.] Berlin: Springer-Verl. 2009.
- [7] Harrington D, D'Agostino RB, Gatsonis C, *et al.* New Guidelines for Statistical Reporting in the Journal. *N Engl J Med* 2019; 381(3): 285–6  
[<https://doi.org/10.1056/NEJMe1906559>][PMID: 31314974]
- [8] Fethney J. Statistical and clinical significance, and how to use confidence intervals to help interpret both. *Aust Crit Care* 2010; 23(2): 93–7  
[<https://doi.org/10.1016/j.aucc.2010.03.001>][PMID: 20347326]
- [9] Field A. *Discovering Statistics Using SPSS*. 3.ed. London: Sage 2009.
- [10] Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates 1988.
- [11] Sullivan GM, Feinn R. Using Effect Size-or Why the P Value Is Not Enough. *J Grad Med Educ* 2012; 4(3): 279–82  
[<https://doi.org/10.4300/JGME-D-12-00156.1>][PMID: 23997866]
- [12] Weiß C, Rzany B. *Basiswissen Medizinische Statistik*. 4., überarbeitete Auflage. Berlin, Heidelberg: Springer Berlin Heidelberg 2008.



- [13] Kim H-Y. Statistical notes for clinical researchers: effect size. *Restor Dent Endod* 2015; 40(4): 328–31  
[<https://doi.org/10.5395/rde.2015.40.4.328>][PMID: 26587420]
- [14] Anlauf M. Beurteilung von Therapien mit der „number needed to treat“. *Deutsches Ärzteblatt* 2006; 103(48): 3254–8.
- [15] Citrome L, Ketter TA. When does a difference make a difference? Interpretation of number needed to treat, number needed to harm, and likelihood to be helped or harmed. *Int J Clin Pract* 2013; 67(5): 407–11  
[<https://doi.org/10.1111/ijcp.12142>][PMID: 23574101]
- [16] Andrade C. Understanding relative risk, odds ratio, and related terms: as simple as it can get. *J Clin Psychiatry* 2015; 76(7): e857-61  
[<https://doi.org/10.4088/JCP.15f10150>][PMID: 26231012]
- [17] Violan C, Foguet-Boreu Q, Flores-Mateo G, *et al.* Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. *PLoS One* 2014; 9(7): e102149  
[<https://doi.org/10.1371/journal.pone.0102149>][PMID: 25048354]
- [18] Sánchez-Fidalgo S, Guzmán-Ramos MI, Galván-Banqueri M, Bernabeu-Wittel M, Santos-Ramos B. Prevalence of drug interactions in elderly patients with multimorbidity in primary care. *Int J Clin Pharm* 2017; 39(2): 343–53  
[<https://doi.org/10.1007/s11096-017-0439-1>][PMID: 28238102]
- [19] Kühlein T, Maibaum T, Roos M, Klemperer D. Quartäre Prävention - die Verhinderung nutzloser Medizin. In: Kochen MM, editor. *Allgemeinmedizin und Familienmedizin*. 5., vollständig überarbeitete und erweiterte Auflage. Stuttgart: Thieme 2017; 594–9.
- [20] Sekhar MS, Vyas N. Defensive medicine: a bane to healthcare. *Ann Med Health Sci Res* 2013; 3(2): 295–6  
[<https://doi.org/10.4103/2141-9248.113688>][PMID: 23919211]
- [21] Große Feldhaus S, Große Feldhaus J, editors. *Arzt und Recht bei Fehlern und Irrtümern: Für Praxis, Klinik und Begutachtung*. 1. Auflage. München, Deutschland: Elsevier; 2019.
- [22] Amann M. Jede zweite Klage gegen Ärzte hat Erfolg; 2008 [cited 2021 February 7] Available from: URL: <https://www.faz.net/aktuell/karriere-hochschule/recht-und-gehalt/medizinrecht-jede-zweite-klage-gegen-aerzte-hat-erfolg-1726396.html>.

- [23] Tonus C, Dittmar RP. BDC-Umfrage: Wirtschaftlichkeit und Kodierung in Praxis und Krankenhaus. *Passion Chirurgie* 2016; 11(6).
- [24] dpa/aerzteblatt.de. SARS-CoV-2: Umfrage zeigt bedingte Impfbereitschaft; 2020 [cited 2021 May 13] Available from: URL: <https://www.aerzteblatt.de/nachrichten/114996>.
- [25] Zylka-Menhorn V. Patienten nehmen lieber Medikamente ein als Sport zu treiben. *Deutsches Ärzteblatt* 2018; 115(16): A764.
- [26] Chong CC. Pros and cons of clinical practice based on guidelines. *Hong Kong Med J* 2018; 24(5): 440–1  
[<https://doi.org/10.12809/hkmj185083>][PMID: 30318475]

## 4.2 Abbildungsverzeichnis der Einordnung in den wiss. Kontext

**Abb.1:** Krankheit als Dichotomie gegenüber dem Kontinuum des Krankseins.

## 4.3 Abkürzungsverzeichnis der Einordnung in den wiss. Kontext

AES absolute Effektstärke

ARR absolute Risikoreduktion

EbM Evidenzbasierte Medizin

ES Effektstärke

LL Leitlinie / Leitlinien

NNH Number needed to harm

NNT Number needed to treat

OR Odds Ratio

RRR relative Risikoreduktion

SMD standard mean difference

WMD weighted mean difference

## 5. Publikationsliste

Morgott M, Heinmüller S, Hueber S, Schedlbauer A, Kühlein T. Do guidelines help us to deviate from their recommendations when appropriate for the individual patient? A systematic survey of clinical practice guidelines. *J Eval Clin Pract.* 2020 Jun;26(3):709-717.

Tallner A, Streber R, Hentschke C, Morgott M, Geidl W, Mäurer M, Pfeifer K. Internet-Supported Physical Exercise Training for Persons with Multiple Sclerosis-A Randomised, Controlled Study. *Int J Mol Sci.* 2016 Sep 30;17(10):1667.

## 6. Danksagung

## 7. Lebenslauf