

Aus dem Institut für Allgemeinmedizin  
Klinikum der Ludwig-Maximilians-Universität München



***Einfluss von Patienten- und Praxismerkmalen auf die Patientenzufriedenheit: Eine gepoolte Analyse der PARADIES- und PRoMPT-Studien.***

Dissertation

zum Erwerb des Doktorgrades der Medizin  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München

vorgelegt von

Christian Alexander Beltz

aus

Hamburg

Jahr

2024

---

Mit Genehmigung der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität München

Erstes Gutachten: Prof. Dr. Jochen Gensichen

Zweites Gutachten: Prof. Dr. Peter Falkai

Drittes Gutachten: Prof. Dr. Jakob Linseisen

ggf. weitere Gutachter: \_\_\_\_\_  
\_\_\_\_\_

Mitbetreuung durch den  
promovierten Mitarbeiter: Dr. Susanne Rospleszcz

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 24.06.2024

# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b> .....	<b>3</b>
<b>Zusammenfassung:</b> .....	<b>6</b>
<b>Abstract (English):</b> .....	<b>6</b>
<b>Abbildungsverzeichnis</b> .....	<b>7</b>
<b>Tabellenverzeichnis</b> .....	<b>8</b>
<b>Abkürzungsverzeichnis</b> .....	<b>9</b>
<b>1. Einleitung</b> .....	<b>10</b>
1.1 Kurzer Überblick über meine Arbeit .....	10
1.2 Patientenzufriedenheit.....	10
1.2.1 Das Konstrukt „Patientenzufriedenheit“ .....	10
1.2.2 Patientenmerkmale und Patientenzufriedenheit .....	11
1.2.3 Merkmale der Behandlung und Patientenzufriedenheit .....	12
1.3 Depression.....	13
1.3.1 Begriffserläuterung.....	13
1.3.2 Prävalenz und Inzidenz .....	13
1.3.3 Soziodemographische Faktoren und Komorbiditäten .....	14
1.3.4 Ätiopathogenese und Risikofaktoren .....	15
1.3.5 Bestimmung des Schweregrades depressiver Symptomatik .....	16
1.3.6 Klinik, Diagnostik und Therapie einer Depression .....	16
1.4 Chronic Care Model (CCM) .....	17
1.4.1 Vorgeschichte und Entwicklung des CCM .....	17
1.4.2 Lösungsansätze zur Überwindung der Probleme.....	17
1.4.3 Entwicklung des Chronic Care Model .....	19
1.4.4 Aufbau und Inhalt des Chronic Care Model.....	20
1.4.5 Überprüfung des Chronic Care Model .....	21
1.5 Patient Assessment of Chronic Illness Care (PACIC) .....	22
1.5.1 Versorgungsqualität aus Sicht der Leistungserbringer: Assessment of Chronic Illness Care (ACIC).....	22
1.5.2 Vorgeschichte und Zweck des PACIC .....	23
1.5.3 Entwicklung des PACIC .....	23
1.5.4 Aufbau und Auswertung des PACIC.....	24
1.5.5 Validierung des PACIC.....	24
1.5.6 Kritik am PACIC .....	25
1.5.7 Weiterentwicklung: Der PACIC mit 11 Items .....	26
1.5.8 Validierung der PACIC-Version mit 11 Items .....	27
1.5.9 Vergleich der PACIC-Items und der Skalenbeschriftung .....	28
1.5.10 PACIC-Versionen in anderen Sprachen.....	28
1.5.11 Andere Versionen des PACIC .....	29
1.5.12 Vergleichbarkeit der PACIC-Fragebögen mit 20 und mit 11 Items.....	29
1.5.13 Umgang mit fehlenden Werten des PACIC.....	30
1.5.14 Zusammenhänge zwischen PACIC und Merkmalen .....	31

---

1.6	Primary Care Evaluation of Mental Disorders (PRIME-MD) und PRIME-MD Patient Health Questionnaire (PHQ).....	32
1.6.1	Primary Care Evaluation of Mental Disorders (PRIME-MD).....	32
1.6.2	Entwicklung des PRIME-MD Patient Health Questionnaire (PRIME-MD PHQ).....	33
1.6.3	PHQ-9: Messung der depressiven Symptome .....	35
<b>2.</b>	<b>Material und Methoden .....</b>	<b>37</b>
2.1	Studien .....	37
2.1.1	Jena-PARADIES (Patient Activation foR Anxiety DisordErS) .....	37
2.1.2	PRoMPT (PRimary care Monitoring for depressive Patients Trial).....	38
2.2	Daten und Variablen .....	39
2.2.1	Patienten- und Praxismerkmale.....	39
2.3	Einschlusskriterien .....	40
2.3.1	Flowchart .....	42
2.4	Statistik.....	43
2.4.1	Umgang mit uneinheitlichen PACIC-Versionen .....	43
2.4.2	Patienten- und Praxismerkmale.....	45
2.4.3	Erstellte Variablen .....	46
2.4.4	PACIC-Werte und Sensitivitätsanalyse .....	47
2.4.5	Berücksichtigung potenzieller Störfaktoren und Bewertung von Effektgrößen .....	48
2.4.6	Software .....	48
2.4.7	Umgang mit fehlenden Werten bei PACIC und PHQ-9.....	48
<b>3.</b>	<b>Ergebnisse .....</b>	<b>49</b>
3.1	Deskriptive Auswertung.....	49
3.1.1	Patientenmerkmale .....	49
3.1.2	Praxismerkmale .....	49
3.2	Vorläufige Studienpopulation: Zusammenhang zwischen PACIC-Werten und der Behandlung.....	52
3.3	Endgültige Studienpopulation: Zusammenhang zwischen PACIC-Werten und der Behandlung.....	54
3.4	PACIC und Patientenmerkmale .....	55
3.5	PACIC und Praxismerkmale.....	58
<b>4.</b>	<b>Diskussion.....</b>	<b>63</b>
4.1	Zusammenfassung der Ergebnisse .....	63
4.2	Interpretation der Ergebnisse .....	63
4.2.1	Relevanz für die Hausarztpraxis.....	63
4.2.2	PACIC: Vergleich der Item-Anzahl.....	64
4.2.3	PACIC: Vergleich der Antwort-Skalen.....	65
4.2.4	PACIC: Umgang mit fehlenden Werten.....	66
4.2.5	PACIC: Zusammenhänge mit Patientenmerkmalen .....	67
4.2.6	PACIC: Zusammenhänge mit Praxismerkmalen .....	68
4.2.7	Einschlusskriterien .....	70
4.2.8	Statistik und Aufbereitung der Daten .....	71
4.3	Limitationen der Arbeit .....	72
4.4	Empfehlung für weiterführende Forschung .....	73

**Affidavit ..... 74**

**Danksagung ..... 75**

**Publikationsliste ..... 76**

**5. Literaturverzeichnis ..... 77**

## Zusammenfassung:

Patientenzufriedenheit beschreibt die Zufriedenheit der Patienten mit der Behandlung sowie die Behandlung beeinflussende Faktoren. Dies können neben Patientenmerkmalen auch Merkmale der medizinischen Einrichtung wie zum Beispiel die Unterscheidung in Einzel- oder Gemeinschaftspraxen sein. Der Patient Assessment of Chronic Illness Care (PACIC) ist ein Fragebogen zur Erfassung der Patientenzufriedenheit mit der medizinischen Behandlung. Im Laufe der Zeit wurde der PACIC durch verschiedene Autoren weiterentwickelt, sodass neben der Originalversion mit 20 Items auch eine kürzere Version mit 11 Items existiert. Beide PACIC-Versionen unterscheiden sich ferner hinsichtlich der Skala als auch der Benennung der Endpunkte. In dieser Arbeit wurde anhand von zwei abgeschlossenen Studien der Zusammenhang zwischen der Patientenzufriedenheit und Merkmalen von Patienten und Arztpraxen untersucht. Sowohl die PARADIES- als auch die PRoMPT-Studie untersuchte den Effekt von Case-Management-Programmen für psychiatrische Erkrankungen. Diese Arbeit zeigt, dass die Patientenzufriedenheit weitgehend unabhängig von Merkmalen der Patienten sowie der medizinischen Einrichtungen ist. Das Alter des Studienarztes war negativ mit der Patientenzufriedenheit verbunden, die Patienten jüngerer Studienärzte berichteten über eine höhere Patientenzufriedenheit.

## Abstract (English):

Patient satisfaction describes the satisfaction of the patients with the treatment as well as factors influencing the treatment. In addition to patient characteristics, these can also be characteristics of the medical practice, such as the distinction between individual or group practices. The Patient Assessment of Chronic Illness Care (PACIC) is a questionnaire designed to assess patient satisfaction with medical treatment. Over time, the PACIC has been further developed by various authors, so that in addition to the original version with 20 items, a shorter version with 11 items also exists. Both PACIC versions differ with regard to the scale and the naming of the endpoints. In this paper, the relationship between patient satisfaction and characteristics of patients and medical practices was investigated on the basis of two completed studies. Both the PARADIES and PRoMPT studies examined the effect of case management programs for psychiatric disorders. This work showed that patient satisfaction is largely independent of patient characteristics as well as medical practice characteristics. Study physician age was negatively associated with patient satisfaction, with patients of younger study physicians reporting higher patient satisfaction.

## Abbildungsverzeichnis

Abbildung 1: Depression und Risikofaktoren (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).....	15
Abbildung 2: Bestimmung des Schweregrades depressiver Symptomatik (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).....	16
Abbildung 3: Das Chronic Care Modell (Gensichen und Rosemann 2007).....	20
Abbildung 4: Flowchart für die Einschlusskriterien und Studienpopulation.....	42
Abbildung 5: Box-Plot der Ergebnisse: PACIC und Studienarme .....	52
Abbildung 6: Box-Plot der Ergebnisse: PACIC und Geschlecht.....	55
Abbildung 7: Box-Plot der Ergebnisse: PACIC und Suizidgedanken .....	56
Abbildung 8: Streudiagramm für Ergebnisse: PACIC und Patientenalter .....	57
Abbildung 9: Box-Plot der Ergebnisse: PACIC und Art der Praxis .....	58
Abbildung 10: Box-Plot der Ergebnisse: PACIC und Lage der Praxis.....	59
Abbildung 11: Box-Plot der Ergebnisse: PACIC und Praxisgründung.....	60
Abbildung 12: Box-Plot der Ergebnisse: PACIC und Zusatzqualifikation "Psychosomatische Grundversorgung" .....	61
Abbildung 13: Streudiagramm für Ergebnisse: PACIC und Alter des Studienarztes .....	62
Abbildung 14: PACIC-5As Seite 1 (Rosemann et al. 2007) ... <b>Fehler! Textmarke nicht definiert.</b>	
Abbildung 15: PACIC-5As Seite 2 (Rosemann et al. 2007) ... <b>Fehler! Textmarke nicht definiert.</b>	
Abbildung 16: PACIC-5As Seite 3 (Rosemann et al. 2007) ... <b>Fehler! Textmarke nicht definiert.</b>	
Abbildung 17: PACIC-5As Seite 4 (Rosemann et al. 2007) ... <b>Fehler! Textmarke nicht definiert.</b>	

## Tabellenverzeichnis

Tabelle 1: Gegenüberstellung der PACIC-Items der Originalversion und der Kurzversion .....	27
Tabelle 2: Gegenüberstellung der Stufen der Likert-Skala der englischen Originalversion und der deutschen Version des PACIC.....	28
Tabelle 3: Gegenüberstellung der Skalenstufen des PACIC mit 11 Items bei 0 % und bei 100 % .....	28
Tabelle 4: Gegenüberstellung verschiedener Grenzwerte fehlender PACIC-Items in durchgeführten Studien.....	30
Tabelle 5: Erste Überlegung zur Umrechnung der Prozent-Skala in die Likert-Skala.....	44
Tabelle 6: Zweite Überlegung zur Umrechnung der Prozent-Skala in die Likert-Skala.....	44
Tabelle 7: Lösungsbeispiel anhand von Probanden aus der PARADIES- und der PRoMPT-Studie .....	45
Tabelle 8: Darstellung der Patienten- und Praxismerkmale der PARADIES-Studie, der PRoMPT-Studie sowie der Studienpopulation .....	51
Tabelle 9: Deskriptive Statistik der Ergebnisse: PACIC und Studienarme .....	53
Tabelle 10: Deskriptive Statistik der Ergebnisse: PACIC und Geschlecht.....	55
Tabelle 11: Deskriptive Statistik der Ergebnisse: PACIC und Suizidgedanken .....	56
Tabelle 12: Deskriptive Statistik der Ergebnisse: PACIC und Art der Praxis .....	58
Tabelle 13: Deskriptive Statistik der Ergebnisse: PACIC und Lage der Praxis.....	59
Tabelle 14: Deskriptive Statistik der Ergebnisse: PACIC und Praxisgründung.....	60
Tabelle 15: Deskriptive Statistik der Ergebnisse: PACIC und Zusatzqualifikation "Psychosomatische Grundversorgung" .....	61
Tabelle 16: PACIC-Version mit 11 Items und 11 Skalenstufen als Prozentskala (Goetz et al. 2012).....	<b>Fehler! Textmarke nicht definiert.</b>



## Abkürzungsverzeichnis

ACIC	Assessment of Chronic Illness Care
CCM	Chronic Care Model
CEG	Clinical evaluation guide
COPD	Chronisch obstruktive Lungenerkrankung
DEGS1-MH	Studie zur Gesundheit Erwachsener in Deutschland
DeMoL	Depressions-Monitoring-Liste
DSM-III-R	Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition
ICD	International Statistical Classification of Diseases and Related Health Problems
ISRCTN	International Standard Randomised Controlled Trial Number
JAMoL	Jena Angst-Monitoring-Liste
KHK	Koronarer Herzkrankheit
KVT	Kognitive Verhaltenstherapie
MFA	Medizinische/r Fachangestellte/r
NAKO	Ehemals: Nationale Kohorte
O-PACIC	Older Patient Assessment of Chronic Illness Care
PACIC	Patient Assessment of Chronic Illness Care
PARADIES	Patient Activation foR Anxiety DisordErS
PRIME-MD	Primary Care Evaluation of Mental Disorders
PHQ-D	Gesundheitsfragebogen für Patienten
PHQ-9	Patient Health Questionnaire: Depressionsmodul
PRIME-MD PHQ	PRIME-MD Patient Health Questionnaire, auch: Patient Health Questionnaire (PHQ)
PRoMPT	PRimary care Monitoring for depressive Patient's Trial

# 1. Einleitung

## 1.1 Kurzer Überblick über meine Arbeit

Diese Forschungsarbeit soll den Einfluss von Patienten- und Praxismerkmalen auf die Zufriedenheit der Patienten mit der ärztlichen Versorgung untersuchen. Die Studienpopulation bildeten depressive Patienten. Die Zufriedenheit mit der Versorgung wurde mithilfe des PACIC-Fragebogens erfasst. Die Studiendaten stammen aus der PARADIES- und der PRoMPT-Studie, die im Methodenteil näher vorgestellt werden.

Die Forschungsergebnisse dieser Arbeit wurden bereits in einem Fachjournal veröffentlicht:

Lukaschek K, Beltz C, Rospleszcz S, Schillok H, Falkai P, Margraf J, Gensichen J. Depressive primary care patients' assessment of received collaborative care. Sci Rep. 2023 Feb 9;13(1):2329. doi: 10.1038/s41598-023-29339-9. PMID: 36759622; PMCID: PMC9911390.

Scientific reports weist für 2021 einen Impact Factor von 4.996 auf.

Als Teil der Forschungsgruppe wurde die Datenkuratierung von mir durchgeführt. Zusammen mit anderen Teammitgliedern war ich an folgenden Aufgaben beteiligt: statistische Auswertung, Erarbeiten der Methodik, Schreiben des Paperentwurfs sowie Verfassen von Stellungnahmen im Rahmen des Reviewprozess des Papers.

## 1.2 Patientenzufriedenheit

*„Patientenzufriedenheit ist ein multidimensionaler Begriff, der verschiedene Aspekte und Komponenten des Versorgungsprozesses betrifft und zugleich bestimmt wird von keineswegs einheitlichen Maßstäben, die Patienten an ihre Behandlung anlegen. Obwohl ihre Erwartungen weitgehend vom „medizinischen System“ selbst [...] geprägt werden, variieren sie je nach Individuum, Situation, Versorgungsbedarf – alle Patienten wollen zu Recht gute Qualität, aber worin sie besteht und worauf es dabei ankommt, ist nicht für alle gleich.“ (Blum et al. 2001, S.29)*

### 1.2.1 Das Konstrukt „Patientenzufriedenheit“

Die Patientenzufriedenheit ist ein von der Kundenzufriedenheit abgeleitetes Konstrukt (Jacob und Bengel 2000; Neugebauer und Porst 2001).

Jacob und Bengel beschreiben in ihrer Arbeit die Evolution der Patientenzufriedenheit seit den 50er Jahren des 20. Jahrhunderts ausgehend von der Kundenzufriedenheit. So hätten gesundheitspolitische und ökonomische Bedingungen das subjektive Befinden des Patienten zu einem Qualitätskriterium gemacht. Der Arztbesuch würde vermehrt als Dienstleistung wahrgenommen und die Zufriedenheit sei abhängig von den Erwartungen (Jacob und Bengel 2000).

Laut Neugebauer und Porst habe sich die Sichtweise auf die Patientenzufriedenheit komplett verändert. Waren Patienten früher „mehr oder weniger ‚willenlos‘“ (Neugebauer und Porst 2001) dem medizinischen Personal ausgeliefert, würden Patienten nun als "Kunden" betrachtet, die sowohl mit guten

medizinischen Ergebnissen als auch mit gutem Service zufrieden sein sollen. Nach allgemeinen Definitionen sei Zufriedenheit eine emotionale Reaktion auf ein Ergebnis eines Soll-Ist-Vergleichs. Ausgehend von der Untersuchung der subjektiv wahrgenommenen Leistung des Anbieters (Ist-Zustand) im Vergleich zu einer erwarteten Leistung wird dieser Vergleich durchgeführt (Soll-Zustand). Dieser Vergleich führt entweder zu einer Bestätigung oder einer Nicht-Bestätigung, wobei letztlich der Grad dieser Ausprägung die Patientenzufriedenheit bestimmt (Neugebauer und Porst 2001).

Um das Konstrukt der Patientenzufriedenheit zu erforschen, führten Hall und Dornan 1988 eine Meta-Analyse über die Zufriedenheit von Patienten mit der medizinischen Versorgung durch. Eingeschlossen wurden 107 Studien, welche die Beziehungen zwischen der Zufriedenheit und anderen Variablen untersuchten. Die wichtigsten Aspekte der Patientenzufriedenheit waren Menschlichkeit und technische Qualität der medizinischen Versorgung. Die Autoren gaben zu bedenken, dass Patientenzufriedenheit aufgrund des Einflusses mehrerer Variablen ein mehrdimensionales Konstrukt ist. Daraus folgt, dass Unterschiede zwischen den Patienten in Bezug auf Werte und Erfahrungen die Studienergebnisse beeinflussen. Insgesamt wurden in der Untersuchung die folgenden zwölf Dimensionen untersucht:

- Gesamtzufriedenheit
- Zugang (Bequemlichkeit, Öffnungszeiten; Entfernung; Leichtigkeit, Termine zu bekommen)
- Kosten
- Gesamtqualität
- Menschlichkeit (Wärme, Respekt, Freundlichkeit etc.)
- Kompetenz (technische Leistung etc.)
- Aufklärung und Information (Erklärungen zu Behandlungen, Therapie oder Diagnosen)
- Organisation (Wartezeit etc.)
- Ausstattung der medizinischen Einrichtung
- Empathie
- Kontinuität der Betreuung
- Behandlungsergebnis

Angemerkt werden muss dabei, dass Studien mit psychiatrischen Patienten aus dem Pool der zu untersuchenden Studien ausgeschlossen wurden (Hall und Dornan 1988).

Hsieh und Kagle beschreiben die Patientenzufriedenheit als ein komplexes Phänomen, das mit den Erwartungen der Patienten, ihrem Gesundheitszustand, persönlichen Merkmalen sowie mit Merkmalen des Gesundheitssystems zusammenhängt. In ihrer Untersuchung wurde die Hypothese, dass die Erwartung der Patienten der beste Prädiktor für die Zufriedenheit ist, bestätigt (Hsieh und Kagle 1991).

### **1.2.2 Patientenmerkmale und Patientenzufriedenheit**

Jacob und Bengel führten in ihrer Arbeit (siehe oben) an, dass die Patientenzufriedenheit mit dem Patientenalter ansteige sowie unabhängig vom Geschlecht zu sein schien. Ferner wird ausgeführt, dass Patienten mit einer höheren Zufriedenheit eine geringere Depressivität angeben (Jacob und Bengel 2000).

Die Meta-Analyse von Hsieh und Kagle hatte gezeigt, dass ältere Patienten eher mit der Gesundheitsversorgung zufrieden sind als jüngere Patienten. Eine Abhängigkeit der Patientenzufriedenheit vom Geschlecht des Patienten konnte dagegen nicht festgestellt werden. Insgesamt seien jedoch die soziodemografischen Faktoren keine starken Prädiktoren für die Patientenzufriedenheit (Hsieh und Kagle 1991).

In einer weiteren Meta-Analyse untersuchten Hall und Dornan Zusammenhänge zwischen soziodemographischen Patientenmerkmalen und ihrer Zufriedenheit mit der medizinischen Versorgung. Hierbei war eine höhere Zufriedenheit signifikant mit höherem Alter und geringerer Bildung verbunden. In Hinblick auf die Patientenzufriedenheit gab es keine geschlechterspezifischen Unterschiede. Generell seien die soziodemographischen Patientenmerkmale die am häufigsten untersuchten Variablen in Hinblick auf die Zufriedenheit (Hall und Dornan 1990).

### **1.2.3 Merkmale der Behandlung und Patientenzufriedenheit**

Batbaatar et al. untersuchten systematisch die Evidenz zu den Determinanten der Patientenzufriedenheit. In ihre Meta-Analyse wurden Studien eingeschlossen, die zwischen Januar 1980 und August 2014 vollständig in begutachteten Fachzeitschriften veröffentlicht wurden und in englischer Sprache abgefasst waren. Dabei wurden 109 Artikel in die Auswertung eingeschlossen (Batbaatar et al. 2017).

Die Studienergebnisse waren hinsichtlich des Einflusses von Faktoren auf die Patientenzufriedenheit unterschiedlich, da es keine weltweit akzeptierte Formulierung der Patientenzufriedenheit und kein einheitliches Messsystem gab. Indikatoren für die Qualität der Gesundheitsversorgung waren in allen Studien die einflussreichsten Determinanten der Patientenzufriedenheit. Unter ihnen war die Qualität der zwischenmenschlichen Betreuung durch die Gesundheitsdienstleister die wichtigste Determinante für die Patientenzufriedenheit. Die soziodemografischen Merkmale waren in der Untersuchung am vielfältigsten (Batbaatar et al. 2017).

Auch an dieser Stelle ist festzuhalten, dass Batbaatar et al. Studien zur Patientenzufriedenheit mit bestimmten Gesundheitseinrichtungen (Psychiatrie, Rehabilitation usw.) oder Krankheiten (Psychosen, Krebserkrankungen, Transplantationen usw.) ausschlossen (Batbaatar et al. 2017).

Hinsichtlich des Einflusses von demografischen und psychologischen Merkmalen auf die Patientenzufriedenheit wurden 71 Studien identifiziert. Die Merkmale waren Alter, Geschlecht, Bildung, sozioökonomischer Status, Familienstand, Ethnie, Religion, geografische Merkmale, Häufigkeit der Besuche, Aufenthaltsdauer, Gesundheitszustand, Persönlichkeit und Erwartungen der Patienten. Die Ergebnisse der Beziehungen zwischen patientenbezogenen Merkmalen und der Patientenzufriedenheit waren dabei schwach, weitgehend uneinheitlich und widersprüchlich in der gesamten Stichprobe (Batbaatar et al. 2017).

Die Mehrzahl der Studien kam zu dem Schluss, dass ältere Patienten mit den Gesundheitsdienstleistungen zufriedener waren als jüngere Patienten. Die Ergebnisse des Zusammenhangs zwischen Geschlecht und Patientenzufriedenheit waren in den einzelnen Studien sehr unterschiedlich (Batbaatar et al. 2017).

Szecsényi et al. untersuchten, ob es einen Zusammenhang zwischen der Patientenzufriedenheit und der Zufriedenheit der Praxisteams in der Primärversorgung gibt. Hierbei wurden 676 Arztpraxen in Deutschland in die Untersuchung einbezogen. Es zeigte sich, dass die Anzahl an Patienten pro Arztpraxis negativ mit der Zufriedenheit der nicht-ärztlichen Praxismitglieder sowie der Zufriedenheit der Patienten zusammenhing. Schlussfolgern lässt sich daraus, dass Patienten und Mitglieder von Praxisteams Praxen mit einer geringeren Patientenzahl vorziehen (Szecsényi et al. 2011).

Victoor et al. untersuchten Faktoren, die die Wahl eines Patienten für eine Arztpraxis beeinflussen. In ihrem Review kommen die Autoren zu dem Schluss, dass die demografischen Parameter eines Arztes zwar die Wahl beeinflussen, dass aber andere Faktoren in der Regel als wichtiger empfunden werden (Victoor et al. 2012).

## **1.3 Depression**

Da es sich bei der Studienpopulation um depressive Patienten handelt, werden nachfolgend Merkmale der Erkrankung Depression erläutert.

### **1.3.1 Begriffserläuterung**

Depression ist eine psychische Erkrankung, die sich durch eine längere Phase ausgeprägter gedrückter Stimmung, Interessenverlust und Antriebsminderung auszeichnet. Sie ist häufig mit einer Reihe von körperlichen Beschwerden verbunden. Die Krankheit Depression beeinträchtigt in der Regel die gesamte Lebensqualität eines depressiven Menschen. Sie leiden unter extremen Selbstzweifeln, Aufmerksamkeitsproblemen, Grübelneigung, Schlaf- und Essstörungen und finden es schwierig oder unmöglich, tägliche Aufgaben zu erledigen. Da diese Erkrankung erhebliche Auswirkungen auf das Wohlbefinden und das Selbstwertgefühl der Patienten hat, sind Depressionen meist mit einem hohen Leidensdruck verbunden (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).

### **1.3.2 Prävalenz und Inzidenz**

Depressionen sind sowohl für den Einzelnen als auch für die Gesellschaft von großer Bedeutung. Sie gehören zu den am weitesten verbreiteten, aber auch am meisten unterschätzten Krankheiten. Sowohl im nationalen als auch im globalen Kontext wird die Lebenszeitprävalenz von Depressionen (aller Arten) auf 16-20 % geschätzt. Nach der repräsentativen Studie zur Gesundheit Erwachsener in Deutschland (DEGS1) und einer Analyse der ebenfalls repräsentativen NAKO-Gesundheitsstudie lag das Lebenszeitrisiko, an einer depressiven Störung zu erkranken, in Deutschland bei 11,6 % bzw. 15,9 % (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).

Der DEGS-Studie zufolge lag die 12-Monats-Prävalenz der Major Depression bei 6,8 % und die der unipolaren Depression bei 8,2 % in der Allgemeinbevölkerung. Dies entspricht fast 5,3 Millionen Patienten mit unipolarer Depression in Deutschland im Laufe eines Jahres. Obwohl die 12-Monats-Prävalenz depressiver Episoden im Vergleich zum Bundesgesundheitsurvey 1998 gleichblieb, zeigt sich

vor allem bei jüngeren Frauen eine höhere Prävalenz. Zudem wurde eine stärkere Symptomschwere depressiver Episoden gemessen (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).

### 1.3.3 Soziodemographische Faktoren und Komorbiditäten

Die Prävalenz und Inzidenz depressiver Störungen werden durch eine Reihe soziodemografischer Merkmale beeinflusst. So werden depressive Störungen bei Frauen fast doppelt so häufig diagnostiziert wie bei Männern, und die Krankheit tritt bei Frauen früher auf (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).

Depressionen können in jedem Alter auftreten. Dem Bundesgesundheitsurvey zufolge traten bei 50 % aller deutschen Patienten die ersten Symptome vor dem 31. Lebensjahr auf. Die Raten steigen bei jüngeren Altersgruppen sowie bei Kindern und Jugendlichen. Depressionen sind die häufigste psychische Erkrankung bei älteren Menschen. Es wird angenommen, dass die Prävalenz von Depressionen in stationären Altenpflegeeinrichtungen 50 % erreicht. Die höchste Suizidprävalenz findet sich ebenfalls bei sehr alten Menschen, insbesondere bei Männern (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).

Andere psychische Erkrankungen und depressive Störungen treten häufig gemeinsam auf. Besonders häufig leiden zum Beispiel depressive Menschen unter Angst- und Panikstörungen. Depressionen treten häufig auch zusammen mit krankhaftem Konsum von Alkohol, Medikamenten und Drogen, mit somatoformen Störungen, Ess-, Persönlichkeits- und Zwangsstörungen auf. Komorbide Erkrankungen erhöhen die Wahrscheinlichkeit einer Chronifizierung, bedingen eine schlechtere Prognose und erhöhen das Suizidrisiko der Patienten (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).

In der S3-Leitlinie zur Behandlung von Angststörungen wird die 12-Monats-Prävalenz einer Panikstörung mit/ohne Agoraphobie mit Verweis auf die DEGS1-Studie mit 2 % angegeben (Männer: 1,2 % und Frauen 2,8 %) (Bandelow et al. 2021).

### 1.3.4 Ätiopathogenese und Risikofaktoren

Es existiert kein einheitliches Krankheitsbild der Depression. Vereinfachte biologische und psychosoziale Modellvorstellungen versuchen die Entstehung einer Depression zu erklären, doch keiner dieser Ansätze hat sich als überzeugende Erklärung erwiesen. Hinter multifaktoriellen Erklärungsansätzen steht die Annahme, dass biologische und psychosoziale Faktoren zusammenwirken, wobei das relative Gewicht der einzelnen Aspekte von Mensch zu Mensch stark variiert. Eine Auflistung verschiedener Risikofaktoren ist *Abbildung 1* zu entnehmen (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).

#### biologische Faktoren

- genetische Vulnerabilität (depressive Störungen in der Familiengeschichte, insbesondere bei Angehörigen ersten Grades)
- körperliche Risikofaktoren, v. a. metabolische Risikofaktoren, Stoffwechselstörungen, Adipositas, Infektionen, chronische Erkrankungen (siehe auch Kapitel 11 Komorbidität)
- hormonelle Umstellung in der Pubertät, Schwangerschaft/Wochenbett, Perimenopause

#### soziodemografische Faktoren

- Geschlecht (vgl. Kapitel 10.3 Genderspezifische Aspekte)
- höheres Alter (vgl. Kapitel 10.1 Ältere Patient\*innen)
- Zugehörigkeit zu einer ethnischen Minderheit (vgl. Kapitel 10.2 Kultur- und migrationspezifische Aspekte)
- niedriger sozioökonomischer Status, Armut

#### psychische Faktoren

- andere psychische Störungen, v. a. Angst-, Substanzgebrauchs- und Persönlichkeitsstörungen

#### psychosoziale Risikofaktoren

- Exposition gegenüber Traumata/Katastrophen (z. B. Missbrauch, Misshandlung, Vernachlässigung oder Verlusterlebnisse in der Kindheit, Krieg, Gewalt)
- Vereinsamung, Verwitwung, soziale Isolation
- aktuell belastende Lebensereignisse (z. B. interpersonelle Konflikte, Beziehungskrisen, Trennungen, Todesfälle, berufliche Enttäuschungen, Diagnose einer schweren Erkrankung u. a.)
- chronischer Stress, Burnout, Überforderung

#### Lebensstilfaktoren

- Ernährung, Rauchen, Bewegungsmangel

**Abbildung 1: Depression und Risikofaktoren (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022)**

### 1.3.5 Bestimmung des Schweregrades depressiver Symptomatik

In *Abbildung 2* sind der Algorithmus zur Bestimmung des Schweregrades der depressiven Symptomatik aufgeführt.

<b>Hauptsymptome</b>	≥ 2	≥ 2	3
<ul style="list-style-type: none"> <li>▪ depressive Stimmung, in einem für die Betroffenen deutlich ungewöhnlichen Ausmaß, die meiste Zeit des Tages, fast jeden Tag, im Wesentlichen unbeeinflusst von den Umständen</li> <li>▪ Interessen- oder Freudeverlust an Aktivitäten, die normalerweise angenehm sind</li> <li>▪ verminderter Antrieb oder gesteigerte Ermüdbarkeit</li> </ul>			
	+	+	+
<b>Zusatzsymptome</b>	≥ 1	≥ 3	≥ 5
<ul style="list-style-type: none"> <li>▪ Verlust des Selbstvertrauens oder des Selbstwertgefühls</li> <li>▪ unbegründete Selbstvorwürfe oder ausgeprägte, unangemessene Schuldgefühle</li> <li>▪ wiederkehrende Gedanken an den Tod oder an Suizid oder suizidales Verhalten</li> <li>▪ Klagen über oder Nachweis eines verminderten Denk- oder Konzentrationsvermögens, Unschlüssigkeit oder Unentschlossenheit</li> <li>▪ psychomotorische Agitiertheit oder Hemmung (subjektiv oder objektiv)</li> <li>▪ Schlafstörungen jeder Art</li> <li>▪ Appetitverlust oder gesteigerter Appetit mit entsprechender Gewichtsveränderung</li> </ul>			
<b>Summe der Symptome</b>	4–5	6–7	≥ 8
<b>Dauer der Symptome ≥ 2 Wochen</b>	↓	↓	↓
<b>Schweregrad nach ICD-10-GM</b>	<b>leicht</b>	<b>mittelgradig</b>	<b>schwer</b>
<b>ohne psychotische Symptome</b>	F32/33.0	F32/33.1	F32/33.2
<b>mit psychotischen Symptomen</b>	–	–	F32/33.3

**Abbildung 2: Bestimmung des Schweregrades depressiver Symptomatik (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022)**

### 1.3.6 Klinik, Diagnostik und Therapie einer Depression

Für weiterführende Informationen zum Thema Depression sowie zu Klinik, Diagnostik und Therapie einer Depression sei an dieser Stelle an die „Nationale VersorgungsLeitlinie der Unipolaren Depression“ verwiesen (Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V. et al. 2022).



## 1.4 Chronic Care Model (CCM)

### 1.4.1 Vorgeschichte und Entwicklung des CCM

Der amerikanische Gesundheitswissenschaftler Edward Wagner schrieb 1998 in einem Artikel, dass die Primärversorgung zunehmend unter der Behandlung alter und chronisch kranker Patienten leidet (Wagner 1998). Trotz der hohen Betreuungsintensität chronischer Patienten war deren Behandlung oft nicht effektiv, die Kontrolle über die Krankheiten oft nicht gegeben und die Patienten schlussendlich unzufrieden (Wagner 1997; Wagner et al. 2001). Da die Betreuung von Patienten mit chronischen Krankheiten gleichzeitig immer stärker einem Kostendruck und einer erhöhten Leistungsüberwachung unterlag, kam es zu starken Bemühungen im Gesundheitssystem, die vorliegenden Missstände zu überwinden (Wagner et al. 1999).

Die Primärversorgung war bisher dafür ausgelegt, einen kranken Patienten mit einem akuten Problem möglichst schnell wieder gesund zu machen. Patienten sollten zeitnah einen Termin bekommen, anhand ihrer Symptomatik zügig eine Diagnostik durchlaufen und therapiert werden. Dies stellte für Patienten mit akuten Problemen eine effiziente Gesundheitsversorgung dar (Wagner et al. 1996). Laut Wagner lag der Fokus auf akuten Erkrankungen und stand somit im Widerspruch zu den Bedürfnissen chronisch kranker Patienten (Wagner 1997).

Akute Notfälle verdrängten viel zu oft die Bedürfnisse komplexer chronischer Krankheiten wie Hypertonie, Depression, Diabetes oder Asthma (Wagner 1998). Bodenheimer et al. fassten diese Situation als „Tyrannie der Dringlichkeit“ zusammen (Bodenheimer et al. 2002).

Ein weiterer Kritikpunkt war, dass die Rolle der Patienten meistens passiv war. Dies schien bei einem Patienten mit einer akuten Krankheit von Tagen oder Wochen kein Problem zu sein. Da ein Patient mit einer chronischen Krankheit lebenslang von ihr betroffen ist, der behandelnde Arzt sich jedoch nur wenige Stunden im Jahr mit dem Patienten beschäftigt, sollte der Patient befähigt werden, selbst zum eigenen Manager seiner Gesundheit zu werden. Der Aufbau in der Primärversorgung machte es Patienten aber meistens schwer, die eigene Krankheit selbst zu managen (Wagner et al. 2001).

Laut Wagner benötigen chronische Patienten und deren Krankheiten ein Gesundheitssystem, das dauerhaft auf das Management chronischer Krankheiten ausgerichtet ist, um Komplikationen zu vermeiden und das Leben der Patienten nachhaltig zu verbessern (Wagner 1998).

### 1.4.2 Lösungsansätze zur Überwindung der Probleme

Die Evidenz zeigte, dass kollaborative, zielorientierte Selbst-Management-Programme das Gesundheitsverhalten der Patienten und damit das Behandlungsergebnis stärker verbesserten als die traditionelle, didaktische Aufklärung des Patienten (Wagner 2019). Alle erfolgreichen Programme zur Behandlung chronischer Krankheiten boten eine Art von Bildungsprogramm an, um den Bedarf der Patienten und Betreuer an Informationen, Verhaltenskompetenzen und psychosozialer Unterstützung zu decken (Wagner 1997).

Randomisierte klinische Studien hatten bereits gezeigt, dass spezielle Interventions-Programme zu deutlich besseren Behandlungsergebnissen führten. Wagner war deshalb der Meinung, dass man das ambulante System grundlegend analysieren und entsprechend verbessern bzw. umbauen müsste (Wagner 1998). Von Korff et al. schrieben, dass eine Kluft zwischen der durch Studien identifizierten Verbesserungsansätze der Behandlung und der tatsächlichen Behandlung eben dieser Patienten bestand. Diese Diskrepanz erfordere eine Neubewertung der derzeitigen Organisation der Gesundheitsversorgung, damit chronisch kranke Patienten mit größerer Wahrscheinlichkeit Leistungen erhalten, die ihnen helfen, so gut wie möglich und so lange wie möglich zu leben (Korff et al. 1997).

Wagner et al. und Korff et al. identifizierten vier wesentliche Elemente, die erfolgreiche Programme ausmachen:

#### 1. Gemeinsame Problemdefinition:

Ein gemeinsames Management beginnt, wenn Patienten und Behandler die Probleme klar definieren. Die Behandler definieren Probleme in der Regel anhand der Diagnose, der mangelnden Therapietreue oder der Fortführung ungesunder Verhaltensweisen (Rauchen, Bewegungsmangel). Die Patienten definieren Probleme eher in Form von Schmerzen, Symptomen, Funktionseinschränkungen, emotionaler Belastung, Schwierigkeiten bei der Durchführung von Behandlungen oder Lebensstiländerungen sowie Ängsten vor unvorhersehbaren gesundheitlichen Folgen der Krankheit (Wagner et al. 1996; Korff et al. 1997).

#### 2. Zielsetzung und Planung:

Die Konzentration auf ein bestimmtes Problem, die Festlegung realistischer Ziele und die Entwicklung eines Aktionsplans zur Erreichung dieser Ziele sind nützliche Schritte im Umgang mit chronischen Krankheiten. Die Zielsetzung ist von entscheidender Bedeutung, da Patienten und Leistungserbringer oft viele Veränderungen zeitgleich einleiten, was zu mangelnder Therapietreue und Entmutigung führen kann. Sobald Patienten und Ärzte ein Problem definieren und sich über dessen Bedeutung einig sind, können sie ein Verhaltensziel und einen Aktionsplan zur Erreichung dieses Ziels aufstellen, der es Patienten und Ärzten ermöglicht, die Fortschritte zu überwachen (Wagner et al. 1996; Korff et al. 1997).

#### 3. Selbstmanagementschulungen und Unterstützungsdienste:

Die Fähigkeit der Patienten, für sich selbst zu sorgen, wird durch Dienstleistungen verbessert, die die notwendigen Fähigkeiten zur Durchführung medizinischer Behandlungen vermitteln, eine Änderung des Gesundheitsverhaltens anleiten und emotionale Unterstützung bieten. Da die Bedürfnisse der Patienten unterschiedlich sind, sollte eine Vielzahl von Diensten zur Verfügung stehen, die von minimalen Selbsthilfensätzen bis zu intensivem Case Management reichen. Es sei wichtig, dass diese Arten von Dienstleistungen individuell gestaltet werden, auf die Motivation und Bereitschaft jedes Patienten zugeschnitten sind und sich an den Prioritäten orientieren, die zwischen Patienten und Anbieter vereinbart wurden (Wagner et al. 1996; Korff et al. 1997).

#### 4. Aktive und nachhaltige Nachsorge:

Ein geplanter und über einen längeren Zeitraum andauernder Kontakt mit den Behandlern des Gesundheitswesens verbessere die Ergebnisse der Patienten. Durch die Kontaktaufnahme mit den Patienten

in bestimmten Abständen können die Leistungserbringer Informationen über den medizinischen und funktionellen Status einholen, potenzielle Komplikationen frühzeitig erkennen, die Fortschritte bei der Umsetzung des Behandlungsplans überprüfen, notwendige Änderungen vornehmen und die Bemühungen der Patienten unterstützen (Wagner et al. 1996; Wagner 1997).

### **1.4.3 Entwicklung des Chronic Care Model**

Es wurde versucht, die in Studien nachgewiesenen Praxis- und Systemänderungen zu identifizieren und zu einem Modell zusammenzufassen, um die Behandlung chronischer Krankheiten zu verbessern (Wagner et al. 1999; Wagner et al. 2001). Mit Unterstützung der Robert Wood Johnson Foundation wurde ein evidenzbasierter Leitfaden – das Chronic Care Model – zur Verbesserung der Behandlung chronischer Krankheiten entwickelt. Dieser Leitfaden stimmt mit den Forschungsergebnissen überein und ermöglicht die Behandlung diverser chronischer Krankheiten (Wagner et al. 2001). Das Modell wurde mithilfe von Forschern, praktischen Ärzten, Krankenschwestern, Health Administrators und Patientenvertretern überarbeitet und 1998 veröffentlicht (Wagner 2019). Das erarbeitete CCM wurde im nationalen Konzept „Improving Chronic Illness Care“ in zahlreichen speziellen Programmen zur Verbesserung der kollaborativen chronischen Behandlung eingesetzt und getestet (Wagner et al. 2001).

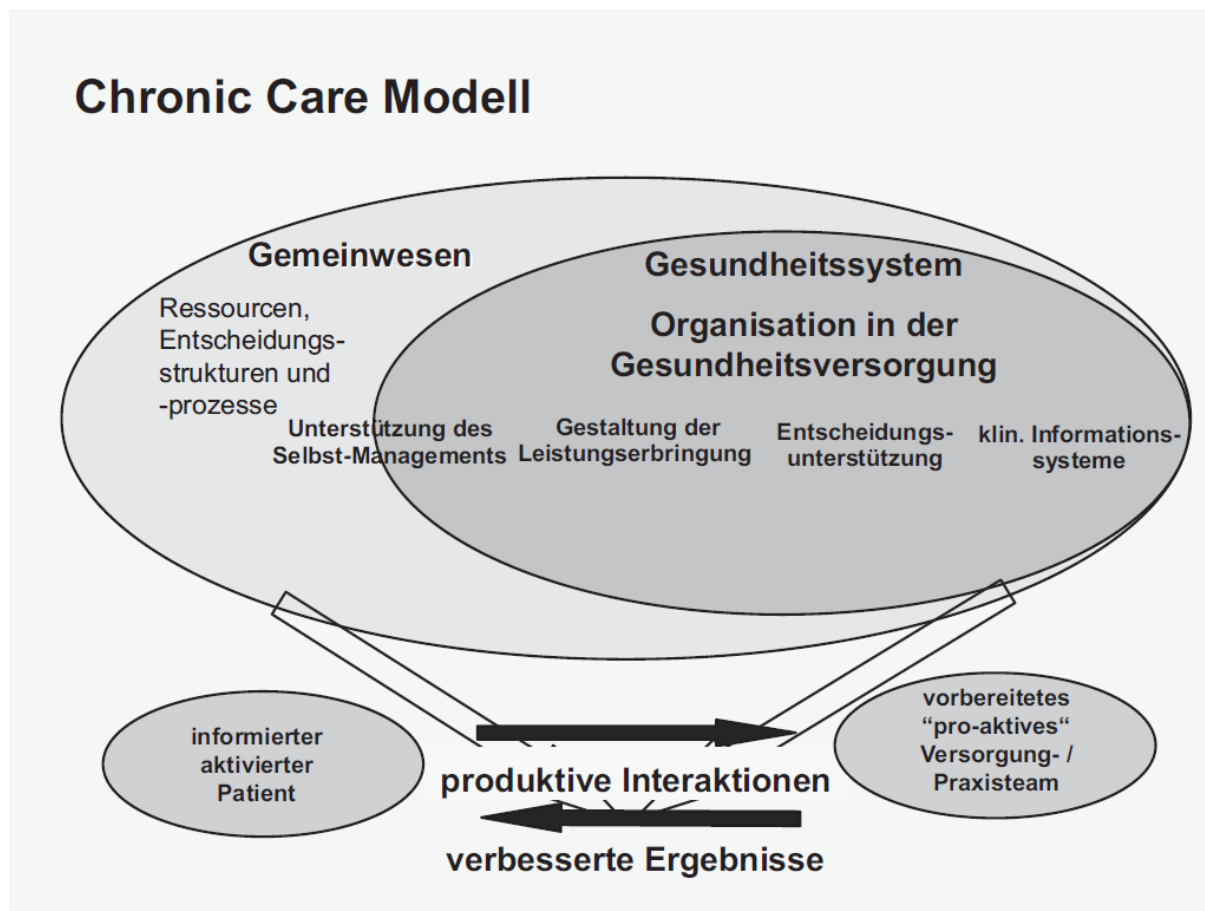
Das überarbeitete Modell betrachtete das Gesundheitssystem als Teil einer größeren Gemeinschaft. Eine wirksame Behandlung chronischer Krankheiten erfordere ein angemessen organisiertes Gesundheitssystem, das mit den notwendigen Ressourcen in der breiteren Gemeinschaft verbunden ist. Das Gesundheitssystem müsse über die Führung, die Anreize und die Ressourcen verfügen, um die Praxis so zu verändern, dass sie den Bedürfnissen chronisch kranker Patienten gerecht wird (Wagner et al. 1999). Dies konnte auch bedeuten, dass zu Beginn deutlich mehr Geld eingesetzt werden müsste, um langfristig Geld zu sparen (Wagner 1998).

Das Hauptziel war eine effiziente Versorgung chronischer Krankheiten durch ein vorbereitetes Praxisteam, das produktiv mit einem aktiven und informierten Patienten an seiner Krankheit arbeitet (Wagner et al. 1999).

Das Chronic Care Model stellte das Gesundheitssystem als Teil einer größeren Gemeinschaft sowie – unabhängig von der Größe - die Praxis als Teil einer Gesundheitsorganisation dar. Ein wirksames Management chronischer Krankheiten erfordere ein angemessen organisiertes Versorgungssystem, das mit ergänzenden, außerhalb der Organisation verfügbaren Gemeinschaftsressourcen verbunden ist. Das Chronic Care Model stellte die Verbesserungen des Gesundheitssystems und der Arztpraxen dar, welche die Interaktion zwischen Ärzten und Patienten verbessern (Wagner et al. 2001). Das Chronic Care Model ist somit ein Leitfaden für die Entwicklung einer effektiven chronischen Versorgung (Bodenheimer et al. 2002).

### 1.4.4 Aufbau und Inhalt des Chronic Care Model

Der Aufbau des Chronic Care Model ist schematisch in *Abbildung 3* dargestellt.



**Abbildung 3: Das Chronic Care Modell (Gensichen und Rosemann 2007)**

#### 1. Gemeinwesen (Community resources and policies):

Um die Versorgung von chronischen Patienten zu verbessern, müssen Arztpraxen Verbindungen zu gemeindebasierten Ressourcen herstellen, z.B. zu Bewegungsprogrammen, Seniorenzentren und Selbsthilfegruppen. Kommunale Verbindungen - z.B. zu Krankenhäusern, die Patientenschulungen anbieten, oder zu häuslichen Pflegediensten, die Fallmanager zur Verfügung stellen - sind besonders für kleine Arztpraxen mit begrenzten Ressourcen hilfreich (Bodenheimer et al. 2002; Wagner et al. 2001).

#### 2. Organisation in der Gesundheitsversorgung (Health Care Organization):

Die Struktur, die Ziele und die Werte einer Arztpraxis und ihre Beziehungen zu Kostenträgern bilden das Fundament, auf dem die übrigen vier Komponenten des Modells für chronische Krankheiten aufbauen. Wenn die Ziele und Führungskräfte einer Organisation die chronische Versorgung nicht als Priorität betrachten, wird es keine Innovation geben. Inwiefern Arztpraxen für verbesserte Bedingungen von den Kostenträgern finanziell entschädigt werden, hat einen großen Einfluss auf Verbesserungen in der chronischen Versorgung. Verbesserungen werden langfristig eher bestehen, wenn sie die Einnahmen erhöhen oder die Ausgaben senken. Wenn die Kostenträger und Versicherer die Qualität der

chronischen Versorgung nicht honorieren, sind Verbesserungen nur schwer aufrechtzuerhalten (Bodenheimer et al. 2002; Wagner et al. 2001).

### 3. Unterstützung des Selbst-Managements (Self-management support)

Da Menschen viele Jahre lang mit chronischen Krankheiten leben, kann man den meisten Patienten wesentliche Teile des Managements ihrer Krankheit beibringen. Teile dieses Managements - Ernährung, Bewegung, Selbstmessung (z.B. mithilfe von Blutzuckermessgeräten oder Personenwaagen) und Medikamenteneinnahme - liegen in der direkten Kontrolle des Patienten. Die Unterstützung des Selbstmanagements umfasst die Zusammenarbeit mit den Patienten und ihren Familien, um ihnen zu helfen, die Fähigkeiten und das Selbstvertrauen für den Umgang mit den chronischen Krankheiten zu erlangen. So sollen Hilfsmittel für das Selbstmanagement (z.B. Blutdruckmanschetten, Blutzuckermessgeräte, Diäten und Verweise auf kommunale Ressourcen) bereitgestellt sowie routinemäßig das Krankheits-Management von Patienten und deren Angehörigen evaluiert werden (Bodenheimer et al. 2002; Wagner et al. 2001).

### 4. Gestaltung der Leistungserbringung (Delivery System Design):

Die Struktur der medizinischen Praxis muss geändert werden, sodass Praxisteams mit einer klaren Arbeitsteilung gebildet werden und die Akutversorgung von der geplanten Behandlung chronischer Erkrankungen getrennt wird. Nichtärztliches Personal wird geschult, um das Selbstmanagement der Patienten zu unterstützen, regelmäßige Routineaufgaben zu erledigen (z.B. Labortests für Diabetiker, Augen- und Fußuntersuchungen) und eine angemessene Nachsorge sicherzustellen. Letzteres betrifft insbesondere geplante Besuche zur Wiedervorstellung (Bodenheimer et al. 2002; Wagner et al. 2001).

### 5. Entscheidungsunterstützung (Decision Support):

Evidenzbasierte Leitlinien für die klinische Praxis bieten Standards für die optimale Versorgung chronischer Erkrankungen und sollten durch Erinnerungshilfen in die Organisation der Praxis integriert werden. Im Idealfall kann das Fachwissen eines ärztlichen Spezialisten telefonisch eingeholt werden, um eine ärztliche Überweisung an einen Facharzt zu vermeiden. Leitliniengerechte Behandlungen werden mithilfe von Fortbildungsveranstaltungen für Praxisteams verstärkt (Bodenheimer et al. 2002; Wagner et al. 2001).

### 6. Klinische Informationssysteme (Clinical Information Systems):

Computergestützte Informationen haben drei wichtige Funktionen: (1) als Erinnerungssystem für leitliniengerechte Behandlungen und als Überweisungshilfe zu Fachärzten, z.B. für regelmäßige Vorstellungen von Diabetikern beim Augenarzt; (2) als Feedback für die Ärzte, das ihnen den Behandlungserfolg chronischer Patienten zeigt; und (3) als Register für die Planung der individuellen Patientenversorgung und für die Durchführung der Versorgung aller Patienten (Bodenheimer et al. 2002; Wagner et al. 2001).

## 1.4.5 Überprüfung des Chronic Care Model

In einer Meta-Analyse von Tsai et al. wurde überprüft, ob Interventionen, die ein oder mehrere Elemente des CCM einbezogen, zu verbesserten Ergebnissen bestimmter chronischer Krankheiten führten. Es wurden randomisierte und nicht-randomisierte kontrollierte Studien eingeschlossen, welche die

klinischen Ergebnisse sowie die Lebensqualität bei Asthma, Herzinsuffizienz, Depression und Diabetes untersuchten. Insgesamt wurden 112 Studien zur Meta-Analyse eingeschlossen, 33 Studien davon betrafen Verbesserungen für die Behandlung von Patienten mit Depression (Tsai et al. 2005).

Interventionen, die ein oder mehrere Elemente des CCM einbezogen, hatten positive Auswirkungen auf die klinischen Ergebnisse und die Versorgungsprozesse. Die Ergebnisse waren bei allen untersuchten Erkrankungen gleich. Die Auswirkungen auf die Lebensqualität waren uneinheitlich, wobei nur die Studien zu Herzinsuffizienz und Depression einen Nutzen zeigten (Tsai et al. 2005).

In einem Übersichts-Review von Coleman et al. wurden 82 ausgewählte Studien dahingehend überprüft, ob die auf dem Chronic Care Model basierenden Veränderungen die Behandlung chronischer Patienten verbesserten. Die Ergebnisse deuteten darauf hin, dass sich die Qualität der Versorgung und die Behandlungsergebnisse chronischer Krankheiten verbesserten. Diesbezüglich waren die Ergebnisse für die USA - aber auch international - übereinstimmend (Coleman et al. 2009).

## **1.5 Patient Assessment of Chronic Illness Care (PACIC)**

### **1.5.1 Versorgungsqualität aus Sicht der Leistungserbringer: Assessment of Chronic Illness Care (ACIC)**

Der Fragebogen „Assessment of Chronic Illness Care“ (ACIC) wurde entwickelt, um Behandlern dabei zu helfen, verbesserungswürdige Bereiche in der Versorgung chronischer Patienten zu identifizieren. Zudem sollten das Ausmaß und die Art der vorgenommenen Verbesserungen bewertet werden können (Bonomi et al. 2000).

Der ACIC basiert auf den sechs Elementen vom Chronic Care Model (CCM) (Bonomi et al. 2002).

Mithilfe des ACIC wird aus Sicht der Arztpraxen bzw. der Praxisteams versucht, die Behandlung chronischer Patienten sowie die generelle Organisation der Versorgung der Patienten in den Bereichen des Chronic Care Models zu evaluieren (Bonomi et al. 2002).

Die Rückmeldungen der teilnehmenden Teams deuteten darauf hin, dass sie den ACIC als äußerst nützlich empfanden. Als nützlich wurde eingeschätzt, dass mit dem ACIC Bereiche identifiziert werden können, in denen Verbesserungen nötig sind und um anschließend die vorgenommenen Anpassungen zu evaluieren (Bonomi et al. 2002). Die Ergebnisse der Studiengruppe ließen den Schluss zu, dass der ACIC auf Veränderungen reagiert, die sich mit den Bemühungen zur Qualitätsverbesserung in der Gesundheitsversorgung ergeben (Bonomi et al. 2002).

Bei der Erarbeitung des ACIC wurden Verbesserungen der Behandlung von Depressionen als auch für die Erkrankungen Diabetes, Herzinsuffizienz sowie Asthma untersucht (Bonomi et al. 2002).

### 1.5.2 Vorgeschichte und Zweck des PACIC

Nachdem mit dem ACIC bereits ein Instrument entwickelt wurde, mit dem Kliniker und Praxisteams beurteilen können, inwieweit Patienten mit chronischen Erkrankungen eine Versorgung erhalten, die dem Chronic Care Model entspricht, rückte die Patientensicht auf die Behandlung in den Fokus (Glasgow et al. 2005a).

Zum damaligen Zeitpunkt existierten keine vergleichbaren Instrumente, die aus Patientensicht die Behandlung ihrer chronischen Erkrankungen im Sinne des CCM erfassen. Eine Beurteilung der Versorgung aus Patientensicht war notwendig, um konvergente Nachweise für die Pflege zu erbringen. Die Patientenperspektive sollte in die Bewertung der Behandlung mit einbezogen werden, um potenzielle Verzerrungen durch die einseitige Sicht aus Behandlerperspektive zu vermeiden (Glasgow et al. 2005a).

Glasgow et al. setzten sich zum Ziel, ein kurzes Instrument zu entwickeln, das in einer Vielzahl von Einrichtungen des Gesundheitswesens eingesetzt werden könnte und für erwachsene Patienten mit einer oder mehreren chronischen Erkrankungen geeignet sei. Der entwickelte PACIC sollte die Inanspruchnahme von CCM-bezogener Pflege bei chronischen Erkrankungen erheben. Mit dem PACIC werden Patientenberichte darüber erfasst, inwieweit ihre Versorgung in den letzten sechs Monaten mit dem CCM in Einklang stehen (Glasgow et al. 2005a).

### 1.5.3 Entwicklung des PACIC

Der PACIC besteht in seiner ursprünglichen Form aus 20 Items. Diese 20 Items wurden aus einem größeren Pool von 46 Items ausgewählt bzw. modifiziert. Diese Items wurden von einer nationalen Gruppe von Experten für die Versorgung chronischer Patienten sowie Experten für das Chronic Care Model entwickelt. Die 46 Items wurden bereits in einer früheren Stichprobe von 130 Patienten getestet (Glasgow et al. 2005a).

Die fünf PACIC-Skalen ließen sich nicht perfekt auf die sechs CCM-Komponenten abbilden. Glasgow et al. führten an, dass die meisten Patienten für sie unsichtbare Praxisabläufe nicht bewerten könnten. Bei diesen Abläufen handelte es sich um Themen wie klinische Informationssysteme oder die Organisation der Gesundheitsversorgung. Die Experten schlugen Items für jede der fünf PACIC-Skalen vor, überprüften frühere Überarbeitungen der Skalen und gaben Feedback. Aus dem größeren Pool wurden diejenigen Items ausgewählt, die eine ausreichende Variabilität aufwiesen, die von den Patienten problemlos verstanden wurden und die das CCM am besten repräsentierten (Glasgow et al. 2005a).

Der PACIC wurde mithilfe von 283 Probanden validiert, die Patienten eines großen Gesundheitssystems in Washington und Idaho waren. Die 283 Probanden litten unter einer oder mehreren chronischen Erkrankungen. Die häufigsten chronischen Erkrankungen waren Bluthochdruck und Arthritis, gefolgt von Depression, Diabetes, Asthma und Schmerzen. Das Durchschnittsalter lag bei 64 Jahren. 56 % der Probanden waren weiblich, und die meisten Probanden litten unter mindestens zwei chronischen Erkrankungen (Glasgow et al. 2005a).

Die Probanden füllten neben dem PACIC-Fragebogen mit 20 Items, einen Fragebogen zu demografischen Faktoren, eine Skala zur Patientenaktivierung sowie ein Instrument zur Bewertung der Primärversorgung aus. Somit wurde gewährleistet, dass Messleistung, Konstruktvalidität sowie Gültigkeit des PACIC bewertet werden konnten (Glasgow et al. 2005a).

### 1.5.4 Aufbau und Auswertung des PACIC

Der PACIC besteht aus 20 Items, zusammengefasst in fünf Subskalen. Die Subskalen des PACIC entsprechen fünf der sechs Schlüsselkomponenten des Chronic Care Model. Wie bereits weiter oben erwähnt ließen sich die sechs CCM-Komponenten nicht perfekt im PACIC abbilden (Glasgow et al. 2005a).

5 Subskalen: (Glasgow et al. 2005a)

- Items 1-3: Patientenaktivierung (Patient Activation/Involvement)
- Items 4-6: Gestaltung des Versorgungssystems/Entscheidungsunterstützung (Delivery System Design/Decision Support)
- Items 7-11: Zielsetzung (Goal Setting/Tailoring)
- Items 12-15: Problemlösung/kontextbezogene Beratung (Problem solving/Contextual Counseling)
- Items 16-20: Nachbereitung/Koordination (Follow-up/Coordination)

Die Probanden bewerten, wie oft sie die in den einzelnen Items beschriebenen Inhalte in den letzten sechs Monaten erlebt haben. Jedes Item wird auf einer Likert-Skala mit fünf Punkten bewertet, die von „1“ („Almost Never“, in deutscher Version „So gut wie nie“) bis „5“ („Almost Always“, in deutscher Version „Fast immer“) reicht. Die Patienten bewerten die komplette Betreuung sowohl durch das Gesundheits- bzw. Praxisteam als auch die Behandlung durch ihren Hausarzt (Glasgow et al. 2005a; Rosemann et al. 2007).

Jede Skala wird durch einfache Mittelung der in der jeweiligen Skala ausgefüllten Items bewertet, und der Gesamt-PACIC wird durch die Mittelung der Punktzahlen aller 20 Items ermittelt (Glasgow et al. 2005a).

Das Ausfüllen der schriftlichen Version des PACIC dauerte 2-5 Minuten. Die Dauer der telefonischen Befragung war sehr unterschiedlich, betrug aber im Durchschnitt etwa 7-8 Minuten (Glasgow et al. 2005a).

### 1.5.5 Validierung des PACIC

Die interne Konsistenz des gesamten PACIC und der verschiedenen Skalen wurde mit Hilfe des Koeffizienten Alpha bewertet. Es ergab sich ein Cronbach's  $\alpha = 0,93$  für die interne Konsistenz der Gesamtskala. Die einzelnen Subskalen lagen zwischen  $\alpha = 0,77$  (Delivery System Design/Decision Support) und  $\alpha = 0,90$  (Problem Solving/Contextual) (Glasgow et al. 2005a).

Es wurde eine konfirmatorische Faktorenanalyse durchgeführt, um zu bewerten, inwieweit die Items des PACIC die fünf Skalen des CCM repräsentieren. Hierbei wurde die vorgeschlagene 5-Skalen-



Struktur bestätigt. Jedoch ließ das hypothetische Modell korrelierte Faktoren zu, da die Autoren davon ausgingen, dass viele der verschiedenen Verbesserungen - gemäß des Chronic Care Model - miteinander verbunden sind (Glasgow et al. 2005a).

Analysiert mithilfe des Korrelationskoeffizienten nach Pearson war der PACIC nur geringfügig mit dem Alter ( $r = 0,16$ ), dem weiblichen Geschlecht ( $r = 0,24$ ) sowie der Anzahl der chronischen Erkrankungen ( $r = 0,13$ ) korreliert. Es gab keine Unterschiede in Bezug auf Bildung, ethnische Zugehörigkeit oder Anzahl der Jahre der chronischen Erkrankung (Glasgow et al. 2005a).

### 1.5.6 Kritik am PACIC

In einer weiteren Studie untersuchten Glasgow et al. anhand einer Stichprobe von 363 Typ-2-Diabetikern den PACIC hinsichtlich Korrelationen mit demografischen oder medizinischen Merkmalen. Die Autoren kamen zu dem Schluss, dass die Subskalen des PACIC so stark miteinander korrelieren, dass eine getrennte Interpretation der Subskalen möglicherweise nicht sinnvoll ist (Glasgow et al. 2005b).

#### 1.5.6.1 Likert-Skala mit fünf Skalenstufen

Gugiu et al. führten eine umfassende Evaluation der Originalversion des PACIC mit 20 Items und der Likert-Skala mit fünf Skalenstufen durch (Gugiu et al. 2009; Gugiu et al. 2010).

Dabei wurde die Likert-Skala mit fünf Skalenstufen des PACIC kritisiert. Für eine Auswertung des Gesamtwerts sowie der Subskalen mittels Mittelwerten sei die verwendete ordinale Skala nicht geeignet. Die verwendete Likert-Skala mit fünf Skalenstufen biete keine Äquidistanz zwischen den einzelnen Punkten auf der Skala. Die Interpretation der Autoren des PACIC beruhe auf einer Annahme der Äquidistanz, die wahrscheinlich nicht gegeben sei. Dies sei nur bei einer Intervall- oder Verhältnisskala möglich (Jöreskog 1994; Gugiu et al. 2009; Gugiu et al. 2010).

Für die eigenen Berechnungen entwickelten Gugiu et al. ausgehend von der Likert-Skala mit fünf Skalenstufen eine Antwortskala mit elf Skalenstufen von 0 % bis 100 %. Der Abstand zwischen den elf Skalenstufen entsprach jeweils 10 %. Während in der Originalversion des PACIC „1“ mit „So gut wie nie“ und „5“ mit „Fast immer“ kodiert war, war „Niemals“ bei Gugiu et al. mit 0 % kodiert und „Immer“ mit 100 %. Die Instruktionaufforderung des Instruments wurde ebenfalls geändert. Nun wurden die Items hinsichtlich der Behandlung nicht mehr anhand des Ausmaßes einzelner Maßnahmen bewertet, sondern der prozentuale Anteil der Maßnahmen an der Behandlungsdauer. Der Zeitraum zur Bewertung der erhaltenen Behandlung blieb mit sechs Monaten unverändert (Gugiu et al. 2009; Gugiu et al. 2010).

#### 1.5.6.2 Struktur mit fünf Faktoren und Faktorenanalyse

Gugiu et al. führten eine explorative und konfirmatorische Faktorenanalyse durch, um die von den Autoren vorgeschlagene Fünf-Faktoren-Struktur des PACIC zu revalidieren. Die Fünf-Faktoren-Struktur entspricht den fünf Subskalen des CCM. Die Ergebnisse der Analyse bestätigten eindeutig *nicht* das von Glasgow et al. vorgeschlagene Fünf-Faktoren-Modell. Vielmehr legten die Ergebnisse nahe, dass der PACIC ein eindimensionales Konstrukt misst. Somit konnte die Anmerkung von Glasgow et al., dass

die Subskalen des PACIC für eine getrennte Interpretation zu stark miteinander korrelieren, bestätigt werden (Gugiu et al. 2010).

### 1.5.7 Weiterentwicklung: Der PACIC mit 11 Items

Wie bereits beschrieben, entwickelten Gugiu et al. den PACIC weiter. Auf Grundlage der Ergebnisse einer explorativen und konfirmatorischen Faktorenanalyse wurden neun der 20 Items aus dem modifizierten PACIC gestrichen. Bevor ein Item eliminiert wurde, musste nachgewiesen werden, dass eine Eliminierung die Modellanpassung verbesserte. Trotz der Reduzierung der Anzahl der Items um 45 % ergab sich eine Verbesserung der Modellanpassung an das hypothetische Konstrukt des PACIC. Die Ergebnisse zeigten zudem, dass der PACIC auf 11 Items reduziert werden konnte, ohne dass sich die psychometrischen Eigenschaften verschlechterten. Es wurde kein Zusammenhang zwischen den demografischen Variablen und dem PACIC festgestellt.

Wie bereits ausgeführt, wurden die Skalenstufen von einer fünfstufigen Likert-Skala auf eine elfstufige Prozent-Skala von 0 % bis 100 % in Schritten von je 10 % umgestellt.

Trotz der Eliminierung von neun Items blieben die fünf Subskalen mit mindestens je einem Item erhalten. In der nachfolgenden Auflistung (*Tabelle 1*) entspricht die Nummerierung dem jeweiligen Item der PACIC-Version mit 20 Items. Diese elf Items blieben bestehen, sind in der *Tabelle 1* mit einem Haken markiert () und repräsentieren folgende Subskalen:

- Item 1: Patientenaktivierung (Patient Activation/Involvement)
- Item 5: Gestaltung des Versorgungssystems/Entscheidungsunterstützung (Delivery System Design/Decision Support)
- Items 8-11: Zielsetzung (Goal Setting/Tailoring)
- Items 13-15: Problemlösung/kontextbezogene Beratung (Problem solving/Contextual Counseling)
- Item 16 und Item 19: Nachbereitung/Koordination (Follow-up/Coordination)

Originalversion (Glasgow et al. 2005a)		Kurzversion (Gugiu et al. 2009)
1. Asked for my ideas when we made a treatment plan.	Patientenaktivierung	
2. Given choices about treatment to think about.		<input checked="" type="checkbox"/>
3. Asked to talk about any problems with my medicines or their effects.		
4. Given a written list of things I should do to improve my health.	Gestaltung des Versorgungssystems	
5. Satisfied that my care was well organized.		<input checked="" type="checkbox"/>
6. Shown how what I did to take care of myself influenced my condition.		
7. Asked to talk about my goals in caring for my condition.	Zielsetzung	

8. Helped to set specific goals to improve my eating or exercise.		<input checked="" type="checkbox"/>
9. Given a copy of my treatment plan.		<input checked="" type="checkbox"/>
10. Encouraged to go to a specific group or class to help me cope with my chronic condition.		<input checked="" type="checkbox"/>
11. Asked questions, either directly or on a survey, about my health habits.		<input checked="" type="checkbox"/>
12. Sure that my doctor or nurse thought about my values, beliefs, and traditions when they recommended treatments to me.	Problemlösung	
13. Helped to make a treatment plan that I could carry out in my daily life.		<input checked="" type="checkbox"/>
14. Helped to plan ahead so I could take care of my condition even in hard times.		<input checked="" type="checkbox"/>
15. Asked how my chronic condition affects my life.		<input checked="" type="checkbox"/>
16. Contacted after a visit to see how things were going.	Nachbereitung	<input checked="" type="checkbox"/>
17. Encouraged to attend programs in the community that could help me.		
18. Referred to a dietitian, health educator, or counselor.		
19. Told how my visits with other types of doctors, like an eye doctor or other specialist, helped my treatment.		<input checked="" type="checkbox"/>
20. Asked how my visits with other doctors were going.		

**Tabelle 1: Gegenüberstellung der PACIC-Items der Originalversion und der Kurzversion**

### 1.5.8 Validierung der PACIC-Version mit 11 Items

Ziel von Goetz et al. war es, die gekürzte Form des PACIC ins Deutsche zu übersetzen und die psychometrischen Eigenschaften bei Patienten mit einer chronischen Erkrankung in Deutschland zu untersuchen. Insgesamt füllten 264 Patienten der Primärversorgung die PACIC-Kurzform aus. In absteigender Reihenfolge waren die häufigsten chronischen Erkrankungen Bluthochdruck, Rückenschmerzen, Osteoarthritis und Typ-2 Diabetes (Goetz et al. 2012).

Die PACIC-Kurzform besteht aus 11 Items und 11 Skalenstufen, die von 0 % ("Nie") bis 100 % ("Immer") reicht. Der Abstand zwischen den einzelnen Skalenstufen liegt bei jeweils 10 % (Goetz et al. 2012).

Die PACIC-Kurzform zeigte eine gute konvergente Konstruktvalidität zum PACIC mit 20 Items (Spearman-Rangkorrelation 0,82,  $p < 0,001$ ) und eine hohe interne Konsistenz (Cronbach's  $\alpha = 0,87$ ). Die Hauptkomponentenanalyse unterstrich die eindimensionale Struktur des Instruments (Goetz et al. 2012).

### 1.5.9 Vergleich der PACIC-Items und der Skalenbeschriftung

Die fünf Stufen der Likert-Skala des ursprünglichen PACIC und des deutschen PACIC mit jeweils 20 Items im Vergleich (in aufsteigender Reihenfolge) sind in der *Tabelle 2* dargestellt:

Originalversion (Glasgow et al. 2005a)	Deutsche Version (Rosemann et al. 2007)
- „Almost Never“	- „So gut wie nie“
- „Generally Not“	- „Meistens nicht“
- „Sometimes“	- „Gelegentlich“
- „Most of the Time“	- „Meistens“
- „Almost Always“	- „Fast immer“

**Tabelle 2: Gegenüberstellung der Stufen der Likert-Skala der englischen Originalversion und der deutschen Version des PACIC**

Die bei 0 % und 100 % verankerten Skalenstufen des PACIC mit 11 Items im Vergleich sind in der *Tabelle 3* dargestellt:

	Originalversion (Gugiu et al. 2009)	Deutsche Version (Goetz et al. 2012)
0 %	„None“	„Nie“
100 %	„Always“	„Immer“

**Tabelle 3: Gegenüberstellung der Skalenstufen des PACIC mit 11 Items bei 0 % und bei 100 %**

### 1.5.10 PACIC-Versionen in anderen Sprachen

Der PACIC wurde außer in die deutsche Sprache noch in diverse andere Sprachen übersetzt:

- Spanisch (Aragones et al. 2008)
- Niederländisch (Wensing et al. 2008)
- Dänisch (Maindal et al. 2012)
- Französisch (Krucien et al. 2014)
- Bengalisch (Koley et al. 2016)
- Finnisch (Simonsen et al. 2018)
- Thailändisch (Zeugfang et al. 2018)
- Malaiisch (Abdul-Razak et al. 2018)
- Serbisch (Pilipovic-Broceta et al. 2018)
- Griechisch (Malliarou et al. 2020)
- Arabisch (Alharbi et al. 2021)
- Ungarisch (Törő et al. 2022)

## 1.5.11 Andere Versionen des PACIC

### 1.5.11.1 PACIC-5A

Das "5A"-Modell stellt einen evidenzbasierten Ansatz zur Herbeiführung einer Verhaltensänderung dar (Goldstein et al. 2004). Es wurde ursprünglich für die Raucherentwöhnung entwickelt und stellt den empfohlenen Beratungsansatz für Verhaltensänderungen gemäß den Empfehlungen der US Preventive Services Task Force dar. Die Schlüsselemente sind die Bewertung des gegenwärtigen Verhaltens („Assess“), die Beratung des Patienten („Advise“), die gemeinsame Vereinbarung realistischer Ziele mit dem Patienten („Agree“), die Unterstützung des Patienten bei der Änderung seines Lebensstils („Assist“) und häufige Nachuntersuchungen („Arrange“) (Glasgow et al. 2005b). Glasgow et al. erweiterten den PACIC um sechs Items, mit denen sie bewerteten, inwieweit die ärztliche Beratung dem 5A-Ansatz entspricht. Sie validierten den PACIC-5A in einer Stichprobe von Diabetespatienten (Glasgow et al. 2005b). In einer Studie untersuchten Rosemann et al. die Kongruenz zwischen der Betreuung in der Allgemeinmedizin in Deutschland und dem CCM sowie zwischen dem Beratungsstil der Allgemeinmediziner und dem 5A-Ansatz (Rosemann et al. 2008).

### 1.5.11.2 O-PACIC

Cramm und Nieboer entwickelten eine Kurzversion des Fragebogens, den „Older Patient Assessment of Chronic Illness Care (O-PACIC)“. Mithilfe des O-PACIC sollte eine Skala zur Bewertung der Pflege älterer chronischer Patienten nach einem Krankenhausaufenthalt entwickelt und validiert werden. Ausgangspunkt des Entwicklungsprozesses war die Originalversion mit 20 Items. Mittels konfirmatorischer Faktorenanalyse wurden zehn der 20 Items ausgeschlossen, sodass der O-PACIC aus zehn Items besteht (Cramm und Nieboer 2014).

## 1.5.12 Vergleichbarkeit der PACIC-Fragebögen mit 20 und mit 11 Items

Arditi et al. führten eine systematische Überprüfung und Metaanalyse aller Studien durch, an der Patienten mit Typ-2 Diabetes teilnahmen und der PACIC erhoben wurde. Dabei wurde unter anderem die Haupthypothese nicht bestätigt, dass Patienten, die eine integrierte Versorgung erhalten, signifikant höhere Punktzahlen aufweisen (Arditi et al. 2018).

Hierbei waren zwei Variablen mit höheren PACIC-Werten verbunden. Erstens war dies ein Instrumentenmerkmal, nämlich die Verwendung einer Likert-Skala mit fünf Skalenstufen (im Vergleich zu einer 11-Punkte-Skala). Die Wahl der verankerten Skalenstufen („Nie“/„Immer“ im Vergleich zu „So gut wie nie“/„Fast immer“) war ebenfalls ein signifikanter Prädiktor (Arditi et al. 2018).

Mit Bezug auf Smith et al. führen Arditi et al. an, dass es nicht überraschend sei, dass die Verwendung einer Likert-Skala mit fünf Skalenstufen zu höheren PACIC-Werten führe. Dies wurde damit begründet, dass die Anzahl der Skalenstufen und die Bezeichnung der Extremwerte die Beantwortung des Fragebogens wesentlich beeinflussen. Wichtige Elemente bei der Beantwortung eines Fragebogens sind die Zustimmung (zu den Items), die Extremität (Bevorzugung des äußersten Punktes) und die Mäßigung (Bevorzugung des mittleren Punktes). Dies beeinflusst die Art und Weise, wie Personen eine Likert-

Skala beantworten (Smith et al. 2016). Folglich sei bei der Interpretation der kombinierten Ergebnisse von Studien Vorsicht geboten, wenn eine unterschiedliche Anzahl an Skalenstufen und unterschiedliche Anker verwendet werden (Arditi et al. 2018).

### 1.5.13 Umgang mit fehlenden Werten des PACIC

Der Umgang mit fehlenden Werten des PACIC wird dadurch erschwert, dass es Versionen mit unterschiedlich vielen Items gibt. Zudem gibt es keine offiziellen Richtlinien für den Umgang mit fehlenden Werten. Deshalb existieren verschiedene Ansätze, die in *Tabelle 4* dargestellt werden:

Mindestanteil beantwortet	Verwendete PACIC-Version	Quelle
25 %	O-PACIC mit 10 Items	(Cramm und Nieboer 2014)
60 %	20 Items	(Rick et al. 2012)
2/3 (66,6 %)	20 Items	(Thom et al. 2015)
75 %	Französisch Version mit 18 Items	(Krucien et al. 2014)
85 %	20 Items	(Simonsen et al. 2018)
100 %	PACIC-5A	(Schwenke et al. 2019)

**Tabelle 4: Gegenüberstellung verschiedener Grenzwerte fehlender PACIC-Items in durchgeführten Studien**

- **Minimum von 25 % beantwortet:** Cramm und Nieboer schlossen Probanden aus ihrer Analyse aus, wenn mehr als 15 der 20 Items fehlten, somit wurden 75 % fehlender Werte beim PACIC akzeptiert (Cramm und Nieboer 2014).
- **Minimum von 60 % beantwortet:** Rick et al. haben das willkürliche Kriterium gewählt, dass die Befragten mindestens 60 % der Items einer Subskala oder der Gesamtskala ausgefüllt haben müssen, um in ihre Analysen einbezogen zu werden (Rick et al. 2012).
- **Minimum von 2/3 beantwortet:** Thom et al. haben die PACIC-Daten als fehlend behandelt, wenn weniger als zwei Drittel der Fragen beantwortet wurden (Thom et al. 2015, S. 687).
- **Minimum von 75 % beantwortet:** Krucien et al. haben einen konservativen Ansatz gewählt und die Befragten, die nicht mindestens 75 % der Items beantwortet haben, von der Analyse ausgeschlossen (Krucien et al. 2014).
- **Minimum von 85 % beantwortet:** Simonsen et al. akzeptierten bei der PACIC-Version mit 20 Items drei fehlende Items (Simonsen et al. 2018).
- **Minimum von 100 % beantwortet:** Schwenke et al. schlossen nur Patienten ein, die keinen fehlenden Wert aufwiesen (Schwenke et al. 2019).

## 1.5.14 Zusammenhänge zwischen PACIC und Merkmalen

### 1.5.14.1 Zusammenhänge zwischen PACIC und Patientenmerkmalen

Rosemann et al. übersetzten den PACIC-5A in die deutsche Sprache und evaluierten diesen anschließend. In dem PraxArt-Projekt wurden in 75 Hausarztpraxen in Baden-Württemberg insgesamt 1.021 Patienten mit Osteoarthritis behandelt. Eine Stichprobe von 300 Patienten füllten die deutsche Version des PACIC aus. Die Auswertung ergab, dass das weibliche Geschlecht sowie das Alter schwach positiv mit dem PACIC-Gesamtwert korrelierten. Ebenfalls im Rahmen des PraxArt-Projekts wurde anschließend anhand von 1.021 Probanden überprüft, inwieweit der PACIC mit dem CCM und dem 5A-Ansatz übereinstimmt. Hierbei ergaben sich höhere PACIC-Werte bei jüngeren Probanden und Probanden mit einer geringeren depressiven Symptomatik (Rosemann et al. 2007; Rosemann et al. 2008).

In einer dänischen Studie mit Diabetes-Patienten waren der PACIC-Gesamtwert sowie die Subskalen des PACIC bis zu einem Alter zwischen 50 und 60 Jahren positiv korreliert; mit zunehmendem Alter ab einem Alter von etwa 60 Jahren nahm der Gesamtwert des PACIC wieder ab (Frølich et al. 2021).

Rick et al. untersuchten anhand von Patienten mit chronischen Erkrankungen in Großbritannien die psychometrischen Eigenschaften des PACIC. Hierbei wiesen Frauen niedrigere PACIC-Werte auf (Rick et al. 2012). In einer weiteren dänischen Studie mit Diabetes-Patienten fanden Bo et al. bei Frauen ebenfalls geringere PACIC-Werte (Bo et al. 2020).

Wiederum fanden Ali et al. bei Patienten mit Herzerkrankungen signifikant höhere PACIC-Werte bei Frauen. Zusätzlich waren höhere PACIC-Werte mit einer geringeren depressiven Symptomatik verbunden. Es fand sich jedoch kein Zusammenhang zwischen PACIC und dem Alter (Ali et al. 2017).

In einer multizentrischen Studie wurden Patienten mit koronarer Herzkrankheit (KHK) (u.a. in Deutschland) hinsichtlich von Faktoren untersucht, die mit der Inanspruchnahme chronischer Behandlung stehen. Das männliche Geschlecht war mit höheren PACIC-Werten verbunden (Ludt et al. 2012).

In einer Studie von Peltzer et al. mit KHK-Patienten waren psychische Erkrankungen mit einem geringfügig niedrigeren PACIC-Wert korreliert (Peltzer et al. 2020).

In einer niederländischen Studie anhand von Patienten mit Herz-Kreislauf-Erkrankungen sowie chronisch obstruktiver Lungenerkrankung (COPD) zeigte sich, dass in beiden Patientengruppen jüngere und weniger depressive Patienten höhere PACIC-Werte erzielten (Cramm und Nieboer 2012).

Jedoch zeigten sich in mehreren Studien keine Zusammenhänge zwischen Patientenmerkmalen und PACIC-Werten (Glasgow et al. 2005b; Aragonés et al. 2008; Taggart et al. 2010).

### 1.5.14.2 Zusammenhänge zwischen PACIC und Praxismerkmalen

In einer Querschnittsstudie chronischer Patienten in deutschen Hausarztpraxen waren auf Praxisebene Einzelpraxen mit signifikant höheren PACIC-Werten verbunden (Petersen et al. 2014).

In einer Studie von Taggart et al. zeigten sich keine Zusammenhänge zwischen Praxismerkmalen und PACIC-Werten (Taggart et al. 2010).

### 1.5.14.3 Psychometrische Eigenschaften des PACIC

Mit der PRoMPT-Studie wollten Gensichen et al. die psychometrischen Eigenschaften des deutschen PACIC in einer Stichprobe von Patienten mit Depression evaluieren. Es konnte gezeigt werden, dass die meisten psychometrischen Eigenschaften der deutschen Version des PACIC gut oder akzeptabel waren, wenn sie bei Patienten mit schweren Depressionen in der Primärversorgung angewendet wurden (Gensichen et al. 2011).

## 1.6 Primary Care Evaluation of Mental Disorders (PRIME-MD) und PRIME-MD Patient Health Questionnaire (PHQ)

Der für diese Arbeit verwendete PHQ-9 ist das Depressionsmodul des Patient Health Questionnaire (PHQ), dem Nachfolger des Primary Care Evaluation of Mental Disorders (PRIME-MD) (Kroenke et al. 2001). Zum besseren Verständnis ist zunächst die Entwicklung des PRIME-MD sowie des PHQ dargestellt.

### 1.6.1 Primary Care Evaluation of Mental Disorders (PRIME-MD)

Der PRIME-MD wurde 1994 von Spitzer et al. mit der „PRIME-MD 1000 Study“ etabliert. Es handelte sich um den ersten psychiatrischen Fragebogen, der für die Primärversorgung entwickelt wurde. Der PRIME-MD ist ein Screeninginstrument für die häufigsten psychischen Störungen sowohl in der Gesamtbevölkerung als auch im Setting der Primärversorgung (Spitzer et al. 1994).

Die psychischen Störungen sind aufgeteilt in fünf Gruppen:

- Störungen der Stimmung (Affektive Störungen)
- Angststörungen
- Somatoforme Störungen
- Störungen in Bezug auf den Alkoholkonsum
- Essstörungen (Spitzer et al. 1994)

Der PRIME-MD besteht aus zwei Bestandteilen:

- Fragebogen (PQ, englisch: patient questionnaire) für den Patienten mit einem Umfang von einer Seite, welcher vor dem Arztbesuch ausgefüllt wird
- Klinische Checkliste, (CEG, englisch: clinical evaluation guide) ein strukturierter Beurteilungsleitfaden mit einem Umfang von 12 Seiten für den Arzt, um auffällige Patientenantworten zu überprüfen (Spitzer et al. 1994)

Der einseitige Patientenfragebogen stellt ein Screeninginstrument für Symptome psychischer Störungen dar. Der Fragebogen besteht aus 26 Fragen im „Ja/Nein“-Format. Beim Ausfüllen der Fragen werden Symptome und Anzeichen psychischer Störungen innerhalb des letzten Monats abgefragt. Eine Abstufung der Intensität oder Häufigkeit findet nicht statt (Spitzer et al. 1994).

Die 26 Items wurden in Anlehnung an die „Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition (DSM-III-R)“ entwickelt. Gemäß der Aufteilung der abgefragten psychischen



Störungen in fünf Gruppen sind die 26 Fragen in fünf Skalen gegliedert. Die Antworten des Patienten auf dem Fragebogen geben dem Arzt Hinweise darauf, welche der fünf Checklisten des CEG, wenn überhaupt, für die Überprüfung der Patientenantworten verwendet werden sollen. Für jede der Skalen existieren Grenzwerte, ab denen eine Skala als „auffällig“ behandelt wird. Jede „auffällige“ Skala wird anschließend für sich genommen mit der passenden Checkliste des CEG im Rahmen eines Arzt-Patienten-Gesprächs überprüft (Spitzer et al. 1994).

Mithilfe des CEG kann der Arzt das Vorhandensein oder Fehlen von 18 psychischen Störungen ermitteln. Neun der 18 psychischen Störungen („threshold disorders“) genügen den Diagnosekriterien der DSM-III-R. Weitere sechs Störungen („subthreshold disorders“) können als PRIME-MD-Störungen erfasst werden; die Diagnosekriterien der DSM-III-R werden hierbei jedoch nicht komplett erfüllt. Die „subthreshold disorders“ werden trotzdem gescreent, da auch eine unterschwellige Symptomatik mit erheblichen Funktionsstörungen verbunden sein kann. Patienten könnten von einer Observation, einer vertieften Anamnese und Therapie profitieren. Des Weiteren wurden drei Ausschlussdiagnosen hinzugefügt, deren Vorhandensein die Diagnostik der psychischen Störungen ausschließen würde (Spitzer et al. 1994).

Spitzer et al. gaben eine gute Übereinstimmung von Diagnosen mittels PRIME-MD und klinischer Evaluation durch klinische Spezialisten an. Für psychiatrische Diagnosen (gemessen an den Diagnosen durch klinische Spezialisten) wurde die Gesamtgenauigkeit mit 88 % angegeben. An der „The PRIME 1000 Study“ hatten 1.000 Patienten teilgenommen. 26 % der Patienten erhielten eine PRIME-MD-Diagnose, die die vollständigen Kriterien für eine bestimmte Störung der DSM-III-R („threshold disorder“) erfüllte. Bei fast der Hälfte der Patienten mit einer PRIME-MD-Diagnose war vorher keine psychische Störung bekannt. Wurde der Patientenfragebogen aufgrund positiver Antworten als „auffällig“ erkannt, dauerte die Überprüfung mithilfe der Checklisten durchschnittlich 8,4 Minuten. Besonders die zeitlichen Einschränkungen in der Primärversorgung wurden von den Autoren als mögliches Hindernis für den Einsatz des PRIME-MD identifiziert (Spitzer et al. 1994).

### **1.6.2 Entwicklung des PRIME-MD Patient Health Questionnaire (PRIME-MD PHQ)**

Wie bereits von Spitzer et al. befürchtet (Spitzer et al. 1994), wurde der PRIME-MD aufgrund der zu langen Bearbeitungszeit zwar in der klinischen Forschung, jedoch kaum in der Primärversorgung eingesetzt. Um die Bearbeitungszeit zu verkürzen und damit den Einsatz des Screeningfragebogens zu erhöhen, wurde mit dem „PRIME-MD Patient Health Questionnaire (PHQ)“ eine Weiterentwicklung des PRIME-MD entwickelt (Spitzer et al. 1999).

Die beiden Komponenten des ursprünglichen PRIME-MD, der Patientenfragebogen und der Beurteilungsleitfaden für den Arzt, wurden in einem dreiseitigen Fragebogen zusammengefasst. Wesentlicher Unterschied zum PRIME-MD ist, dass der Fragebogen komplett selbstständig vom Patienten ausgefüllt wird. Anschließend scannt der Arzt den ausgefüllten Fragebogen, überprüft die positiven Antworten und wendet die diagnostischen Algorithmen an, die unten auf jeder Seite in Kurzform angegeben sind (Spitzer et al. 1999).

Mit dem ursprünglichen PRIME-MD wurden 18 psychische Störungen erfasst. Durch die Zusammenfassung mehrerer Stimmungs-, Angst- und somatoformer Kategorien psychischer Erkrankungen wurde die Differentialdiagnostik mittels des PRIME-MD PHQ erheblich vereinfacht, da nur acht Störungen erfasst werden. Wie beim ursprünglichen PRIME-MD sind diese Störungen unterteilt in Störungen, die den diagnostischen Kriterien genügen („threshold disorders“, z.B. Major Depression, Panikstörung, andere Angststörung und Bulimia nervosa) oder die Diagnosekriterien nicht vollständig erfüllen („subthreshold disorders“, z.B. andere depressive Störung, wahrscheinlicher Alkoholmissbrauch oder -abhängigkeit sowie somatoforme Störungen und Essstörungen) (Spitzer et al. 1999).

Eine wichtige Änderung wurde bei den Antwortkategorien für depressive und somatoforme Symptome vorgenommen, die beim ursprünglichen PRIME-MD dichotom (Ja/Nein) waren. Im PRIME-MD PHQ sind die Antwortkategorien erweitert. Die Patienten geben für jedes der neun Symptome einer Depression an, ob das Symptom in den letzten zwei Wochen „Überhaupt nicht“, „An einzelnen Tagen“, „An mehr als der Hälfte der Tage“ oder „Beinahe jeden Tag“ vorlag. Durch diese Änderung wird der PRIME-MD PHQ zu einem erweiterten Diagnoseinstrument, das auch den Schweregrad der Depression erfasst. Dies kann sowohl zu Behandlungsbeginn als auch zur Therapieüberwachung hilfreich sein (Spitzer et al. 1999).

Wie bei der ursprünglichen PRIME-MD-Studie wird erwartet, dass der Arzt vor der endgültigen Diagnose körperliche Ursachen für die Depression und Angstzustände sowie im Falle einer Depression einen normalen Trauerfall und eine manische Episode in der Vergangenheit ausschließt (Spitzer et al. 1999).

Die Datenerhebung für die Validierungsstudie fand zwischen Mai 1997 und November 1998 statt. Es sollte ermittelt werden, ob der selbstständig auszufüllende PRIME-MD PHQ eine vergleichbare Validität und Nützlichkeit für die Diagnose psychischer Störungen in der Primärversorgung besitzt wie der PRIME-MD. Insgesamt wurden die Fragebögen von 3.000 Probanden ausgewertet. 585 der 3.000 Probanden wurden zusätzlich von einem Psychologen interviewt, um die Übereinstimmung zwischen PRIME-MD PHQ und der Diagnostik des Psychologen zu überprüfen. Zusammengenommen hatten 825 (28 %) der 3.000 Probanden eine PHQ-Diagnose. Von den 585 interviewten Probanden waren es 170 (29 %). Die Gesamtgenauigkeit wurde mit 85 %, die Sensitivität mit 75 % und die Spezifität mit 90 % angegeben (Spitzer et al. 1999).

Der Zeitaufwand des Arztes für die Überprüfung des PRIME-MD PHQ betrug bei 85 % der Probanden weniger als drei Minuten, während dies beim ursprünglichen PRIME-MD nur bei 16 % der Probanden der Fall war (Spitzer et al. 1999).

Die Übereinstimmung zwischen der PHQ-Diagnose einer schweren depressiven Störung und der Diagnose des Psychologen wurde erhöht, indem die Diagnoseschwelle der Symptommhäufigkeit von „Beinahe jeden Tag“ auf „An mehr als der Hälfte der Tage“ abgesenkt wurde. Dies erhöhte die Sensitivität von zunächst 37 % auf 73 % (Spitzer et al. 1999).

Die deutsche Version des PRIME-MD PHQ, der „Gesundheitsfragebogen für Patienten (PHQ-D) (Löwe et al. 2002) wurde von Gräfe et al. validiert. Dabei zeigte sich eine ausgezeichnete Kriteriumsvalidität des PHQ-D insbesondere bei der Diagnose der Major Depression mit einer Sensitivität von 95 % und einer Spezifität von 86 % (Gräfe et al. 2004). Die Autoren gaben ebenso Grenzwerte für

Screeningzwecke an. Da dabei eine hohe Sensitivität notwendig ist, wurden PHQ-9-Grenzwerte von 11 Punkten für die Diagnose einer Major Depression und 9 Punkten für die Diagnose aller depressiven Störungen angegeben (Gräfe et al. 2004).

### 1.6.3 PHQ-9: Messung der depressiven Symptome

#### 1.6.3.1 PHQ-9

Der PHQ-9 ist das Depressionsmodul des Patient Health Questionnaire (PHQ), in dem jedes der neun DSM-IV-Kriterien einer Depression auf einer Skala von 0 („Überhaupt nicht“) bis 3 („Beinahe jeden Tag“) bewertet wird. Als Maß für den Schweregrad kann der PHQ-9-Score somit Werte von 0 bis 27 erreichen. Dadurch ermöglicht der PHQ-9 nicht nur eine auf Kriterien basierende Diagnostik depressiver Störungen, sondern ist auch ein zuverlässiges und valides Maß für den Schweregrad einer Depression (Kroenke et al. 2001).

Eine „schwere Depression“ (Major Depression) wird diagnostiziert, wenn 5 oder mehr der 9 Kriterien für depressive Symptome in den letzten 2 Wochen an mindestens "An mehr als der Hälfte der Tage" vorhanden waren und eines der Symptome eine gedrückte Stimmung oder Anhedonie ist. Eine „andere Depression“ wird diagnostiziert, wenn 2, 3 oder 4 depressive Symptome in den letzten 2 Wochen an mindestens "An mehr als der Hälfte der Tage" vorhanden waren und eines der Symptome depressive Stimmung oder Anhedonie ist. Eines der 9 Symptomkriterien ("Gedanken, dass Sie lieber tot wären oder sich Leid zufügen möchten“) zählt, wenn es wenigstens „An einzelnen Tagen“ vorhanden ist, unabhängig von der Dauer (Kroenke et al. 2001).

Für die Abstufung der depressiven Symptomatik wurden verschiedene Schweregrade festgelegt. PHQ-9-Werte von mindestens 5, 10, 15 oder 20 Punkten standen für leichte, mäßige, mäßig schwere bzw. schwere Depression (Kroenke et al. 2001). Für das Screening auf eine Depression empfehlen die Autoren einen PHQ-9-Wert von 10 oder höher (Kroenke und Spitzer 2002).

Fehlende Werte des PHQ-9-Fragebogens werden durch den Mittelwert der Punktzahl ersetzt, wenn weniger als 20 % der Antworten fehlen (Kroenke et al. 2010).

In der Validierungsstudie des PHQ-9 wurde ein strukturiertes Telefoninterview durch einen klinischen Psychologen als Standardkriterium eingesetzt. Dabei hatte ein PHQ-9-Wert  $\geq 10$  eine Sensitivität von 88 % und eine Spezifität von 88 % für eine Major Depression. PHQ-9-Werte von 5, 10, 15 und 20 standen für leichte, mäßige, mäßig schwere bzw. schwere Depressionen (Kroenke et al. 2001).

Löwe et al. untersuchten, inwiefern mittels des PHQ-9 auch Veränderungen der depressiven Symptomatik, beispielsweise durch eine Therapie, erfasst werden können. Neben dem PHQ-9 wurde das strukturierte klinische Interview SCID (First 1997) bei der Erst- und Nachuntersuchung ausgefüllt bzw. durchgeführt. Die Empfindlichkeit des PHQ-9 gegenüber Veränderungen in drei Gruppen von Patienten, deren Depressionsstatus sich im Laufe der Zeit entweder verbesserte, unverändert blieb oder sich verschlechterte, wurde untersucht. Die Änderungswerte des PHQ-9 unterschieden signifikant zwischen den drei Gruppen. Daraus wurde geschlossen, dass der PHQ-9 sowohl für Längs- als auch für Querschnittstudien verwendet werden kann (Löwe et al. 2004a).

In einer weiteren Arbeit von Löwe et al. wurde der minimal klinisch wichtige Unterschied zur Beurteilung der individuellen depressiven Symptomatik mittels des PHQ-9-Werts überprüft. Die Veränderung des PHQ-9-Werts, der auf eine Änderung der depressiven Symptomatik schließen lässt, wurde mit 5 Punkten angegeben. Es wurde zusammengefasst, dass der PHQ-9 aufgrund der Änderungssensitivität ein attraktives Instrument für die Messung des Therapieerfolgs ist (Löwe et al. 2004b).

In einer Übersichtsstudie von Gilbody et al. wurden 14 Validierungsstudien des PHQ-9 in Bezug auf die Diagnose „Major Depressive Disorder“ überprüft. Hierbei betrug die Sensitivität 80 % sowie die Spezifität 92 % (Gilbody et al. 2007).

## 2. Material und Methoden

### 2.1 Studien

#### 2.1.1 Jena-PARADIES (Patient Activation foR Anxiety DisordErS)

##### 2.1.1.1 Kurzbeschreibung

Die PARADIES-Studie war eine Cluster-randomisierte Studie mit zwei Studienarmen. Die Cluster bildeten die Hausarztpraxen, die entweder der Interventionsgruppe oder der Kontrollgruppe zugeordnet waren. Die Probanden waren Patienten der jeweiligen Hausarztpraxen (Gensichen et al. 2014).

Die Datenerhebung erfolgte zwischen August 2012 und Oktober 2014 (Teismann et al. 2018).

Bei der Intervention handelte es sich um ein Praxisteam-unterstütztes Behandlungsprogramm für Patienten mit Panikstörung mit/ohne Agoraphobie. Das Behandlungsprogramm vereinte Elemente der kognitiven Verhaltenstherapie (KVT; z.B. Expositionstraining) sowie des Chronic Care Model zu einem Hausarzt-basiertem Case Management System. Ziel des Case Managements war die Aktivierung des Patienten zur eigenständigen Durchführung therapeutischer Expositionsübungen unter Supervision des Hausarztes (Hiller et al. 2018).

Insgesamt nahmen 73 Hausarztpraxen und 419 Patienten mit Panikstörung mit/ohne Agoraphobie (ICD-10: F40.01 oder F41.0) teil. Zur Interventionsgruppe zählten 36 Praxen mit 230 Patienten, zur Kontrollgruppe zählten 37 Praxen mit 189 Patienten (Hiller et al. 2018).

##### 2.1.1.2 Zeitlicher Ablauf, Intervention

Über einen Zeitraum von 23 Wochen fanden für jeden Patienten vier Termine mit dem Hausarzt sowie zehn protokollbasierte Telefonkontakte mit einer/einem Medizinischen Fachangestellte/r (MFA) statt. Zusätzlich erhielten die Patienten ein Selbsthilfehandbuch (Gensichen et al. 2014).

Inhalte der Termine mit dem Hausarzt waren „KVT-orientierte[n] Therapieelemente Psychoedukation (Aufklärungsgespräch), interozeptive Exposition („Körperübungen“), in vivo Exposition („Situationsübungen“) und Rückfallprophylaxe“ (Hiller et al. 2018); zu diesen Inhalten erhielten die Patienten zusätzlich ein Selbsthilfehandbuch. Die Telefonkontakte mit der/dem MFA fanden circa alle zwei Wochen statt und dienten der kontinuierlichen Überwachung der Angstsymptome sowie des Behandlungsverlaufs (Gensichen et al. 2014; Hiller et al. 2018).

Die Patienten wurden angehalten, mindestens zweimal pro Woche Expositionsübungen durchzuführen. Die Einhaltung des Behandlungsprotokolls wurde mittels eines Selbstauskunftsfragebogens gemessen (Gensichen et al. 2014).

##### 2.1.1.3 Monitoring, Measures, Konsequenzen

Eigens für das Telefonmonitoring wurde eine Checkliste entwickelt, die „Jena Angst-Monitoring-Liste (JAMoL)“. Mithilfe dieses Tools wurden die Angstsymptome und die Compliance der Patienten mittels

einer Ampel-Kodierung erfasst und bewertet. Verstärkten sich die Symptome oder war die Compliance nicht ausreichend, nahm der Hausarzt Kontakt zum Patienten auf, um die Behandlungssituation zu verbessern (Gensichen et al. 2014; Hiller et al. 2018).

Analog zum Chronic Care Model wurde die JAMoL als „klinisches Informationssystem“ genutzt (Hiller et al. 2018).

Depressivität wurde mithilfe des Depressionsmoduls PHQ-9 des PHQ-D (Löwe et al. 2002) erfasst, die Sicht der Patienten auf die Versorgung mithilfe der Kurzform des PACIC mit einer Prozent-Skala und elf Skalenstufen (Goetz et al. 2012; Gensichen et al. 2014).

Die Daten wurden erhoben zur Baseline (T0), sechs ( $\pm 1$ ) Monate nach der Baseline (T1) sowie zwölf ( $\pm 1$ ) Monate nach der Baseline (T2) erhoben (Gensichen et al. 2014).

#### 2.1.1.4 Kontrollgruppe

Die Patienten der Kontrollgruppe erhielten den leitliniengerechten Behandlungsstandard (Gensichen et al. 2014).

#### 2.1.1.5 Ethik

Das Studienprotokoll wurde von der Ethikkommission der Medizinischen Fakultät der Friedrich-Schiller-Universität Jena zugelassen (Genehmigungsnummer: 3484-06/1). Die Studie wurde bei ISRCTN registriert (ISRCTN64669297).

### 2.1.2 PRoMPT (PRimary care Monitoring for depressive Patients Trial)

#### 2.1.2.1 Kurzbeschreibung

Bei der PRoMPT-Studie handelte es sich um eine Cluster-randomisierte Studie mit zwei Studienarmen. Die Hausarztpraxen bildeten die Cluster, die entweder zur Interventionsgruppe oder zur Kontrollgruppe gehörten (Gensichen et al. 2005).

Die Studie begann im April 2005 und endete im September 2007 (Gensichen et al. 2009).

Ziel der Studie war die Untersuchung der Wirksamkeit einer Intervention für Patienten mit Depressionen. Die Intervention war ein durch die Hausarztpraxis durchgeführtes Case Management, das in Übereinstimmung mit dem Chronic Care Model konzipiert wurde (Gensichen et al. 2005; Gensichen et al. 2009).

An der Studie nahmen 74 Hausarztpraxen und 626 Patienten mit einer schweren Depression teil. Zur Interventionsgruppe gehörten 35 Praxen mit 310 Patienten, zur Kontrollgruppe 39 Praxen mit 316 Patienten (Gensichen et al. 2009).

#### 2.1.2.2 Zeitlicher Ablauf, Intervention

Die Interventionsgruppe erhielt eine leitliniengerechte Behandlung sowie eine telefonische Abfrage durch eine/n MFA: im ersten Monat im Abstand von zwei Wochen, für die folgenden elf Monate im Abstand von vier Wochen. Die Dauer der Intervention betrug für jeden Patienten zwölf Monate (Gensichen et al. 2005; Gensichen et al. 2009).

### 2.1.2.3 Monitoring, Measures, Konsequenzen

Um das Telefonmonitoring strukturiert durchführen und auswerten zu können, wurde ein Testinstrument entwickelt, die „Depressions-Monitoring-Liste (DeMoL)“. Dieses Instrument besteht zusätzlich zu dem PHQ-9 (Kroenke et al. 2001) aus speziell entwickelten Items zur Untersuchung der Depressionssymptomatik. Diese Items fragen die Medikamenteneinnahme (Frage 10), die Einhaltung der Behandlungsziele ab (z.B. Ausführung von geplanten Tätigkeiten um „das Rückzugsverhalten zu reduzieren“, Item 11) (Gensichen et al. 2006) sowie in einer offenen Frage Wünsche oder Probleme des Patienten (Frage 12) (Gensichen et al. 2005; Gensichen et al. 2006).

Die Auswertung der Telefonabfrage erfolgte nach dem Ampelschema, und nach jedem Patientenkontakt wurde dem Hausarzt ein kurzer Bericht über den Behandlungsverlauf vorgelegt. Je nach Ergebnis nahm der Arzt Kontakt zum Patienten auf (bei Verdacht auf Suizidalität nach PHQ-9 Item 9), wurde die Medikation verändert, ein Arzttermin ausgemacht oder die Behandlung unverändert fortgeführt (Gensichen et al. 2005; Gensichen et al. 2006).

Die Hauptvariable der Studie war die Depressivität, welche anhand des PHQ-9-Fragebogens erfasst wurde (Kroenke et al. 2001; Gensichen et al. 2005; Gensichen et al. 2009).

Um die Sicht der Patienten auf das Case Management abzufragen, wurde der PACIC verwendet. Verwendet wurde die PACIC-Version mit 20 Items sowie der fünfstufigen Likert-Skala mit den Skalenstufen 1 = „So gut wie nie“ bis 5 = „Fast immer“ (Gensichen et al. 2009).

Die Daten wurden zur Baseline (T0), sechs Monate nach der Baseline (T1) sowie zwölf Monate nach der Baseline (T2) erhoben (Gensichen et al. 2005; Gensichen et al. 2009).

### 2.1.2.4 Kontrollgruppe

Die Patienten der Kontrollgruppe erhielten den leitliniengerechten Behandlungsstandard (Gensichen et al. 2005).

### 2.1.2.5 Ethik

Das Studienprotokoll wurde von der Ethikkommission der Medizinischen Fakultät der Goethe-Universität Frankfurt am Main zugelassen (Genehmigungsnummer: E 26/05). Die Studie wurde bei ISRCTN registriert (ISRCTN66386086).

## 2.2 Daten und Variablen

### 2.2.1 Patienten- und Praxismerkmale

Folgende Variablen wurden aus den originalen Studien-Dateien zu einer Gesamtdatei zusammengefügt:

- Patienten ID: Identifizierungsnummer des jeweiligen Probanden (in der jeweiligen Studie)
- Praxis ID: Identifizierungsnummer der jeweiligen Praxis (in der jeweiligen Studie)
- Studienarm: Kontroll- oder Interventionsgruppe

- 0 = Kontrolle PARADIES
- 1 = Intervention PARADIES
- 2 = Kontrolle PRoMPT
- 3 = Intervention PRoMPT
- Alter der Patienten: in Jahren
- Geschlecht der Patienten: männlich oder weiblich
- Datumsangaben zu T0, T1 und T2
- PHQ-9: Punktwerte der einzelnen Items (jeweils zu T0, T1 und T2)
- PACIC
  - PARADIES: Punktwerte der 11 Items (jeweils zu T0, T1 und T2)
  - PRoMPT: Punktwerte der 20 Items (T2)
- Art der Praxis: Einzel- oder Gemeinschaftspraxis
- Anzahl der Ärzte in der Praxis
- Lage der Praxis: Stadt oder Land
- Alter des Studienarztes: in Jahren
- Jahr der Praxis-Niederlassung: als Jahreszahl
- Studienarzt mit Zusatzqualifikation „Psychosomatische Grundversorgung“: „Ja“ oder „Nein“

## 2.3 Einschlusskriterien

Die Einschlusskriterien der Untersuchung sind in Form einer Flowchart in *Abbildung 4* dargestellt.

In einer vorläufigen Analyse wurden die Interventions- und Kontrollgruppen sowohl der PARADIES- als auch der PRoMPT-Studie analysiert, um zu zeigen, dass die Studien hinsichtlich des PACIC-Wertes miteinander verglichen werden können.

Die gesamte Studienpopulation stellte  $n = 1.045$  Probanden dar. Die PARADIES-Studie wurde mit  $n = 419$  Probanden und die PRoMPT-Studie mit  $n = 626$  Probanden durchgeführt.

Da der Fokus der PARADIES-Studie auf Patienten mit Panikstörung mit/ohne Agoraphobie lag (Hiller et al. 2018), wurden in der PARADIES-Studie auch Patienten eingeschlossen, die gemäß des PHQ-9 keine depressive Symptomatik aufwiesen. Diejenigen Probanden wurden von der statistischen Auswertung ausgeschlossen, die zum Messzeitpunkt T0 – gemäß des PHQ-9 – keine depressive Symptomatik aufwiesen. Dies betraf  $n = 155$  Probanden der PARADIES-Studie.

Obwohl der Schwerpunkt der PRoMPT-Studie auf Patienten mit Depressionen lag und ein PHQ-9-Wert von mindestens 9 Punkten zum Messpunkt T0 ein Einschlusskriterium war (Gensichen et al. 2005), erfüllten  $n = 3$  Probanden das Einschlusskriterium (PHQ-9-Wert zum Zeitpunkt T0  $\geq 9$  Punkten) nicht. Bei diesen Probanden fehlte in den statistischen Daten die Angabe des PHQ-9-Wertes, sodass diese Probanden von der statistischen Auswertung ausgeschlossen wurden.

Der PHQ-9 kann ausgewertet werden, wenn maximal 20 % der 9 Items (1,8 Items) fehlen. Wir haben festgelegt, dass maximal 2 der 9 Items fehlen dürfen. Es sollte nicht nur die depressive Symptomatik zu Studienbeginn T0, sondern auch die Entwicklung der depressiven Symptomatik zwischen T1 und T2



analysiert werden. Wenn mehr als 2 Items pro Messpunkt fehlten wurden diese Probanden ausgeschlossen; dies betraf  $n = 86$  Probanden der PARADIES-Studie sowie  $n = 146$  Probanden der PRoMPT-Studie.

Wie unter „Umgang mit fehlenden Werten des PACIC“ beschrieben, existieren keine bindenden Vorgaben zum Umgang mit fehlenden Werten bei der Auswertung des PACIC-Fragebogens. Wir entschieden daher, dass maximal 6 der 11 Items fehlen dürfen. Vom Ausschluss von der Studienpopulation waren somit  $n = 11$  Probanden der PARADIES-Studie sowie  $n = 65$  Probanden der PRoMPT-Studie betroffen.

Zur vorläufigen Studienpopulation wurden insgesamt  $n = 579$  Probanden zugelassen (PARADIES-Studie  $n = 167$ , PRoMPT-Studie  $n = 412$ ). Um in der anschließenden statistischen Auswertung gezielt den Einfluss der jeweiligen Intervention messen zu können, wurden die Kontrollgruppen beider Studien mit ihren Probanden ausgeschlossen.

Die endgültige Studienpopulation bestand aus  $n = 280$  Probanden, davon  $n = 86$  Probanden aus der PARADIES- und  $n = 194$  Probanden aus der PRoMPT-Studie.

## 2.3.1 Flowchart

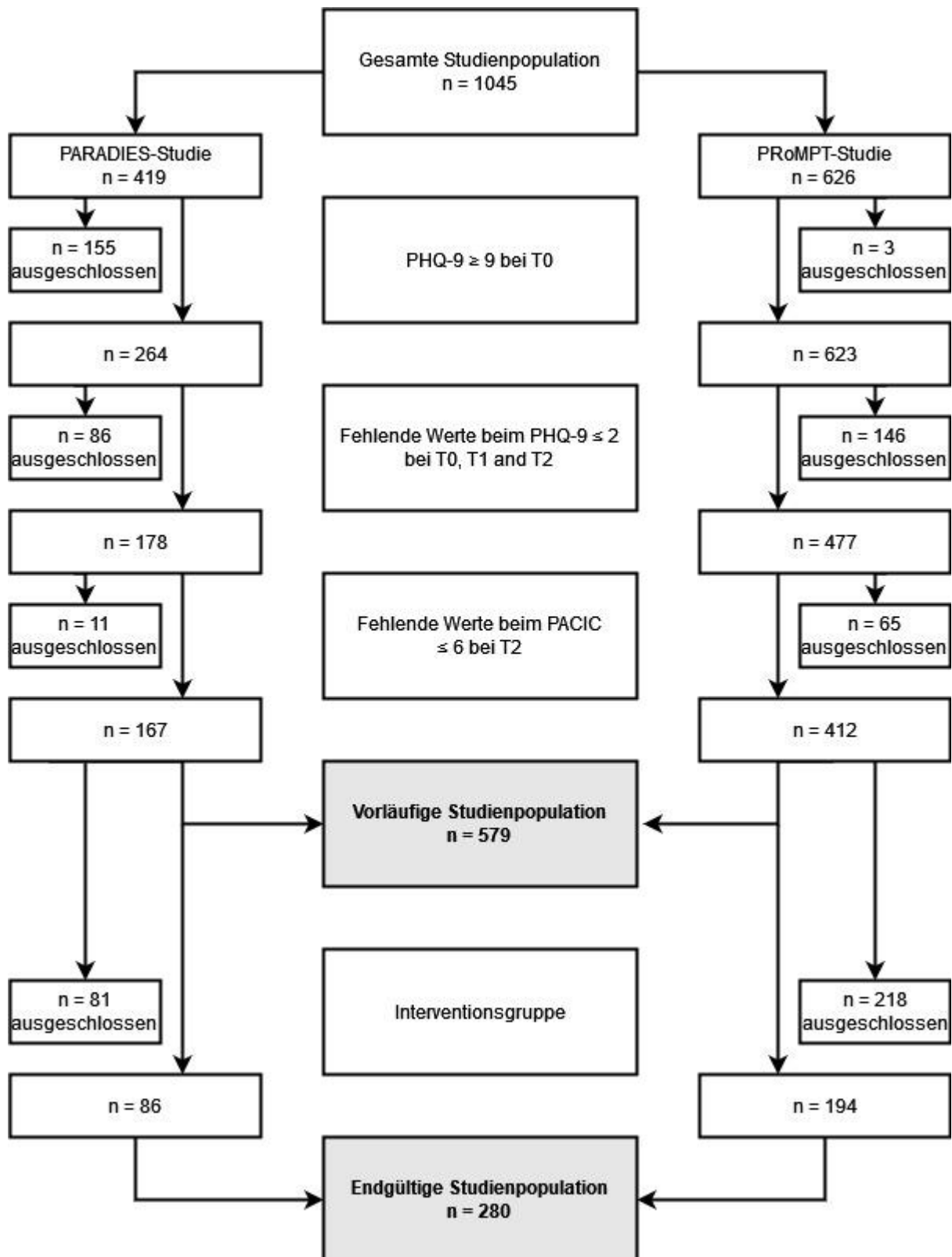


Abbildung 4: Flowchart für die Einschlusskriterien und Studienpopulation

## 2.4 Statistik

Retrospektiv wurden anhand eines gepoolten Datensatzes aus zwei randomisierten kontrollierten Studien Zusammenhänge zwischen Patienten- und Praxismerkmalen und dem PACIC-Wert der Patienten berechnet.

In einer vorläufigen Analyse wurden die Interventions- und Kontrollgruppen sowohl der PARADIES- als auch der PRoMPT-Studie analysiert. Damit sollte gezeigt werden, dass die Studien hinsichtlich des PACIC-Wertes miteinander verglichen werden können.

### 2.4.1 Umgang mit uneinheitlichen PACIC-Versionen

Folgende Unterschiede erschweren den Vergleich zwischen verschiedenen PACIC-Versionen:

- Anzahl von Items (20-Item-Version vs. 11-Item-Version)
- Skalen (5-Punkte-Skala vs. 11-Punkte-Skala)
- Benennung der Skalenstufen („So gut wie nie“ und „Fast immer“ bzw. „Nie“ und „Immer“)

Nachstehend wird der gedankliche Prozess unseres statistischen Vorgehens geschildert.

Die PRoMPT-Studie nutzte den PACIC in der Langform mit 20 Items und fünf Skalenstufen. Die äußeren Skalenstufen waren mit „So gut wie nie“ und „Fast immer“ definiert.

Die PARADIES-Studie nutzte den PACIC in der Kurzform mit elf Items und elf Skalenstufen. Die äußeren Skalenstufen waren mit „Nie“ und „Immer“ definiert.

Problem 1: Unterschiedlich viele Items

Obwohl man prinzipiell beim PACIC die 20-Item-Version mit der 11-Item-Version vergleichen könnte, verzerrt man durch die Einbeziehung des kompletten 20-Item-PACIC womöglich die Ergebnisse. Da die Kurzversion neun Items weniger umfasst als die Langversion, wurden bei der Langversion diese neun Items für die statistische Auswertung ausgeschlossen. Somit werteten wir beim PACIC immer nur die elf identischen Items aus.

Mithilfe von SPSS wurde aus den einzelnen Studiendaten eine Gesamtdatei zur statistischen Analyse erstellt. Hier musste beachtet werden, dass nur identisch kodierte Variablen beider Studien zu einer gemeinsamen Variable in der Gesamtdatei zusammengeführt werden können. Diese Voraussetzungen umfassen „Name“, „Typ“, „Breite“, „Dezimalstellen“, „Beschriftung“, „Werte“ und „Messniveau“ (Skalenmaße). Bei Unterschieden in diesen Punkten kann SPSS die zwei Variablen nicht in eine Variable zusammenführen.

Problem 2: Unterschiedliche Skalen und Kodierung der Minimal- und Maximalwerte

Zu Beginn gab es Überlegungen, inwiefern man die elfstufige Prozent-Skala der Kurzversion in die Likert-Skala mit fünf Skalenstufen umrechnen könnte. Nachfolgend ist dargestellt, welche Skalenstufen der Prozent-Skala dabei welchen Skalenstufen der Likert-Skala entsprechen könnten:

Die 1. Überlegung, dass nur die Extremwerte der Prozent-Skala („0 %“ und „100 %“) den Extremwerten der Likert-Skala entsprechen, ist in *Tabelle 5* dargestellt:

0%	10% - 20% - 30%	40% - 50% - 60%	70% - 80% - 90%	100%
So gut wie nie	Meistens nicht	Gelegentlich	Meistens	Fast immer

**Tabelle 5: Erste Überlegung zur Umrechnung der Prozent-Skala in die Likert-Skala**

Die 2. Überlegung berücksichtigte, dass die Extremwerte der Likert-Skala keine absolut definierten Extremwerte (zum Beispiel „Nie“ und „Immer“) aufweisen (siehe *Tabelle 6*). Deshalb hätte die Prozent-Skala beispielsweise bei „10 %“ und bei „90 %“ ebenfalls den Endpunkten „So gut wie nie“ und „Fast immer“ entsprechen können.

0% - 10%	20% - 30%	40% - 50% - 60%	70% - 80%	90% - 100%
So gut wie nie	Meistens nicht	Gelegentlich	Meistens	Fast immer

**Tabelle 6: Zweite Überlegung zur Umrechnung der Prozent-Skala in die Likert-Skala**

Wie an den beiden Beispielen illustriert, gestaltet sich jedoch die Umwandlung als schwierig. Das liegt vor allem daran, dass die Minimal- und Maximalwerte der beiden Skalen unterschiedlich kodiert sind. Während die Likert-Skala von „So gut wie nie“ bis „Fast immer“ reicht, reicht die Prozent-Skala von „Nie“ bis „Immer“. Die Umrechnung erwies sich somit als wenig praktikabel und statistisch fragwürdig.

Lösungsansatz: Prozentwert vom Maximum

Es wurde der Punktedurchschnitt der gültigen Antworten gebildet und durch den maximal möglichen Gesamtwert dividiert. Der maximal mögliche Gesamtwert käme zustande, wenn alle Items - gemäß der jeweiligen Skala - die maximale Punktzahl erzielen würden.

Um die unterschiedlichen Skalen miteinander vergleichen zu können, wurden folgende Hilfsvariablen erstellt:

- „PACIC\_Maximaler\_Punktwert\_pro\_Item“
  - o Maximale Punktzahl pro Antwort
    - Originalversion (PRoMPT) = 5 (Likert-Skala von 1-5)
    - Kurzversion (PARADIES) = 100 (Prozent-Skala von 0 % bis 100 %)
- „PACICMeanT2\_11\_identische\_Items“
  - o Punktedurchschnitt der gültigen Antworten
- „PACICT2Prozentual\_vom\_Maximum\_11\_identische\_Items“
  - o Prozentualer Anteil des Gesamtwertes am maximal möglichen Gesamtwert

$$\left( \frac{\text{Summe der Antworten}}{\text{Anzahl der Antworten}} \right) * 100 \%$$

Beispielhaft ist der Vergleich von Probanden beider Studien in *Tabelle 7* dargestellt:

Studie	PARADIES			PRoMPT		
	P032-Pat02	P032-Pat03	P034-Pat02	904	905	1301
Item #1	10	-	40	5	-	5
Item #2	60	100	40	4	5	5
Item #3	100	90	40	5	4	-
Item #4	0	100	50	1	4	-
Item #5	0	0	-	5	4	-
Item #6	0	100	50	1	5	5
Item #7	10	90	30	1	1	5
Item #8	10	90	0	4	4	5
Item #9	0	90	10	5	4	5
Item #10	0	-	10	1	2	5
Item #11	30	90	40	1	1	-
Fehlende Werte	0	2	1	0	1	4
Punktedurchschnitt gültiger Werte	20	83,33	31	3	3,4	5
Maximalwert pro Item	100	100	100	5	5	5
Prozent-vom-Maximum	20 %	83,33 %	31 %	60 %	68 %	100 %

**Tabelle 7: Lösungsbeispiel anhand von Probanden aus der PARADIES- und der PRoMPT-Studie**

Beispiel anhand von Proband P032-Pat02 aus der PARADIES-Studie:

$$\left( \frac{10 + 60 + 100 + 0 + 0 + 0 + 10 + 10 + 0 + 0 + 30}{11} \right) * 100 \% = 20 \%$$

Beispiel anhand von Proband 905 aus der PRoMPT-Studie:

$$\left( \frac{5 + 4 + 4 + 4 + 4 + 5 + 1 + 4 + 4 + 2 + 1}{10} \right) * 100 \% = 68 \%$$

## 2.4.2 Patienten- und Praxismerkmale

Kontinuierliche Variablen wie Alter des Patienten, Alter des Studienarztes oder Jahr der Praxis-Niederlassung werden als Mediane mit erstem und drittem Quartil dargestellt. Der Median wird hier statt des Mittelwerts für deskriptive Zwecke verwendet, da er robuster gegenüber Extremwerten innerhalb der Daten ist.

Unterschiede in kontinuierlichen Merkmalen werden mithilfe des Mann-Whitney-U-Test ausgewertet.

Kategoriale Variablen werden als Zählungen und Prozentsätze dargestellt. Diese Variablen beschreiben das Geschlecht des Patienten, Art der Arztpraxis (Einzel- oder Gemeinschaftspraxis), Standort der Praxis (städtisch oder ländlich) sowie die Zusatzqualifikation „Psychosomatische Grundversorgung“ des Studienarztes.

Unterschiede in kategorialen Merkmalen werden mithilfe des  $\chi^2$ -Tests (Chi-Quadrat-Test) ausgewertet.

Das Alter einer Praxis wird entsprechend dem studienspezifischen Median des Gründungsjahres in „neu“ oder „alt“ unterteilt.

### 2.4.3 Erstellte Variablen

Folgende Variablen wurden anhand der Variablen unter „2.3.1 Patienten- und Praxismerkmale“ errechnet und für die statistische Auswertung benötigt:

#### **TageT0bisT1**

Zeitraum in Tagen zwischen T0 und T1

#### **TageT0bisT2**

Zeitraum in Tagen zwischen T0 und T2

#### **PHQMissingT0**

Anzahl der fehlenden Items des PHQ-9 zum Zeitpunkt T0

#### **PHQMissingT1**

Anzahl der fehlenden Items des PHQ-9 zum Zeitpunkt T1

#### **PHQMissingT2**

Anzahl der fehlenden Items des PHQ-9 zum Zeitpunkt T2

#### **PHQMissingGes**

Anzahl der insgesamt fehlenden Items des PHQ-9 zu den Zeitpunkten T0, T1 und T2

#### **PHQGesT0**

Gesamtwert des PHQ-9 zum Zeitpunkt T0

#### **PHQGesT1**

Gesamtwert des PHQ-9 zum Zeitpunkt T1

#### **PHQGesT2**

Gesamtwert des PHQ-9 zum Zeitpunkt T2

#### **PACIC\_Version**

Bei den Probanden der PARADIES-Studie wird der Wert analog der verwendeten Version des PACIC mit 11 Items mit „11“ angegeben.

Bei den Probanden der PRoMPT-Studie wird der Wert analog der verwendeten Version des PACIC mit 20 Items mit „20“ angegeben.

#### **PACIC\_Maximaler\_Punktwert\_pro\_Item**

Diese Variable gibt an, wie viele Punkte pro Antwort des PACIC in der jeweiligen Version erreicht werden konnte (PARADIES-Studie: 100 Punkte; PRoMPT-Studie: 5 Punkte).

#### **Suizidalität\_zu\_T2\_nein1\_womöglich2**

Diese Variable bezieht sich auf das Item 9 des PHQ-9. Dieses Item fragt nach „Gedanken, dass Sie lieber tot wären oder sich Leid zufügen möchten“. Dieses wird bereits als positiv gewertet, wenn „An einzelnen Tagen“ angekreuzt wird. Negativ wird das Item nur bei „Überhaupt nicht“ gewertet. (Kroenke et al. 2001)

#### **PACICMissingItemsT2\_11\_identische\_Items**

Diese Variable stellt die Anzahl der fehlenden Items zum Zeitpunkt T2 dar. Dabei bezieht sich die Angabe fehlender Items lediglich auf die in beiden PACIC-Versionen identischen Items.

#### **PACICMeanT2\_11\_identische\_Items**

Diese Variable stellt den Punktedurchschnitt der Antworten auf die 11 identischen Items zum Zeitpunkt T2 dar.

#### **PACICT2Prozentual\_vom\_Maximum\_11\_identische\_Items**

Diese Variable beschreibt das primäre Ergebnis *PACIC: Prozentwert vom Maximum* unserer Studie und beschreibt den prozentualen Anteil, den der erreichte Gesamtwert gemessen an der maximalen Punktzahl darstellt.

#### **Praxis\_ID\_NEU**

Diese Variable wurde erstellt, damit es zwischen identischen Praxis-IDs beider Studien (PARADIES- und PRoMPT-Studie) nicht zu Überschneidungen kommt. Dies war erforderlich, da die Praxis-IDs im selben Stil kodiert wurden und manche Praxis-IDs in beiden Studien vorkamen (Beispiele: „1“, „2“, „4“ usw.) – dies hätte eine Analyse der Cluster unmöglich gemacht. Die Praxis-ID besteht aus der Bezeichnung des Studienarmes und der ursprünglichen ID der Praxis in der jeweiligen Studie.

### **2.4.4 PACIC-Werte und Sensitivitätsanalyse**

Unterschiede in den PACIC-Werten nach kategorialen Merkmalen wurden grafisch durch Box-Plots dargestellt, und quantitativ mittels Mann-Whitney-U-Test ausgewertet.

Korrelationen zwischen PACIC-Werten und kontinuierlichen Merkmalen wurden grafisch durch Streudiagramme dargestellt und quantitativ mittels Pearson's Korrelationskoeffizient mit 95 %-Konfidenzintervallen ausgewertet.

Im Rahmen einer Sensitivitätsanalyse wurden einzelne Items des PACIC-Wertes explorativ mit dem Depressionsstatus verglichen.

### **2.4.5 Berücksichtigung potenzieller Störfaktoren und Bewertung von Effektgrößen**

Zur Berücksichtigung potenzieller Störfaktoren und zur Bewertung von Effektgrößen, die über Korrelationen hinausgehen, wurden lineare gemischte Regressionsmodelle berechnet. Während Korrelationen relevante Informationen über die Beziehung zwischen zwei Variablen liefern, ist ihr Nutzen durch mögliche Störfaktoren begrenzt. Die Regressionsmodelle zur abhängigen Variable PACIC wurden mit zufälligen Achsenabschnitten („random intercepts“) pro Praxis berechnet, um Beta-Koeffizienten und 95 %-Konfidenzintervalle als Maß für die Assoziation von Patienten- und Praxismerkmalen zu erhalten.

Hierbei haben wir uns entschieden, zufällige Effekte pro Praxis zu modellieren, da die Merkmale von Patienten, die eine bestimmte Praxis aufsuchen, ähnlich sein können, was zu einer Clusterstruktur mit der Praxis als Cluster führen würde.

Die Beta-Koeffizienten stellen die Veränderungen der PACIC-Werte der Patienten pro Veränderung der unabhängigen Merkmale um eine Einheit oder Kategorie dar.

### **2.4.6 Software**

Für alle Berechnungen wurden SPSS (Version 26) und R (Version 4.0.3) verwendet. P-Werte  $< 0,05$  gelten als statistische Signifikanz. Es wurde keine zusätzliche Korrektur für Mehrfachtests durchgeführt.

### **2.4.7 Umgang mit fehlenden Werten bei PACIC und PHQ-9**

Bei der Berechnung des PACIC-Wertes „Prozentwert vom Maximum“ wird zunächst der Durchschnitt der gültigen Antworten errechnet und anschließend ins Verhältnis zur maximal möglichen Punktzahl gesetzt.

Beim PHQ-9-Wert wurde ähnlich vorgegangen, da fehlende Werte durch den Durchschnitt der vorhandenen Werte ersetzt wurden. Um ausgehend von dem Durchschnitt auf den PHQ-9-Gesamt schließen zu können, wurde anschließend der Punktedurchschnitt mit 9 (der Anzahl der PHQ-9-Items) multipliziert.



## 3. Ergebnisse

### 3.1 Deskriptive Auswertung

In der nachfolgenden *Tabelle 8* sind die Patienten- und Praxismerkmale der Studienpopulation aufgeführt und als gesamte Studienpopulation sowie die PARADIES- und PRoMPT-Kohorte dargestellt.

#### 3.1.1 Patientenmerkmale

Mit  $n = 194$  Probanden stellte die PRoMPT-Studie 69,3 % der gesamten Studienpopulation. Im Schnitt waren die Probanden der PRoMPT-Studie mit einem Durchschnittsalter von 52,0 Jahren um 5,7 Jahre älter als die Probanden der PARADIES-Studie (Durchschnittsalter = 46,3 Jahre). Der Anteil der Männer war mit 26,7 % in der PARADIES-Studie und 24,7 % in der PRoMPT-Studie ähnlich.

Die Probanden der PRoMPT-Studie wiesen zu den Messpunkten T0, T1 und T2 durchweg eine höhere depressive Symptomatik (gemäß des PHQ-9) auf als die Probanden der PARADIES-Studie. Zum Messpunkt T2 lag bei 59,8 % der Probanden der PRoMPT-Studie eine depressive Symptomatik vor, während dies bei der PARADIES-Kohorte nur bei 37,2 % der Fall war. Ebenso hatten die Probanden der PRoMPT-Studie zum Messzeitpunkt T2 in 34,5 % der Fälle Suizidgedanken, während dies in der PARADIES-Kohorte nur bei 18,6 % zutraf.

Der *PACIC: Prozentwert vom Maximum* betrug zum Messpunkt T2 64,7 % in der PARADIES-Studie und 69,2 % in der PRoMPT-Studie.

#### 3.1.2 Praxismerkmale

Die Auswertung der Praxismerkmale basiert auf den Angaben von  $n = 242$  Probanden, da diesbezüglich die Daten von  $n = 38$  Probanden unvollständig waren.

In Einzelpraxen wurden 62,4 % der Probanden behandelt, in Gemeinschaftspraxen hingegen 37,6 %. Während das Verhältnis in der PARADIES-Studie weitgehend ausgeglichen war (Einzelpraxen: 48,8 %; Gemeinschaftspraxen: 51,2 %), wurden in der PRoMPT-Studie 69,9 % der Probanden in Einzelpraxen behandelt.

Deutliche Unterschiede zeigten sich ferner beim Standort der Praxis. In der PARADIES-Studie wurden 25,6 % in städtischen Arztpraxen behandelt, während dies in der PRoMPT-Studie zu 88,5 % der Fall war. Mit 66,1 % überwog die Behandlung in städtischen Arztpraxen in der gesamten Studienpopulation deutlich der Behandlung in ländlichen Arztpraxen (33,9 %).

Von den Probanden der PARADIES-Studie wurden 67,4 % von einem Studienarzt mit der Zusatzqualifikation "Psychosomatische Grundversorgung" behandelt, während dies in der PRoMPT-Studie nur bei 6,4 % der Probanden der Fall war.

Das Alter des Studienarztes unterschied sich zwischen beiden Studien im Durchschnitt nur um 0,6 Jahre (PARADIES-Studie: 50,9 Jahre; PRoMPT-Studie: 50,3 Jahre).

		Studienpopulation	PARADIES-Studie	PRoMPT-Studie	p-Wert
		N = 280	N = 86 (30.7%)	N = 194 (69.3%)	
<b>Patientenmerkmale</b>					
Alter (in Jahren)	Median	51.0 [41.8, 58.0]	47.0 [36.2, 53.0]	53.0 [43.0, 61.8]	0.002
	Mittelwert	50.2 ± 13.6	46.3 ± 11.8	52.0 ± 14.0	<0.001
Männer		71 (25.4%)	23 (26.7%)	48 (24.7%)	0.837
PHQ-9-Wert bei T0 (in Punkten)	Median	16.0 [13.0, 19.0]	14.0 [11.0, 17.0]	17.0 [14.0, 19.0]	<0.001
	Mittelwert	16.3 ± 3.9	14.5 ± 4.2	17.0 ± 3.6	<0.001
PHQ-9-Wert bei T1 (in Punkten)	Median	11.0 [6.0, 15.0]	9.0 [5.0, 12.0]	12.0 [7.0, 16.0]	<0.001
	Mittelwert	11.0 ± 5.6	9.0 ± 4.9	11.9 ± 5.6	<0.001
PHQ-9-Wert bei T2 (in Punkten)	Median	9.0 [6.0, 14.0]	8.0 [4.0, 12.0]	10.0 [6.2, 14.0]	<0.001
	Mittelwert	9.9 ± 5.8	8.1 ± 5.2	10.7 ± 5.9	<0.001
Veränderung des PHQ-9-Wertes von T2 zu T0 (in Punkten)	Median	-6.6 [-10.0, -3.0]	-6.1 [-9.0, -2.0]	-7.0 [-10.0, -3.0]	0.636
	Mittelwert	-6.3 ± 5.3	-6.3 ± 5.3	-6.3 ± 5.3	0.614
Depression bei T2 (PHQ-9 ≥ 9)		148 (52.9%)	32 (37.2%)	116 (59.8%)	0.001
Mögliche Suizidalität bei T2		83 (29.6%)	16 (18.6%)	67 (34.5%)	0.011
PACIC, % vom Maximum bei T2	Median	69.1 [54.5, 83.6]	71.8 [43.2, 83.6]	69.1 [56.4, 83.6]	0.519
	Mittelwert	67.8 ± 20.0	64.7 ± 25.0	69.2 ± 17.2	0.314

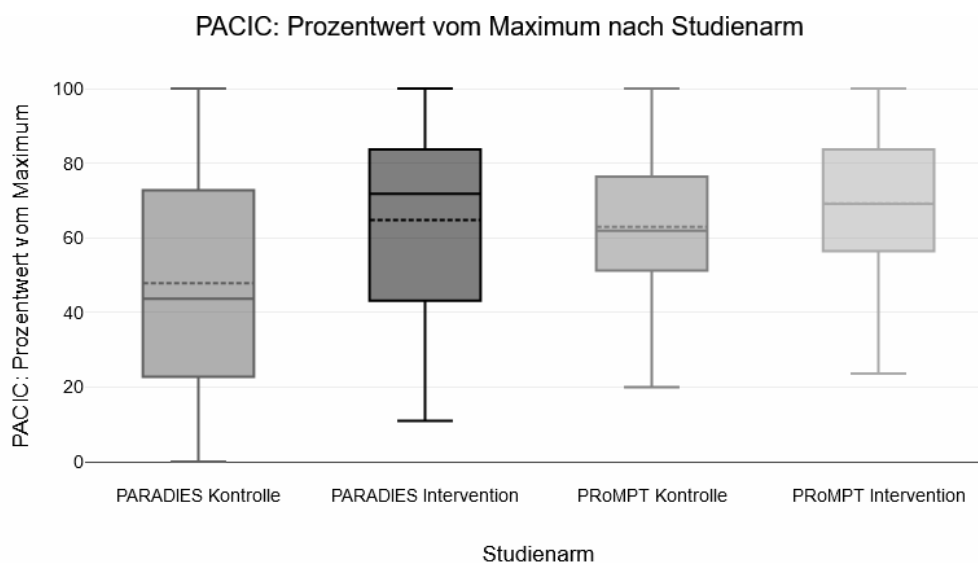
		Studienpopulation	PARADIES-Studie	PRoMPT-Studie	p-Wert
<b>Praxismerkmale</b>					
<i>Basierend auf n = 242 Patienten</i>					
Praxistyp					0.002
Einzelpraxis		151 (62.4%)	42 (48.8%)	109 (69.9%)	
Gemeinschaftspraxis		91 (37.6%)	44 (51.2%)	47 (30.1%)	
Standort der Praxis					<0.001
Städtisch		160 (66.1%)	22 (25.6%)	138 (88.5%)	
Ländlich		82 (33.9%)	64 (74.4%)	18 (11.5%)	
Zusatzqualifikation "Psychosomatische Grundversorgung" des Studienarztes		68 (28.1%)	58 (67.4%)	10 (6.4%)	<0.001
Jahr der Praxis-Gründung, #	Median	1997.0 [1987.0, 2001.0]	1999.0 [1993.8, 2006.0]	1989.0 [1985.5, 2000.0]	<0.001
Praxis-Gründung nach dem studienspezifischen Median		130 (53.7%)	50 (58.1%)	80 (51.3%)	0.277
Alter des Studienarztes (in Jahren)	Median	49.0 [44.0, 56.0]	51.0 [47.0, 54.5]	49.0 [43.0, 56.0]	0.387
	Mittelwert	50.5 ± 7.3	50.9 ± 6.9	50.3 ± 7.5	0.572

**Tabelle 8: Darstellung der Patienten- und Praxismerkmale der PARADIES-Studie, der PRoMPT-Studie sowie der Studienpopulation**

### 3.2 Vorläufige Studienpopulation: Zusammenhang zwischen PACIC-Werten und der Behandlung

Wie auf S. 43 beschrieben wurden in der vorläufigen Analyse die Interventions- und Kontrollgruppen sowohl der PARADIES- als auch der PRoMPT-Studie analysiert. Diese vorläufige Studienpopulation bestand aus  $n = 579$  Probanden. Durch diese vorläufige Analyse sollte gezeigt werden, dass die Studien hinsichtlich des PACIC-Wertes miteinander verglichen werden können. Im Anschluss an diese vorläufige Analyse wurden die Probanden der Kontrollgruppen beider Studien ausgeschlossen, sodass die endgültige Studienpopulation wie beschrieben aus  $n = 280$  Probanden bestand.

Die PACIC-Werte unterschieden sich in beiden Studien signifikant zwischen der Kontroll- und der Interventionsgruppe. Die medianen PACIC-Werte für die PARADIES-Studie lagen bei 44 % und 72 % ( $p < 0,001$ ) und für die PRoMPT-Studie bei 62 % und 70 % ( $p < 0,001$ ) für die Kontroll- bzw. Interventionsgruppe. Bei den PACIC-Werten gab es keinen signifikanten Unterschied zwischen den Interventionsgruppen der beiden Studien ( $p = 0,519$ ). Die Kontrollgruppe der PRoMPT-Studie zeigte statistisch signifikant höhere PACIC-Werte ( $p < 0,001$ ). Die Daten sind als Box-Plot in *Abbildung 5* und die deskriptive Statistik in *Tabelle 9* dargestellt. Beim Box-Plot sind die Medianwerte durch durchgezogene Linien, die Mittelwerte durch gestrichelte Linien gekennzeichnet.



**Abbildung 5: Box-Plot der Ergebnisse: PACIC und Studienarme**

	<b>Kontrolle PARADIES</b>	<b>Intervention PARADIES</b>	<b>Kontrolle PRoMPT</b>	<b>Intervention PRoMPT</b>
<b>N gültig/fehlend</b>	81/0	86/0	218/0	194/0
<b>Mittelwert</b>	47,8011	64,7418	62,9068	69,1664
<b>Median</b>	43,6364	71,8182	61,9091	69,0909
<b>Standardabweichung</b>	31,67064	24,96069	16,54269	17,18727
<b>Spannweite</b>	100,00	89,09	80,00	76,36
<b>Minimum</b>	0,00	10,91	20,00	23,64
<b>Maximum</b>	100,00	100,00	100,00	100,00

**Tabelle 9: Deskriptive Statistik der Ergebnisse: PACIC und Studienarme**

### 3.3 Endgültige Studienpopulation: Zusammenhang zwischen PACIC-Werten und der Behandlung

Wie bereits beschrieben, wurden nach der vorläufigen Analyse (siehe 3.2) die Probanden der Kontrollgruppen beider Studien ausgeschlossen, sodass die endgültige Studienpopulation aus  $n = 280$  Probanden bestand.

Innerhalb der endgültigen Studienpopulation wurden die Patienten- und Praxismerkmale der Probanden und Arztpraxen bzw. Studienärzte nicht zwischen den jeweiligen Studien (PARADIES oder PRoMPT) verglichen.

In einem linearen gemischten Modell, das für das Alter, das Geschlecht und die Suizidalität der Patienten zum Zeitpunkt T2 mit zufälligem Intercept pro Praxis adjustiert wurde, waren höhere PHQ-9-Werte zu T2 und eine stärkere Verbesserung der PHQ-9-Werte von Baseline zu T2 signifikant mit höheren PACIC-Werten assoziiert (Beta = 0,67, 95%-CI [0,02, 1,34]).

In einer Unteranalyse testeten wir für jedes der 11 PACIC-Items, ob sich das Item zwischen der Gruppe mit Depression (PHQ-9  $\geq 9$ ) und der Gruppe ohne Depression unterscheidet. Nominell signifikante Unterschiede gab es für zwei Items aus der PACIC-Subskala "Zielsetzung"

- "In den letzten 6 Monaten erhielt ich einen schriftlichen Behandlungsplan ausgehändigt" ( $p = 0,027$ )
- "In den letzten 6 Monaten wurde ich ermutigt zu einer bestimmten (Selbsthilfe-)Gruppe oder Einrichtung für meine Erkrankung zu gehen" ( $p = 0,025$ )

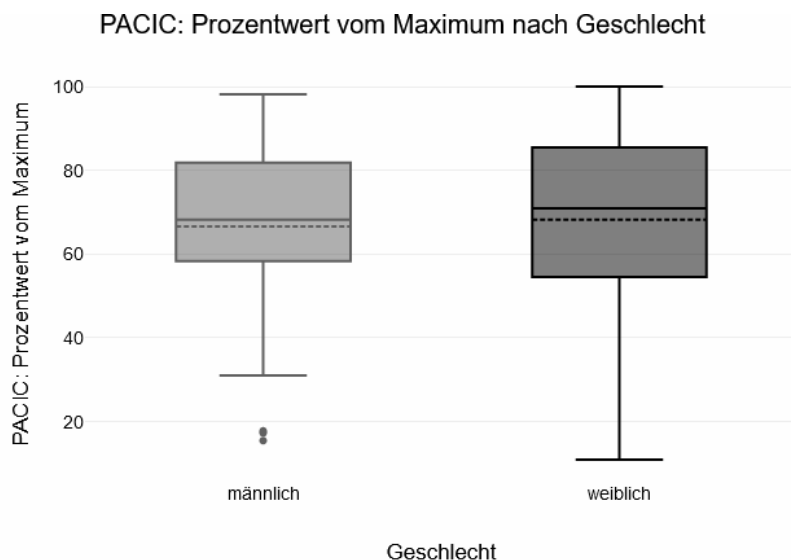
und für zwei Items der PACIC-Subskala "Problemlösung/kontextbezogene Beratung"

- "In den letzten 6 Monaten wurde mit mir ein Plan für meine Behandlung entworfen, den ich in meinem Alltag umsetzen kann" ( $p = 0,019$ )
- "In den letzten 6 Monaten wurde mir im Voraus erklärt, wie ich mit meiner Erkrankung auch in schwierigen Phasen umgehen kann" ( $p = 0,012$ ).

Es sei jedoch darauf hingewiesen, dass diese Unterschiede nach Korrektur für Mehrfachtests statistisch nicht mehr signifikant wären.

### 3.4 PACIC und Patientenmerkmale

Es gab keine signifikanten Unterschiede der PACIC-Werte zwischen männlichen und weiblichen Teilnehmern ( $p = 0,473$ ). Die Daten sind als Box-Plot in *Abbildung 6* und die deskriptive Statistik in *Tabelle 10* dargestellt. Beim Box-Plot sind die Medianwerte durch durchgezogene Linien, die Mittelwerte durch gestrichelte Linien gekennzeichnet.

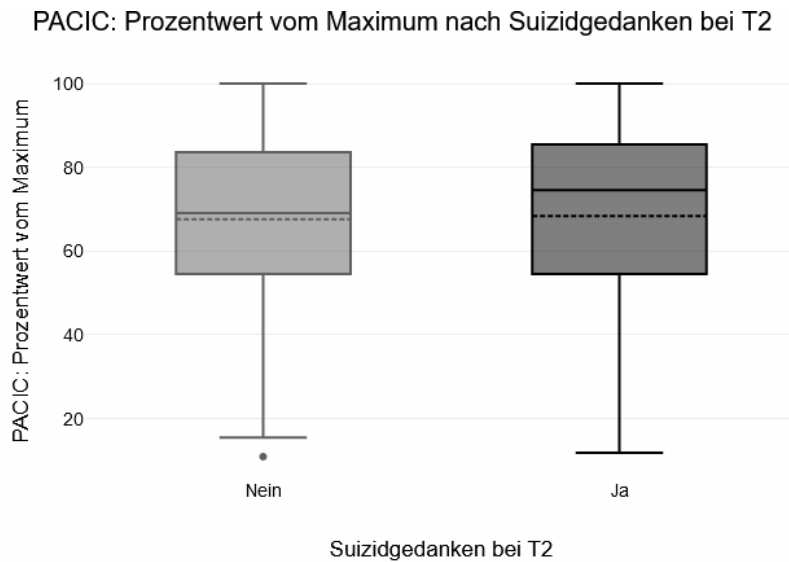


**Abbildung 6: Box-Plot der Ergebnisse: PACIC und Geschlecht**

	Männliches Geschlecht	Weibliches Geschlecht
<b>N gültig/fehlend</b>	71/0	209/0
<b>Mittelwert</b>	66,5578	68,2319
<b>Median</b>	68,1818	70,9091
<b>Standardabweichung</b>	19,16069	20,24957
<b>Spannweite</b>	82,73	89,09
<b>Minimum</b>	15,45	10,91
<b>Maximum</b>	98,18	100,00

**Tabelle 10: Deskriptive Statistik der Ergebnisse: PACIC und Geschlecht**

Es gab keine signifikanten Unterschiede der PACIC-Werte zwischen suizidgefährdeten und nicht suizidgefährdeten Personen gemäß PHQ-9 Item 9 zum Messzeitpunkt T2 ( $p = 0,751$ ). Die Daten sind als Box-Plot in *Abbildung 7* und die deskriptive Statistik in *Tabelle 11* dargestellt. Beim Box-Plot sind die Medianwerte durch durchgezogene Linien, die Mittelwerte durch gestrichelte Linien gekennzeichnet.



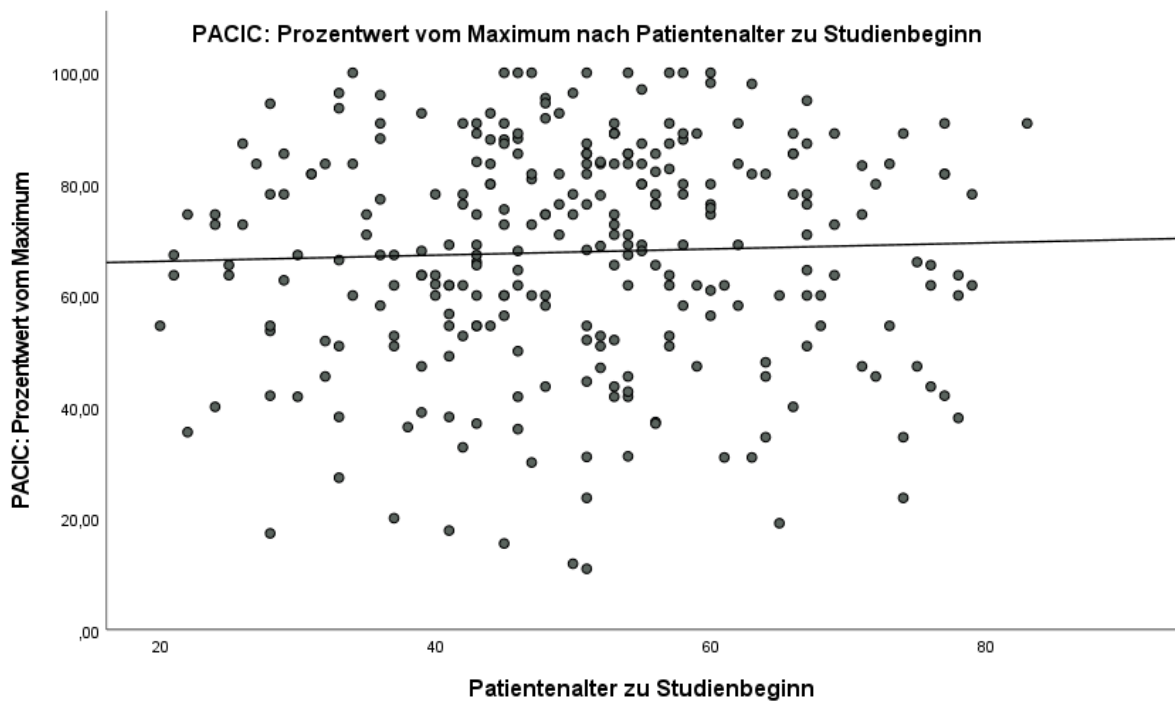
**Abbildung 7: Box-Plot der Ergebnisse: PACIC und Suizidgedanken**

	Keine Suizidgedanken vorhanden	Womöglich Suizidgedanken vorhanden
<b>N gültig/fehlend</b>	197/0	83/0
<b>Mittelwert</b>	65,5730	68,3636
<b>Median</b>	69,0909	74,5455
<b>Standardabweichung</b>	19,86389	20,29183
<b>Spannweite</b>	89,09	88,18
<b>Minimum</b>	10,91	11,82
<b>Maximum</b>	100,00	100,00

**Tabelle 11: Deskriptive Statistik der Ergebnisse: PACIC und Suizidgedanken**



Das Alter der Patienten war nicht signifikant mit den PACIC-Werten korreliert ( $r = 0,04$ , 95%-CI [-0,07, 0,15],  $p = 0,531$ ). Das Streudiagramm ist in *Abbildung 8* dargestellt.

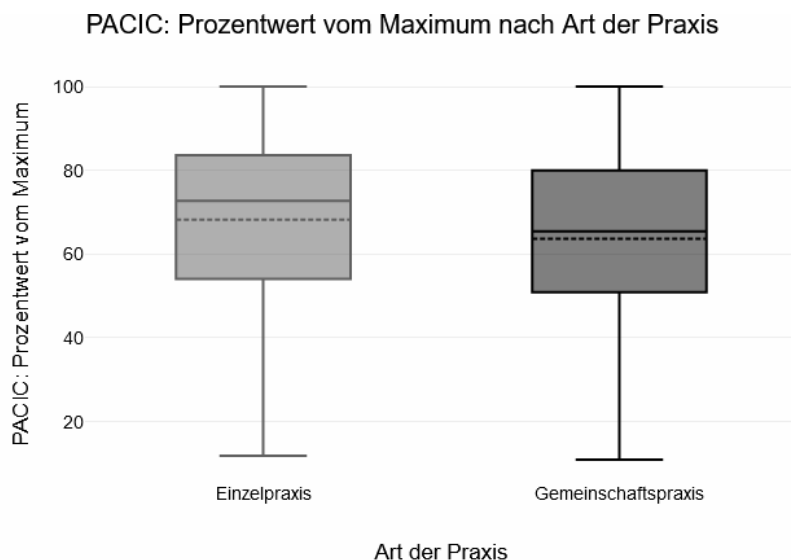


**Abbildung 8: Streudiagramm für Ergebnisse: PACIC und Patientenalter**

In einem linearen gemischten Modell, das das Alter, das Geschlecht und die Suizidgedanken der Patienten mit einem zufälligen Intercept pro Praxis einschloss, war keiner dieser Faktoren signifikant mit den PACIC-Werten verbunden (Alter: Beta = 0,04, 95%-CI [-0,14, 0,22]; weibliches Geschlecht: Beta = 1,16, 95%-CI [-4,11, 6,47]; Suizidgedanken: Beta = -1,30, 95%-CI [-7,07, 4,46]).

### 3.5 PACIC und Praxismerkmale

Es gab keine signifikanten Unterschiede der PACIC-Werte zwischen der Behandlung in einer Einzelpraxis oder in einer Gemeinschaftspraxis ( $p = 0,129$ ). Die Daten sind als Box-Plot in *Abbildung 9* und die deskriptive Statistik in *Tabelle 12* dargestellt. Beim Box-Plot sind die Medianwerte durch durchgezogene Linien, die Mittelwerte durch gestrichelte Linien gekennzeichnet.

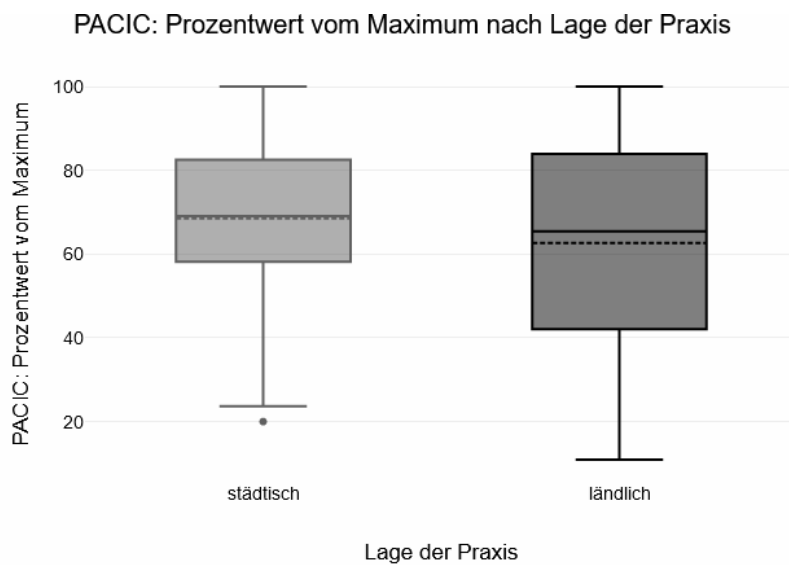


**Abbildung 9: Box-Plot der Ergebnisse: PACIC und Art der Praxis**

	Einzelpraxis	Gemeinschaftspraxis
<b>N gültig/fehlend</b>	151/0	91/0
<b>Mittelwert</b>	68,2175	63,69,05
<b>Median</b>	72,7273	65,4545
<b>Standardabweichung</b>	19,29185	21,64835
<b>Spannweite</b>	88,18	89,09
<b>Minimum</b>	11,82	10,91
<b>Maximum</b>	100,00	100,00

**Tabelle 12: Deskriptive Statistik der Ergebnisse: PACIC und Art der Praxis**

Es gab keine signifikanten Unterschiede der PACIC-Werte zwischen städtischen und ländlichen Praxen ( $p = 0,203$ ). Die Daten sind als Box-Plot in *Abbildung 10* und die deskriptive Statistik in *Tabelle 13* dargestellt. Beim Box-Plot sind die Medianwerte durch durchgezogene Linien, die Mittelwerte durch gestrichelte Linien gekennzeichnet.

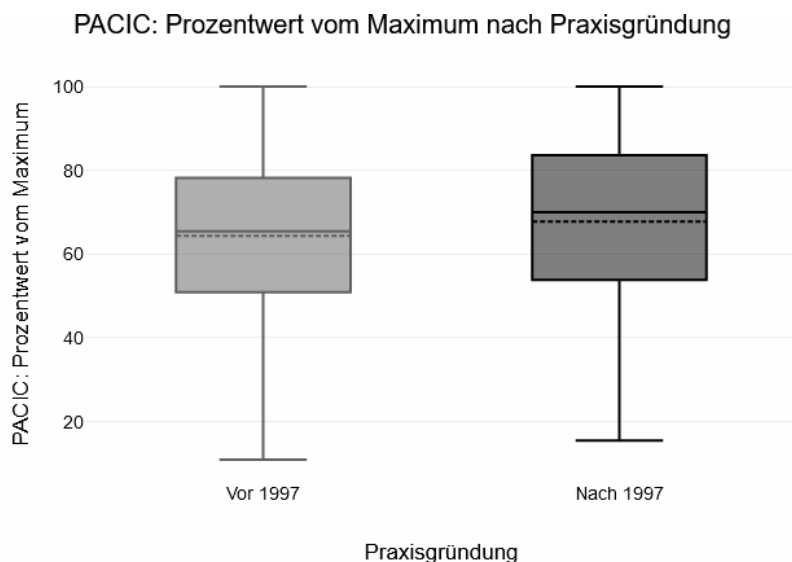


**Abbildung 10: Box-Plot der Ergebnisse: PACIC und Lage der Praxis**

	Städtische Lage	Ländliche Lage
<b>N gültig/fehlend</b>	160/0	82/0
<b>Mittelwert</b>	68,4951	62,6520
<b>Median</b>	69,0909	65,4545
<b>Standardabweichung</b>	16,68474	25,59296
<b>Spannweite</b>	80,00	89,09
<b>Minimum</b>	20,00	10,91
<b>Maximum</b>	100,00	100,00

**Tabelle 13: Deskriptive Statistik der Ergebnisse: PACIC und Lage der Praxis**

Es gab keine signifikanten Unterschiede der PACIC-Werte zwischen Praxen, die vor oder nach dem studienspezifischen Median (Jahr 1997) gegründet wurden ( $p = 0,143$ ). Die Daten sind als Box-Plot in *Abbildung 11* und die deskriptive Statistik in *Tabelle 14* dargestellt. Beim Box-Plot sind die Medianwerte durch durchgezogene Linien, die Mittelwerte durch gestrichelte Linien gekennzeichnet.



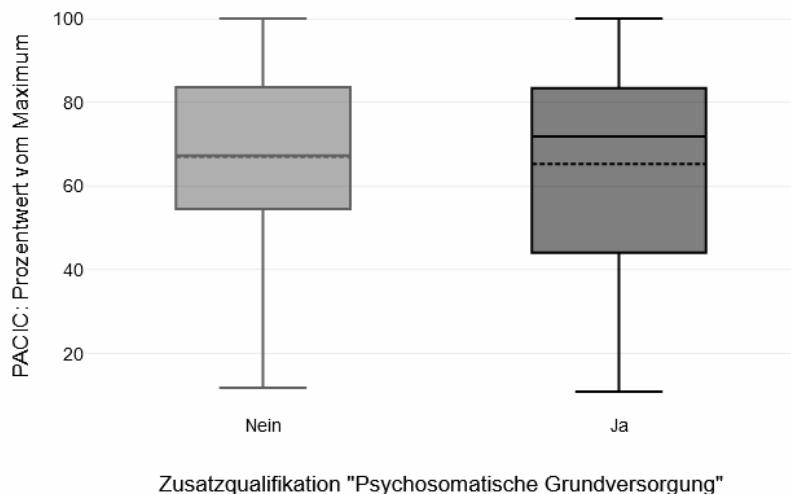
**Abbildung 11: Box-Plot der Ergebnisse: PACIC und Praxisgründung**

	Praxisgründung vor 1997	Praxisgründung nach 1997
<b>N gültig/fehlend</b>	117/0	106/0
<b>Mittelwert</b>	64,3796	67,7546
<b>Median</b>	65,4545	70,0000
<b>Standardabweichung</b>	20,15510	20,62021
<b>Spannweite</b>	89,09	84,55
<b>Minimum</b>	10,91	15,45
<b>Maximum</b>	100,00	100,00

**Tabelle 14: Deskriptive Statistik der Ergebnisse: PACIC und Praxisgründung**

Es gab keine signifikanten Unterschiede der PACIC-Werte zwischen Ärzten mit oder ohne Zusatzqualifikation "psychosomatische Grundversorgung" ( $p = 0,785$ ). Die Daten sind als Box-Plot in *Abbildung 12* und die deskriptive Statistik in *Tabelle 15* dargestellt. Beim Box-Plot sind die Medianwerte durch durchgezogene Linien, die Mittelwerte durch gestrichelte Linien gekennzeichnet.

PACIC: Prozentwert vom Maximum nach Zusatzqualifikation „Psychosomatische Grundversorgung“

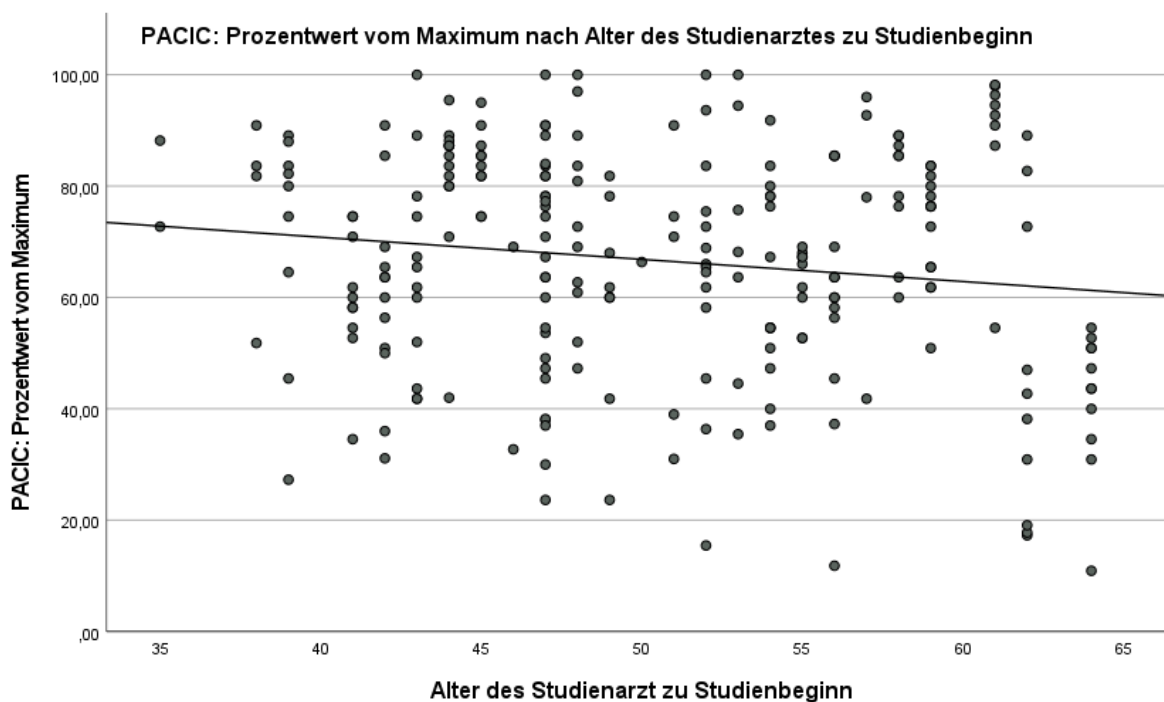


**Abbildung 12: Box-Plot der Ergebnisse: PACIC und Zusatzqualifikation "Psychosomatische Grundversorgung"**

	Zusatzqualifikation „Psychosomatische Grundversorgung“ vorhanden	Zusatzqualifikation „Psychosomatische Grundversorgung“ nicht vorhanden
<b>N gültig/fehlend</b>	174/0	68/0
<b>Mittelwert</b>	66,9950	65,2873
<b>Median</b>	67,2727	71,8182
<b>Standardabweichung</b>	18,44526	24,48470
<b>Spannweite</b>	88,18	89,09
<b>Minimum</b>	11,82	10,91
<b>Maximum</b>	100,00	100,00

**Tabelle 15: Deskriptive Statistik der Ergebnisse: PACIC und Zusatzqualifikation "Psychosomatische Grundversorgung"**

Dennoch gab es nominale Unterschiede in den PACIC-Werten nach diesen Merkmalen. Dies könnte durch das Alter des behandelnden Arztes erklärt werden: Das Alter des Arztes war signifikant negativ mit den PACIC-Werten korreliert ( $r = -0,14$ , 95%-CI  $[-0,27, -0,02]$ ,  $p = 0,03$ , siehe Abbildung). Ärzte in städtischen Praxen waren signifikant jünger als Ärzte in ländlichen Praxen (Median 47 Jahre vs. 53 Jahre,  $p < 0,001$ ). Ärzte in jüngeren Praxen waren signifikant jünger als Ärzte in älteren Praxen (Median 45 Jahre vs. 56 Jahre,  $p < 0,001$ ). *Abbildung 13* zeigt die Korrelation zwischen den PACIC-Werten (y-Achse) und dem Alter des Hausarztes in Jahren (x-Achse), das signifikant negativ korreliert war.



**Abbildung 13: Streudiagramm für Ergebnisse: PACIC und Alter des Studienarztes**

Es gab keine signifikanten Altersunterschiede zwischen Ärzten in Einzel- und Gruppenpraxen, zwischen Ärzten in den beiden Studien oder zwischen Ärzten mit und ohne Zusatzqualifikation "Psychosomatische Grundversorgung".

In einem Regressionsmodell mit gemischten Effekten, das die durch die Hausarztpraxen gegebene Clusterstruktur berücksichtigt, war ein höheres Alter des behandelnden Arztes signifikant mit niedrigeren PACIC-Werten assoziiert (Beta =  $-0,62$ , 95%-CI  $[-1,18, -0,07]$ ; dies entspricht einem Rückgang des PACIC um 0,62 Prozentpunkte bei einem Anstieg des Alters des behandelnden Arztes um ein Jahr). Der Zusammenhang blieb auch nach Adjustierung für den Praxistyp und das Vorhandensein der Zusatzqualifikation "Psychosomatische Grundversorgung" bestehen, schwächte sich jedoch nach Adjustierung für den Praxisstandort und den Zeitpunkt der Gründung ab.

## 4. Diskussion

### 4.1 Zusammenfassung der Ergebnisse

Diese Arbeit untersuchte, inwieweit die Zufriedenheit depressiver Patienten mit ihrer Behandlung in einer Studiumgebung von Patienten- oder Praxismerkmalen beeinflusst wird. Die Patientenzufriedenheit wurde hierbei mithilfe des PACIC-Fragebogens gemessen. Eine zusätzliche Schwierigkeit bestand darin, dass die von uns untersuchten Studien unterschiedliche Versionen des PACIC-Fragebogens hinsichtlich der Anzahl der Items und der Art der Antwortskalen nutzten. Um eine größtmögliche Vergleichbarkeit zu ermöglichen, wurden nur die identischen Items der Fragebögen in die statistische Auswertung eingeschlossen. Bezüglich der verschiedenen Antwortskalen erstellten wir die Variable *PACIC: Prozentwert vom Maximum*, sodass wir die verwendeten Antwortskalen nicht auf das Maß einer anderen Skala umrechnen mussten. Dadurch erhielten wir eine dimensionslose Zahl, die ebenfalls als Prozentwert abgelesen und mit den Werten der anderen Studien verglichen werden kann.

In einer zu Beginn durchgeführten Analyse wurde eine höhere Patientenzufriedenheit in den Interventionsgruppen – verglichen mit den Kontrollgruppen – bestätigt. Die Unterschiede der jeweiligen Studie zwischen Kontroll- und Interventionsgruppen waren statistisch signifikant.

In der anschließenden Analyse konnten auf Ebene der Patientenmerkmale keine signifikanten Unterschiede der Patientenzufriedenheit berichtet werden. Dies betraf die Patientenmerkmale *Alter*, *Geschlecht* und das Vorhandensein von Suizidgedanken.

Auf Ebene der Praxismerkmale waren die *Art der Praxis* (Einzel- oder Gemeinschaftspraxis), die *Lage der Praxis* (städtisch oder ländlich), die Einteilung in ältere oder neuere Praxen sowie die mögliche Zusatzqualifikation *Psychosomatische Grundversorgung* des Studienarztes nicht signifikant mit der Patientenzufriedenheit verbunden. Hingegen war das Alter des Studienarztes statistisch negativ signifikant mit der Patientenzufriedenheit assoziiert, sodass Patienten jüngerer Studienärzte von einer statistisch signifikant höheren Zufriedenheit berichteten.

### 4.2 Interpretation der Ergebnisse

#### 4.2.1 Relevanz für die Hausarztpraxis

Die Patientenzufriedenheit spielt in einer Hausarztpraxis eine entscheidende Rolle und sollte stets im Auge behalten werden.

Wie ich unter 1.3 Depression ausgeführt habe, sind Depressionen eine Erkrankung mit heterogenem Erscheinungsbild sowie einem chronischen Verlauf, die häufig in der Primärversorgung auftritt. Ärzte der Primärversorgung spielen eine wichtige Rolle bei der Behandlung von Depressionen, und stehen bei der Diagnose und Behandlung der Erkrankung vor Herausforderungen (Ferenchick et al. 2019).

Zufriedene Patienten sind ein Indikator für eine qualitativ hochwertige medizinische Versorgung. Battaar et al. zufolge sind Indikatoren für die Qualität der Gesundheitsversorgung die einflussreichsten

Determinanten der Patientenzufriedenheit. Hierbei sei insbesondere die Qualität der zwischenmenschlichen Betreuung durch das Praxisteam die wichtigste Determinante für die Patientenzufriedenheit (Batbaatar et al. 2017). Williams & Calnan haben die Bedeutung der Arzt-Patienten-Beziehung für die Patientenzufriedenheit ebenfalls belegt (Williams und Calnan 1991). Bei einer hohen Patientenzufriedenheit ist anzunehmen, dass sich Patienten gut betreut fühlen, und eher bereit sind, offen über ihre gesundheitlichen Probleme, wie z.B. Depressionen, zu sprechen. Dies wiederum verbessert nicht nur die Diagnostik, sondern im Folgenden auch die Behandlung der Patienten. Eine hohe Patientenzufriedenheit ermöglicht somit eine bessere medizinische Versorgung und steigert den Behandlungserfolg.

Die Einbeziehung der Patientensicht auf die erhaltene Behandlung ist ein wichtiger Forschungsansatz, um eine adäquate Behandlung zu gewährleisten. Diese Forschungsarbeit trug Erkenntnisse dazu bei, welche Faktoren die Patientenzufriedenheit depressiver Patienten in Case-Management-Programmen beeinflussen. Diese Erkenntnisse sollen einen direkten Einfluss auf weitere Forschungsarbeiten und letztlich die Behandlung betreffender Patienten haben.

#### **4.2.2 PACIC: Vergleich der Item-Anzahl**

Während die PARADIES-Studie den kürzeren PACIC-Fragebogen mit 11 Items nutzte, wurde in der PRoMPT-Studie der längere PACIC-Fragebogen mit 20 Items, also 9 zusätzlichen Items, verwendet. Wir entschieden uns dazu, nur die 11 identischen Items auszuwerten.

Generell ist die Vergleichbarkeit von Fragebögen ein komplexes Problem, das auftritt, wenn ein Fragebogen geändert wird. Der PACIC-Fragebogen wurde von zunächst 20 auf 11 Items gekürzt. Die Vergleichbarkeit von Fragebögen bezieht sich auf das Ausmaß, in dem die Ergebnisse eines Fragebogens im Zeitverlauf oder zwischen verschiedenen Gruppen von Befragten verglichen werden können. Wenn ein Fragebogen zusätzlich noch verändert wird, zum Beispiel gekürzt wird, kann die Vergleichbarkeit aufgrund mehrerer Faktoren eingeschränkt sein.

Einer der Hauptgründe für die erschwerte Vergleichbarkeit ist, dass der längere Fragebogen zu einer höheren Belastung der Probanden führen kann. Die Probanden könnten aufgrund der zusätzlichen Items schneller das Interesse verlieren, was zu einem höheren Risiko nachlässiger oder unvollständiger Antworten führen könnte. Dies kann zu Messfehlern führen und die Zuverlässigkeit der Daten verringern (Sharma 2022). Wenn das Ausfüllen des Fragebogens aufgrund zusätzlicher Items mehr Zeit in Anspruch nimmt, kann sich dies auf die Art und Weise auswirken, wie die Probanden mit dem Fragebogen umgehen, wie sie die Items bearbeiten und beantworten (Chudoba und B. 2020).

Darüber hinaus könnten die zusätzlichen Items einen anderen Schwierigkeitsgrad oder eine andere kulturelle Prägung aufweisen oder von den Probanden unterschiedlich interpretiert oder verstanden werden, was die Vergleichbarkeit der Ergebnisse beeinträchtigen könnte.

Um die bestmögliche Vergleichbarkeit zu gewährleisten, entschieden wir uns dazu, nur die 11 identischen Items auszuwerten. Natürlich macht es jedoch einen Unterschied, dass die Probanden der PRoMPT-Studie den längeren PACIC-Fragebogen ausfüllten. Wir schätzen diese Unsicherheit jedoch als geringfügig ein, da die identischen 11 Items über den Fragebogen mit 20 Items verteilt waren. Es



handelt sich bei den 11 identischen Items somit nicht um die zusammenhängenden letzten Items des Fragebogens, bei deren Beantwortung die Konzentration bereits signifikant niedriger sein könnte.

### 4.2.3 PACIC: Vergleich der Antwort-Skalen

Während die PARADIES-Studie den PACIC-Fragebogen mit einer Prozent-Skala und 11 Skalenstufen nutzte, wurde in der PRoMPT-Studie der PACIC-Fragebogen mit einer Likert-Skala und 5 Skalenstufen verwendet. Beim PACIC-Fragebogen mit der Prozent-Skala wurden nur die Extremwerte mit „Nie“ und „Immer“ benannt, beim PACIC-Fragebogen mit der Likert-Skala hingegen alle 5 Stufen von „So gut wie nie“ bis letztlich „Fast immer“ benannt.

Antwortskalen sind ein wichtiger Bestandteil der Forschung, da sie eine standardisierte Möglichkeit bieten, die Antworten verschiedener Teilnehmer zu messen und zu vergleichen. Der Vergleich verschiedener Antwortskalen kann jedoch aufgrund verschiedener Faktoren schwierig sein.

Wie Arditi et al. in ihrer Metaanalyse zeigten, war die Verwendung einer Likert-Skala mit 5 Skalenstufen - verglichen mit einer Skala mit 11 Skalenstufen - signifikant mit höheren PACIC-Werten verbunden (Arditi et al. 2018). Dies wird damit begründet, dass die Anzahl der Antwortmöglichkeiten und die Bezeichnung der einzelnen Skalenstufen die Zustimmung zu den Mittel-, den Minimal- und den Maximalwerten erhöhe.

Zwar können Skalen mit mehr Stufen präziser, jedoch auch kognitiv anspruchsvoller für die Probanden sein. Generell ist es jedoch fraglich, welcher Informationsgewinn zu erwarten ist, wenn Patienten mithilfe einer elfstufigen Prozent-Skala hinsichtlich ihrer Patientenzufriedenheit der letzten 6 Monate befragt werden, wobei sie angeben sollen, in wie viel Prozent der Fälle verschiedene Items zutrafen. Eine adäquate Antwort sollte auch dann hilfreich sein, wenn nur 5 Antwortmöglichkeiten zur Verfügung stehen. Bezieht man bei der Beantwortung der Fragebögen die Tendenzen zu Mittel- und Extremwerten (Kalton et al. 1980; Greenleaf 1992) mit ein, erscheint die Wahl einer Skala mit 11 statt 5 Antwortmöglichkeiten hinsichtlich ihres Nutzens fragwürdig. Diese Überlegung folgt der Berechnung, dass Mittel- und Extremwerte bei einer Skala mit 5 Antwortmöglichkeiten 60 %, auf einer Skala mit 11 Antwortmöglichkeiten nur knapp 27 % der Antwortmöglichkeiten darstellen.

Ein weiterer Grund für die Schwierigkeit, verschiedene Antwortskalen miteinander zu vergleichen, liegt in der unterschiedlichen Bezeichnung der Minimal- und Maximalwerte. So verwendet der PACIC-Fragebogen mit der Likert-Skala die Werte „So gut wie nie“ und „Fast immer“ für die Extremwerte, während die Prozent-Skala die niedrigste Stufe mit „Nie“ und die höchste Stufe mit „Immer“ definiert. Generell kann die Bezeichnung der Extremwerte die Art und Weise beeinflussen, wie Teilnehmer diese Items interpretieren und beantworten, was einen Vergleich von Antworten unterschiedlicher Skalen erschwert.

Wie im Methodenteil dargestellt, bestand zunächst die Überlegung, die Prozent-Skala auf eine Likert-Skala umzurechnen. Jedoch zeigten sich hier aufgrund der unterschiedlichen Bezeichnung der Extremwerte erhebliche Unsicherheiten in der statistischen Auswertung. Sind „100 %“ der Prozent-Skala gleichbedeutend mit „Fast immer“ der Likert-Skala? Da „100 %“ rein logisch mehr als „Fast immer“ sind, müssten auch „90 %“ der Prozent-Skala mit „Fast immer“ identisch sein.

Um durch die Umrechnung verschiedener Antwortskalen möglichst keinen Informationsverlust zu riskieren, entschieden wir uns dazu, die Variable *PACIC: Prozentwert vom Maximum* zu erstellen und für unsere Analyse auszuwerten. Somit konnten die statistischen Daten unberührt und lediglich ins Verhältnis zum Maximalwert des jeweiligen Fragebogens gesetzt werden. Dieser Ansatz erwies sich als praktikabel, sodass sich die Ergebnisse der unterschiedlichen PACIC-Fragebögen gut vergleichen ließen. Obwohl die Umrechnung der Werte einer Prozent-Skala in die Werte einer Likert-Skala mithilfe von Statistikprogrammen wie SPSS unkompliziert ist, entschieden wir uns dagegen, um einen Informationsverlust zu vermeiden. Wenngleich der Unterschied zwischen „90 %“ und „100 %“ bei der Beantwortung einer Frage minimal zu sein scheint, bleibt durch unseren Ansatz trotzdem die – manchmal auch nur geringe – Diskrimination zwischen einzelnen Skalenstufen bestehen.

Gugiu et al. kritisierten die Likert-Skala mit 5 Skalenstufen, da für eine Auswertung des Gesamtwerts sowie die Subskalen Mittelwerte von Ordinalskalen ungeeignet seien (Gugiu et al. 2009; Gugiu et al. 2010). Die notwendige Äquidistanz zwischen den einzelnen Stufen der Antwortskala sei nur bei einer Intervall- oder Verhältnisskala gegeben (Jöreskog 1994). Wir schätzten dieses Risiko der Unsicherheit als gering ein, da wir mit der Variable *PACIC: Prozentwert vom Maximum* eine dimensionslose Zahl als zusätzliche Variable einführten, die das Konstrukt des jeweiligen Fragebogens nicht berührt oder beeinflusst. Da aber bei der Likert-Skala alle Skalenstufen eine genaue Bezeichnung haben („So gut wie nie“, „Meistens nicht“, „Gelegentlich“, „Meistens“ und „Fast immer“), könnte man auch hier Prozentzahlen von 0 % bis 100 % für die einzelnen Stufen vergeben und als Gesamtwert den arithmetischen Mittelwert bilden. Während „So gut wie nie“ dann wahrscheinlich mit „0 %“ gleichgesetzt würde, würde „Meistens nicht“ „25 %“, „Gelegentlich“ „50 %“, „Meistens“ „75 %“ und „Fast immer“ „100 %“ entsprechen.

#### 4.2.4 PACIC: Umgang mit fehlenden Werten

Bei den Einschlusskriterien unserer Studienpopulation gab es die Herausforderung, dass es für den PACIC-Fragebogen keine offiziellen Richtlinien für den Umgang mit fehlenden Werten gibt. Wie im Methodenteil dargestellt, gab es bisher verschiedene Ansätze, wie Forschende mit fehlenden Werten umgingen.

Der Prozentsatz der fehlenden Antworten in einer Studie, der in einem Fragebogen toleriert werden kann, wird hier als Schwellenwert für fehlende Daten bezeichnet. Der Schwellenwert sollte auf der Grundlage des studienspezifischen Studiendesigns und der allgemeinen Forschungsfrage festgelegt werden. Wird der Schwellenwert zu hoch angesetzt, kann dies zu einer Verzerrung der Ergebnisse führen, da sich Probanden mit fehlenden Daten systematisch von Probanden unterscheiden können, die ausreichend Daten geliefert haben (Mack et al. 2018). Außerdem kann ein hoher Schwellenwert (es werden mehr fehlende Daten akzeptiert) zu einer Verringerung der statistischen Aussagekraft führen, wodurch es schwieriger wird, aussagekräftige Unterschiede in der Stichprobe festzustellen. Andererseits kann ein zu niedrig angesetzter Schwellenwert (es werden weniger fehlende Daten akzeptiert) zum Ausschluss von Probanden führen, deren Daten zwar unvollständig sind, die aber trotzdem wertvolle Informationen für die Studie beisteuern könnten.

Wir entschieden uns dazu, dass nur Probanden eingeschlossen werden, bei denen maximal 6 der 11 Items fehlten. Wir folgten damit den Empfehlungen von Goetz et al. und Gugiu et al., da der PACIC-Fragebogen mithilfe des arithmetischen Mittelwerts der Items ausgewertet wird und eine hohe interne Konsistenz aufweist. Dadurch ist der PACIC-Fragebogen unempfindlich gegenüber einzelnen fehlenden Items.

#### **4.2.5 PACIC: Zusammenhänge mit Patientenmerkmalen**

In unserer statistischen Auswertung war die Patientenzufriedenheit statistisch nicht signifikant mit Patientenmerkmalen verbunden. Dies betraf das Alter und das Geschlecht der Patienten sowie das mögliche Vorhandensein von Suizidgedanken. Die Zufriedenheit mit der Behandlung aus Sicht des Patienten ist ein multidimensionales Konzept, das von einer Vielzahl von Faktoren beeinflusst wird, darunter die Erkrankung, das Behandlungsteam und der Arzt, aber auch Patientenmerkmale wie Alter oder Geschlecht.

Während Rosemann et al. (Rosemann et al. 2007; Rosemann et al. 2008), und Ali et al. (Ali et al. 2017) eine höhere Versorgungsqualität bei Frauen maßen, war bei Rick et al. (Rick et al. 2012) und Bo et al. (Bo et al. 2020) das weibliche Geschlecht mit einer niedrigeren Patientenzufriedenheit verbunden. Wohingegen bei Ludt et al. (Ludt et al. 2012) das männliche Geschlecht mit einer höheren Patientenzufriedenheit verbunden war. Gemeinsam haben all diese Studien, dass sie nicht speziell Patienten mit einer psychiatrischen Erkrankung untersuchten. In den von uns untersuchten Studien hingegen lag der Fokus explizit auf psychiatrischen Patienten, teilweise mit mehreren psychiatrischen Diagnosen wie Depression und Agoraphobie, wohingegen die oben genannten Studien Patienten mit Erkrankungen wie Osteoarthritis, Diabetes oder Herzerkrankungen untersuchten. Da bei den untersuchten Erkrankungen dieser Studien womöglich auch das Geschlecht durch eine höhere Prävalenz oder eine andere Symptomatik mit der Erkrankung assoziiert ist, könnte dies die geschlechtsspezifischen Unterschiede der PACIC-Werte erklären. Bei psychiatrischen Erkrankungen wie Depression könnten die Einschränkungen und Bedürfnisse bei Männern und Frauen ähnlich sein, sodass eine ähnliche Patientenzufriedenheit berichtet wird.

Eine weitere Erklärung dafür, dass in unserer Analyse die Patientenmerkmale nicht statistisch signifikant mit der Patientenzufriedenheit verbunden waren, liegt in dem Studiumfeld begründet. Generell ist anzunehmen, dass Patientenmerkmale wie Alter oder Geschlecht eine Rolle bei der Zufriedenheit mit der Behandlung aus Patientensicht spielen. In einem standardisierten Studiumfeld mit geplanten Abläufen und Maßnahmen sind die Auswirkungen der Patientenmerkmale jedoch möglicherweise minimal. Wenn die Dauer eines Arztbesuches oder der Ablauf der Therapie aufgrund standardisierter Bedingungen nicht variiert, könnten Patientenmerkmale nur eine untergeordnete Rolle spielen. Eine mögliche Erklärung ist außerdem, dass sich die Patienten der wissenschaftlich geprägten Studiensituation bewusst waren. Dies könnte dazu geführt haben, dass die Probanden die Fragebögen möglichst objektiv beantworten wollten, obwohl möglicherweise die Zufriedenheit mit der Versorgung aus Patientensicht mit dem Alter oder dem Geschlecht verbunden war.

Zusätzlich lässt sich anführen, dass auch Glasgow et al. (Glasgow et al. 2005b), Aragonés et al. (Aragonés et al. 2008) und Taggart et al. (Taggart et al. 2010) keine Assoziationen zwischen Patientenmerkmalen und der Patientenzufriedenheit berichteten.

#### 4.2.6 PACIC: Zusammenhänge mit Praxismerkmalen

Unsere statistische Auswertung ergab hinsichtlich der Praxismerkmale *Art der Praxis*, *Lage der Praxis*, der Einteilung der Praxen bezogen auf den studienspezifischen Median in *neuere oder ältere Praxen* sowie der *Zusatzqualifikation „Psychosomatische Grundversorgung“* des Studienarztes keine statistisch signifikanten Unterschiede der Patientenzufriedenheit. Statistisch signifikant war jedoch, dass Patienten jüngerer Studienärzte eine höhere Patientenzufriedenheit berichteten.

Die Zufriedenheit mit der Behandlung aus Sicht der Patienten kann von verschiedenen Merkmalen der Arztpraxis abhängen, z.B. davon, ob es sich um eine Einzel- oder Gemeinschaftspraxis handelt, ob sie in einem städtischen oder ländlichen Gebiet liegt oder ob es sich um eine neuere oder ältere Praxis handelt.

Ein Merkmal der Praxis, das sich ebenfalls auf die Zufriedenheit mit der Behandlung aus Patientensicht auswirken kann, ist die Frage, ob es sich um eine Einzel- oder Gemeinschaftspraxis handelt. Gemeinschaftspraxen bieten möglicherweise ein breiteres Spektrum an Dienstleistungen an und beschäftigen mehr Ärzte, was zu einer umfassenderen und effizienteren Versorgung führen kann. Andererseits bieten Einzelpraxen möglicherweise eine individuellere und kontinuierlichere Betreuung, da die Patienten in der Regel von demselben Arzt behandelt werden. Dies kann zu einem besseren Verständnis der gesundheitlichen Vorgeschichte des Patienten, seiner Präferenzen und Erwartungen an die Behandlung führen, sodass Patienten ein höheres Vertrauen sowie eine höhere Zufriedenheit empfinden.

Dass sich hinsichtlich der Art der Praxis keine signifikanten Unterschiede zwischen Einzel- und Gemeinschaftspraxis fanden, könnte daran gelegen haben, dass sich die genannten Effekte ausgeglichen haben. Patienten einer Gemeinschaftspraxis könnten von einer besseren Organisation der Arbeitsabläufe innerhalb des Praxisteam profitieren. Da der PACIC-Fragebogen speziell die Organisation der Therapie und die Betreuung des Patienten abfragt, könnte eine bessere Organisation der Arbeitsabläufe in einer Gemeinschaftspraxis generell die Versorgungsqualität aus Patientensicht erhöhen. Andererseits könnte die persönlichere Betreuung in einer Einzelpraxis sowie das Wissen des Praxisteam um die persönlichen Bedürfnisse der Patienten eine höhere Versorgungsqualität aus Patientensicht bedingen. Es ist grundsätzlich möglich, dass die Art der Praxis einen Einfluss auf die Patientenzufriedenheit hat, die Vor- und Nachteile der Praxisformen sich jedoch ausgleichen. In der Studie von Petersen et al. war die Behandlung in einer Einzelpraxis mit höheren PACIC-Werten verbunden als die Behandlung in einer Gemeinschaftspraxis (Petersen et al. 2014). Diese Ergebnisse können daran liegen, dass in dieser Studie multimorbide Patienten untersucht wurden, deren individuelle Bedürfnisse im vertrauteren Rahmen einer Einzelpraxis womöglich besser entsprochen werden kann. In der Studie von Taggart et al. wiederum waren Praxismerkmale nicht mit den PACIC-Werten verbunden (Taggart et al. 2010).

Die Tatsache, dass unsere Auswertung keine signifikanten Unterschiede zwischen der Behandlung in einer städtischen oder einer ländlichen Arztpraxis ergab, könnte an dem Studiendesign und damit der standardisierten Betreuung der Patienten liegen. Generell kann es in ländlichen Gebieten aufgrund der sinkenden Anzahl an Hausärzten schwierig sein, einen Hausarzt zu finden, der noch Kapazitäten für neue Patienten hat (Bundesministerium für Gesundheit 2016). Ebenso verfügen Praxen in ländlichen Gebieten möglicherweise über weniger personelle Ressourcen, um eine bessere Versorgung der Patienten zu ermöglichen. Jedoch können sich durch gefestigtere Strukturen in ländlichen Gebieten auch Effekte ergeben, dass die persönliche Beziehung zwischen den Patienten und dem gesamten Praxisteam verbessern und somit eine höhere Versorgungsqualität ermöglichen. Andererseits können städtische Praxen zwar womöglich über mehr Ressourcen verfügen, jedoch auch mit Herausforderungen wie langen Wartezeiten oder einer steigenden Patientenzahl konfrontiert sein.

Auch der Zeitraum seit Gründung der Arztpraxis kann sich auf die Patientenzufriedenheit auswirken. In neueren Praxen werden möglicherweise eher elektronische Gesundheitsakten und neue Technologien zur Patientenverwaltung und -behandlung zur Verbesserung der Effizienz und Qualität der Versorgung eingesetzt; möglicherweise verfügen die Praxisteams aber nicht über das gleiche Maß an Erfahrung etablierter Beziehungen zu den Patienten. Ältere Praxen hingegen verfügen möglicherweise über mehr Erfahrung und etablierte Beziehungen zu den Patienten sowie eine Kontinuität in der Versorgung. Ähnlich wie bei der Unterscheidung zwischen der Art und der Lage der Praxis könnte es auch bei dem Alter einer Arztpraxis sein, dass sich die jeweiligen Effekte nicht gegensätzlich, sondern gleichgerichtet, auf die Patientenzufriedenheit auswirken.

Unsere Auswertung hat gezeigt, dass die Zufriedenheit mit der Versorgung aus Patientensicht nicht signifikant von der Zusatzqualifikation „Psychosomatische Grundversorgung“ abhängt. Hierfür gibt es mehrere Erklärungen. Erstens verfügen Hausärzte möglicherweise bereits über die erforderlichen Fähigkeiten und Kenntnisse für eine wirksame Behandlung von Depressionen, sodass eine zusätzliche Schulung keinen wesentlichen Mehrwert bringt. Laut Bayerischer Landesärztekammer beträgt der Umfang der Weiterbildung 80 Unterrichtseinheiten à 45 Minuten sowie „eine mindestens 3-jährige Erfahrung in selbstverantwortlicher ärztlicher Tätigkeit in einer psychosomatisch orientierten Krankheitslehre“ (Bayerische Landesärztekammer (BLÄK)). Obwohl man annehmen könnte, dass eine Weiterbildung von nicht einmal 80 Zeitstunden sowie selbstverantwortliche Erfahrungen nicht ausreichen, um eine höhere Versorgungsqualität auszumachen, muss an dieser Stelle differenziert werden. Die Frage, ob die Versorgungsqualität von der Zusatzqualifikation „Psychosomatische Grundversorgung“ abhängt, ist schwierig zu beantworten, da die Fortbildung primär dem Therapieergebnis dient. Das Therapieergebnis ist aber nicht prinzipiell mit der Patientenzufriedenheit gleichzusetzen. Es ist also anzunehmen, dass Ärzte mit der Zusatzqualifikation „Psychosomatische Grundversorgung“ die Therapie ihrer Patienten verbessern. Ferner spiegeln die standardisierten Bedingungen einer klinischen Studie möglicherweise nicht genau die Komplexität der hausärztlichen Praxis wider. Dies unterstreicht den Punkt, dass das Therapieergebnis nicht mit der Versorgungsqualität im Sinne der Organisation und Betreuung des Patienten gleichzusetzen ist.

Unsere statistische Auswertung hat gezeigt, dass die Patienten jüngerer Studienärzte eine signifikant höhere Patientenzufriedenheit berichten. Eine mögliche Erklärung ist, dass jüngere Ärzte aufgrund der geringeren Zeit zwischen dem Medizinstudium und ihrer Tätigkeit eher auf dem neuesten Stand des medizinischen Fortschritts und der Behandlung sind. Es könnte sein, dass jüngere Ärzte besser mit den aktuellen Leitlinien und Therapiealgorithmen vertraut sind und dadurch eine stärker auf Evidenz basierte Versorgung anbieten.

Eine weitere Erklärung ist, dass jüngere Ärzte offener für neue Ideen und Versorgungsansätze sind. Jüngere Ärzte sind möglicherweise eher bereit, neue Technologien wie moderne und intelligente Patientenverwaltungsprogramme in ihren Arztpraxen einzusetzen. Diese Programme können eine bessere Koordinierung der Patientenversorgung, z.B. durch gezielte Einladungen zu Routineuntersuchungen sowie eine verbesserte Kommunikation mit ärztlichen Kollegen, bewirken. Dies kann zu besseren Behandlungsergebnissen führen. Auch könnte es sein, dass jüngere Ärzte in ihrer Behandlungsroutine noch relativ unerfahren sind und noch keine „professionelle Rolle“ eingenommen haben und sich daher eher an die Vorgaben eines Studiensettings halten (Dreyfus und Dreyfus 2016; Durand et al. 2018). Generell könnte unter jüngeren Ärzten ein größeres Interesse an Qualitätsverbesserung für ihre Praxen vorliegen. Möglicherweise sind jüngere Ärzte auch empfänglicher für Feedback und Kritik und sind eher bereit, auf Beschwerden der Patienten Änderungen in ihren Abläufen vorzunehmen.

Es soll in der Diskussion unserer Ergebnisse jedoch nicht der Eindruck erweckt werden, als ob ältere Ärzte jüngeren Ärzten unterlegen sind. Ältere Ärzte verfügen über einen größeren Erfahrungsschatz und ein umfangreiches Wissen, was besonders bei komplexen und schwierigen Patientenfällen von großer Bedeutung sein kann. So ist es z.B. möglich, dass die Patienten in den untersuchten Studien nur über eine höhere Patientenzufriedenheit berichteten, weil die jüngeren Studienärzte sich stärker an die inhaltlichen Angaben des Studiendesigns hielten. Ihre geringere Erfahrung könnte die jüngeren Studienärzte dahingehend beeinflusst haben, dass sie den Patientengesprächen mehr Zeit eingeräumt und sich vermehrt der Zufriedenheit ihrer Patienten versichert haben. Dies könnten Patienten als besondere Zuwendung des Arztes empfunden haben.

#### **4.2.7 Einschlusskriterien**

Wir haben die Einschlusskriterien für unsere statistische Analyse so definiert, dass nur Probanden mit einem PHQ-9-Gesamtwert von mindestens 9 Punkten in die Studienpopulation einbezogen wurden. Obwohl die Autoren des PHQ-9 einen Grenzwert von 10 Punkten für das Screening vorschlagen (Kroenke und Spitzer 2002), haben wir diesen Wert analog zur PRoMPT-Studie auf 9 Punkte festgelegt. Da die PRoMPT-Studie speziell Patienten mit Depressionen untersucht hat, wurden nur Probanden ausgeschlossen, bei denen die Daten gänzlich fehlten. Die PARADIES-Studie hingegen untersuchte Patienten mit Panikstörung mit/ohne Agoraphobie. Dies führte dazu, dass knapp 37 % der Probanden bereits aufgrund eines PHQ-9-Gesamtwerts kleiner als 9 ausgeschlossen wurden.

Wie von den Autoren des PHQ-9-Fragebogens angegeben, können Fragebögen ausgewertet werden, bei denen maximal 20 % der Items fehlen (Kroenke et al. 2010). Da 20 % von 9 Items rein rechnerisch 1,8 Items ergeben, haben wir entschieden, dass maximal 2 von 9 Items fehlen dürfen. Um zusätzlich

auch den Verlauf der depressiven Symptomatik untersuchen zu können, wurde dieses Einschlusskriterien für die Messpunkte T0, T1 und T2 angewendet.

In einer vorläufigen Analyse wurden die Interventions- und Kontrollgruppen sowohl der PARADIES- als auch der PRoMPT-Studie analysiert, um zu zeigen, dass die Studien hinsichtlich des PACIC-Wertes miteinander verglichen werden können. Durch dieses Vorgehen konnten wir die Unterschiede der Versorgungsqualität der Kontroll- und Interventionsgruppen beider Studien gegenüberstellen. In beiden Studien zeigten die Interventions- gegenüber den Kontrollgruppen signifikant höhere PACIC-Werte. Während sich zwischen den Interventionsgruppen beider Studien keine signifikanten Unterschiede hinsichtlich der PACIC-Werte zeigten, waren die PACIC-Werte der Kontrollgruppe der PRoMPT-Studie signifikant höher als die PACIC-Werte der Kontrollgruppe der PARADIES-Studie.

Dass die Probanden der Interventionsgruppen beider Studien signifikant höhere PACIC-Werte zeigten, ist wahrscheinlich durch das jeweilige Studiendesign begründet. In beiden Studien erhielten die Probanden der Interventionsgruppe eine intensivere Betreuung im Sinne häufigerer Kontakte mit ihrem Praxisteam, und es wurde explizit die psychiatrische Symptomatik abgefragt. Dies könnte die höhere Patientenzufriedenheit anhand der PACIC-Werte der Interventionsgruppen erklären.

Im Vergleich der PACIC-Werte beider Interventionsgruppen zeigten sich keine statistisch signifikanten Unterschiede. Dies könnte dadurch begründet sein, dass der Grad der Betreuung sehr ähnlich war. Die beiden Studien ähnelten sich auch hinsichtlich des telefonischen Monitorings der psychiatrischen Symptomatik, sodass die Unterschiede zwischen der Behandlung der Interventionsgruppen nur gering waren.

Dass sich die Probanden beider Kontrollgruppen hinsichtlich ihrer PACIC-Werte statistisch signifikant voneinander unterschieden, könnte beispielsweise an den untersuchten Erkrankungen liegen. Mit den Kontrollprobanden der PRoMPT-Studie gaben jene Probanden signifikant höhere PACIC-Werte an, die – bedingt durch das Studiendesign – an Depressionen litten. Dies könnte darin begründet sein, dass eine unipolare Depression mit einer 12-Monats-Prävalenz von 8,2 % - verglichen mit 2 % einer Panikstörung mit/ohne Agoraphobie - eine höhere Prävalenz aufweist und Hausärzte in einem höheren Maße für die Bedürfnisse depressiver Patienten sensibilisiert sind, als sie es bei Patienten mit Panikstörung sind.

#### **4.2.8 Statistik und Aufbereitung der Daten**

Zunächst sondierten wir die verfügbaren statistischen Daten beider Studien. Die Aufbereitung der statistischen Daten stellte eine Herausforderung dar, da die Studiendaten der PRoMPT-Studie auf 5 einzelne Dateien aufgeteilt waren. Zunächst wurde überprüft, welche Variablen in den einzelnen Datensätzen vorlagen und anhand welcher Patienten- oder Praxis-Identifikationsnummern die einzelnen Dateien fehlerfrei zu einer Gesamtdatei zusammengefügt werden konnten. Dies war besonders für die Daten der PRoMPT-Studie wichtig, um aus den 5 Dateien eine Gesamtdatei zu erstellen, in welcher die Daten der jeweiligen Messpunkte korrekterweise den jeweiligen Probanden bzw. Arztpraxen zugeordnet sind.

Bei der Arbeit mit SPSS ist es hierbei notwendig, dass die Variablen „Name“, „Typ“, „Breite“, „Dezimalstellen“, „Beschriftung“, „Werte“ und „Messniveau“ (Skalenmaße) identisch sind. Bei Unterschieden in

diesen Punkten könnte SPSS die zwei Variablen nicht in eine Variable zusammenführen, sodass eine statistische Analyse nicht möglich wäre. Nur bei korrektem Vorgehen können beispielsweise PHQ-9-Werte vom Messpunkt T0 sowohl der PARADIES- als auch der PRoMPT-Studie miteinander in einer übereinstimmenden Variable zusammengefügt und später gemeinsam analysiert werden.

Eine weitere Herausforderung waren die uneinheitlichen PACIC-Versionen. Während die PARADIES-Studie den Fragebogen mit 11 Items verwendet hat, wurde bei der PRoMPT-Studie der Fragebogen in seiner Originalversion mit 20 Items eingesetzt. Das Item 1 des PACIC-Fragebogens mit 20 Items entspricht beispielsweise nicht dem Item 1 der PACIC-Version mit 11 Items. Die Variablen der identischen Items beider Fragebogen-Versionen wurden speziell benannt, sodass beim Zusammenfügen der Daten eine einzige Variable entstand. Um bei dem genannten Beispiel zu bleiben, waren Item 2 der PACIC-Originalversion und Item 1 der PACIC-Kurzversion identisch.

Mithilfe von Hilfsvariablen wurde die Variable *PACIC: Prozentwert vom Maximum* für die statistische Analyse erstellt. Diese Hilfsvariablen haben beispielsweise beim Umrechnen der Skalenwerte, bezogen auf ihre ursprüngliche Skala, geholfen. Während in der PACIC-Version mit 20 Items maximal 5 Punkte pro Item möglich sind, sind dies bei der PACIC-Version mit 11 Items 100 Punkte. Damit die Ergebnisse fehlerfrei blieben, wurde mit der Hilfsvariable „Maximaler Punktwert pro Item“ die Relation wiederhergestellt. Somit wurde gewährleistet, dass 5 von maximal 5 Punkten auf der Likert-Skala ebenso der vollen Punktzahl entsprechen wie 100 % von maximal 100 % auf der Prozent-Skala.

Eine weitere Herausforderung stellten die Identifikationsnummern der Arztpraxen dar. Während die Formatierung der Variable zur Identifikation der einzelnen Probanden zwischen beiden Studien unterschiedlich war, wurden die Praxen fortlaufend nummeriert. Um in der Gesamtdatei nicht die Probanden der Praxis 2 aus der PARADIES-Studie mit den Probanden der Praxis 2 aus der PRoMPT-Studie in nur eine Praxis zusammenzuführen, wurde unter Einbeziehung des Studienarmes eine neue Formatierung gewählt.

### 4.3 Limitationen der Arbeit

Auch bei dieser Forschungsarbeit liegen Beschränkungen der Forschung vor, welche in Stichpunkten aufgeführt werden:

- Zwar wurden nur Patienten mit depressiver Symptomatik in die Studienpopulation eingeschlossen, jedoch blieben (psychiatrische) Nebenerkrankungen bei der statistischen Auswertung außen vor. Obwohl Modelle gerechnet wurden, die die unterschiedliche depressive Symptomatik berücksichtigten, wurde die depressive Symptomatik vorrangig als Einschlusskriterium und nicht als primäre Variable behandelt. Die primäre Variable war die Zufriedenheit mit der Behandlung aus Patientensicht.
- Obwohl sich die PARADIES- und PRoMPT-Studie hinsichtlich ihres Studiendesigns durch das Case Management der Interventionsgruppen ähnelten, unterschieden sich die Behandlungsprogramme der Interventionsgruppen. Während für die Probanden der PARADIES-Studie 4 Arzttermine innerhalb von 23 Wochen vorgesehen waren, erhielten die Probanden der



PRoMPT-Studie eine Leitlinien-gerechte Therapie. Ähnlich war, dass die Probanden beider Interventionsgruppen mithilfe eines telefonischen Symptom-Monitoring kontaktiert wurden.

- Die Praxen der Studien unterschieden sich zum Teil erheblich hinsichtlich der Art der Praxis (Einzel- oder Gemeinschaftspraxis) sowie der Lage der Praxis (städtische oder ländliche Region). Von der PARADIES-Studie wurden knapp 50 % der Probanden in Einzelpraxen behandelt und nur circa 25 % der Probanden in städtischen Praxen, während von der PRoMPT-Studie circa 70 % der Probanden in Einzelpraxen behandelt wurden und knapp 90 % in städtischen Praxen.
- Obwohl für die statistische Auswertung nur die 11 identischen Items der PACIC-Versionen ausgewertet wurden, so unterscheiden sich maßgeblich die Umstände der Beantwortung der Fragebögen. Die Vergleichbarkeit erscheint eingeschränkt, da die Probanden der PRoMPT-Studie jeweils einen um 9 Items längeren Fragebogen ausfüllten. Hiermit verbundene Effekte, die sich auf die Qualität der Daten auswirken, sind möglich.
- Wenngleich wir mit der Variable *PACIC: Prozentwert vom Maximum* einen Ansatz zur Vergleichbarkeit unterschiedlicher PACIC-Versionen liefern, so bleiben Skalenunterschiede hinsichtlich der Anzahl der Skalenstufen sowie der Benennung der Extremwerte bestehen. Es ist gut möglich, dass diese Merkmale die Angaben und damit die Qualität der Studiendaten beeinflussten.

#### 4.4 Empfehlung für weiterführende Forschung

Aus dieser Forschungsarbeit gehen Empfehlungen für die zukünftige Forschung hervor, welche in Stichpunkten aufgeführt werden:

- Um zu überprüfen, ob auch andere Studien zu dem Ergebnis kommen, dass Patienten jüngerer Studienärzte höhere PACIC-Werte angeben, wäre eine Meta-Analyse dieser Fragestellung wünschenswert. Es ließe sich zudem überprüfen, ob die PACIC-Werte von Erkrankungen der Patienten abhängen.
- Bei der Erfassung der Zufriedenheit mit der Behandlung aus Patientensicht wären Variablen wünschenswert, die die Dauer der Behandlung in der jeweiligen Praxis, mögliche (psychiatrische) Nebendiagnosen der Patienten, Fallzahlen der Arztpraxen mit psychiatrischen Patienten sowie Vorerfahrungen der Studienärzte mit psychiatrischen Patienten erheben. Hierdurch wäre es noch besser möglich, individuelle Unterschiede dieser Patienten- und Praxismerkmale zu erfassen.
- Bei der Arbeit mit dem PACIC-Fragebogen wären definierte Grenzwerte fehlender Antworten für den Ausschluss von der statistischen Auswertung hilfreich. Wie im Kapitel 1.5.13 „Umgang mit fehlenden Werten des PACIC“ ausgeführt, herrscht Uneinigkeit über geeignete Grenzwerte, sodass die Studienpopulationen unterschiedlicher Studien nur schwer miteinander verglichen werden können.

## Affidavit



### Eidesstattliche Versicherung

Beltz, Christian

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Einfluss von Patienten- und Praxismerkmalen auf die Patientenzufriedenheit: Eine gepoolte Analyse der PARADIES- und PRoMPT-Studien.

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den 09.07.2024

Ort, Datum

Christian Beltz

Unterschrift Doktorandin bzw. Doktorand

## Danksagung

Ich möchte mich bei allen bedanken, die mich bei der Anfertigung dieser Dissertation unterstützt haben.

Danke für die Betreuung durch Prof. Dr. Jochen Gensichen, PD Dr. Karoline Lukaschek und Prof. Dr. Peter Falkai.

Danke für die fachliche Unterstützung durch Dr. Susanne Rospleszcz.

Danke für die persönliche Unterstützung durch meine Familie und Freunde.

Christian Beltz

## Publikationsliste

Lukaschek K, Beltz C, Rospleszcz S, Schillok H, Falkai P, Margraf J, Gensichen J. Depressive primary care patients' assessment of received collaborative care. *Sci Rep.* 2023 Feb 9;13(1):2329. doi: 10.1038/s41598-023-29339-9. PMID: 36759622; PMCID: PMC9911390.

## 5. Literaturverzeichnis

Abdul-Razak, S.; Ramli, A. S.; Badlishah-Sham, S. F.; Haniff, J. (2018): Validity and reliability of the patient assessment on chronic illness care (PACIC) questionnaire: the Malay version. In: *BMC Family Practice* 19 (1), S. 119. DOI: 10.1186/s12875-018-0807-5.

Alharbi, N.; Alsubki, N.; Alotabi, F.; Alotabi, M.; Alhrabi, N.; Lusgnian, S. de; Jonas, S. (2021): Translation into Arabic and validation of the Patient Assessment of Care for Chronic Conditions questionnaire for diabetes. In: *Eastern Mediterranean Health Journal* 27 (2), S. 142–150. DOI: 10.26719/emhj.20.136.

Ali, S.; Chessex, C.; Bassett-Gunter, R.; Grace, S. L. (2017): Patient satisfaction with cardiac rehabilitation: association with utilization, functional capacity, and heart-health behaviors. In: *Patient preference and adherence* 11, S. 821–830. DOI: 10.2147/PPA.S120464.

Aragones, A.; Schaefer, E. W.; Stevens, D.; Gourevitch, M. N.; Glasgow, R. E.; Shah, N. R. (2008): Validation of the Spanish translation of the Patient Assessment of Chronic Illness Care (PACIC) survey. In: *Preventing chronic disease* 5 (4), A113. Online verfügbar unter <https://pubmed.ncbi.nlm.nih.gov/18793501/>.

Arbeitsgemeinschaft für Neuropsychopharmakologie und Pharmakopsychiatrie e.V.; Arzneimittelkommission der deutschen Ärzteschaft; Arzneimittelkommission der Deutschen Apotheker; Bundesarbeitsgemeinschaft Selbsthilfe von Menschen mit Behinderung, chronischer Erkrankung und ihren Angehörigen e. V.; Bundespsychotherapeutenkammer; Bundesverband der Angehörigen psychisch erkrankter Menschen e. V. et al. (2022): Nationale VersorgungsLeitlinie Unipolare Depression - Langfassung: Bundesärztekammer (BÄK); Kassenärztliche Bundesvereinigung (KBV); Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF).

Arditi, C.; Iglesias, K.; Peytremann-Bridevaux, I. (2018): The use of the Patient Assessment of Chronic Illness Care (PACIC) instrument in diabetes care: a systematic review and meta-analysis. In: *International Journal for Quality in Health Care* 30 (10), S. 743–750. DOI: 10.1093/intqhc/mzy091.

Bandelow, B.; Aden, I.; Alpers, G. W.; Benecke, A.; Benecke, C.; Beutel, M. E. et al. (2021): Deutsche S3-Leitlinie Behandlung von Angststörungen, Version 2 (2021). Online verfügbar unter [www.awmf.org/leitlinien/detail/ll/051-028.html](http://www.awmf.org/leitlinien/detail/ll/051-028.html), zuletzt geprüft am 19.01.2023.

Batbaatar, E.; Dorjdagva, J.; Luvsannyam, A.; Savino, M. M.; Amenta, P. (2017): Determinants of patient satisfaction: a systematic review. In: *Perspectives in public health* 137 (2), S. 89–101. DOI: 10.1177/1757913916634136.

Bayerische Landesärztekammer (BLÄK): Psychosomatische Grundversorgung. Bayerische Landesärztekammer (BLÄK). Website. Online verfügbar unter <https://www.blaek.de/fortbildung/seminare-veranstaltungen-der-blaek/psychosomatische-grundversorgung>, zuletzt geprüft am 12.01.2023.

Blum, K.; Satzinger, W.; Buck, R. (2001): Patientenbefragungen und Qualitätsmanagement. Eine Einführung in die Thematik. In: W. Satzinger, A. Trojan und P. Kellermann-Mühlhoff (Hg.): *Patientenbefragungen in Krankenhäusern. Konzepte*. Sankt Augustin: Asgard-Verlag, S. 29.

Bo, A.; Jensen, N. H.; Bro, F.; Nicolaisen, S. K.; Maindal, H. T. (2020): Higher patient assessed quality of chronic care is associated with lower diabetes distress among adults with early-onset type 2 diabetes: Cross-sectional survey results from the Danish DD2-study. In: *Primary care diabetes* 14 (5), S. 522–528. DOI: 10.1016/j.pcd.2020.02.003.

- Bodenheimer, T.; Wagner, E. H.; Grumbach, K. (2002): Improving primary care for patients with chronic illness. In: *Journal of the American Medical Association* 288 (14), S. 1775–1779. DOI: 10.1001/jama.288.14.1775.
- Bonomi, A. E.; Glasgow, R.; Wagner, E. H.; Davis, C.; Sandhu, N. (2000): Assessment of Chronic Illness Care: How Well Does Your Organization Provide Care for Chronic Illness? Institute for Healthcare Improvement National Congress, 2000.
- Bonomi, A. E.; Wagner, E. H.; Glasgow, R. E.; Korff, M. von (2002): Assessment of chronic illness care (ACIC): a practical tool to measure quality improvement. In: *Health Services Research* 37 (3), S. 791–820. DOI: 10.1111/1475-6773.00049.
- Bundesministerium für Gesundheit (2016): Landärzte. Hg. v. Bundesministerium für Gesundheit. Online verfügbar unter <https://www.bundesgesundheitsministerium.de/service/begriffe-von-a-z/llandaerzte.html>, zuletzt aktualisiert am 18.01.2023.
- Chudoba; B. (2020): How much time are respondents willing to spend on your survey? Online verfügbar unter [https://www.surveymonkey.com/curiosity/survey\\_completion\\_times/](https://www.surveymonkey.com/curiosity/survey_completion_times/), zuletzt geprüft am 17.01.2023.
- Coleman, K.; Austin, B. T.; Brach, C.; Wagner, E. H. (2009): Evidence on the Chronic Care Model in the new millennium. In: *Health affairs (Project Hope)* 28 (1), S. 75–85. DOI: 10.1377/hlthaff.28.1.75.
- Cramm, J. M.; Nieboer, A. P. (2012): The chronic care model: congruency and predictors among patients with cardiovascular diseases and chronic obstructive pulmonary disease in the Netherlands. In: *BMC health services research* 12, S. 242. DOI: 10.1186/1472-6963-12-242.
- Cramm, J. M.; Nieboer, A. P. (2014): Development and Validation of the Older Patient Assessment of Chronic Illness Care (O-PACIC) Scale After Hospitalization. In: *Social Indicators Research* 116 (3), S. 959–969. DOI: 10.1007/s11205-013-0314-y.
- Dreyfus, H. L.; Dreyfus, S. E. (2016): What is Moral Maturity? In: Hubert L. Dreyfus und Mark A. Wrathall (Hg.): *Skillful coping. Essays on the phenomenology of everyday perception and action*. 1st publ. in pbk. Oxford: Oxford University Press, S. 183–201.
- Durand, M.-A.; DiMilia, P. R.; Song, J.; Yen, R. W.; Barr, P. J. (2018): Shared decision making embedded in the undergraduate medical curriculum: A scoping review. In: *PloS one* 13 (11), e0207012. DOI: 10.1371/journal.pone.0207012.
- Ferenchick, E. K.; Ramanuj, P.; Pincus, H. A. (2019): Depression in primary care: part 1-screening and diagnosis. In: *The BMJ* 365, l794. DOI: 10.1136/bmj.l794.
- First, M. B. (1997): *SCID-I. Clinician version*. Washington, D.C.: American Psychiatric Press.
- Frølich, A.; Nielsen, A.; Glümer, C.; Eriksen, C. U.; Maindal, H. T.; Kleist, B. H. et al. (2021): Patients' assessment of care for type 2 diabetes: Results of the Patient Assessment of Chronic Illness Care scale in a Danish population. In: *BMC health services research* 21 (1), S. 1069. DOI: 10.1186/s12913-021-07051-6.
- Gensichen, J.; Hiller, T. S.; Breitbart, J.; Teismann, T.; Brettschneider, C.; Schumacher, U. et al. (2014): Evaluation of a practice team-supported exposure training for patients with panic disorder with or without agoraphobia in primary care - study protocol of a cluster randomised controlled superiority trial. In: *Trials* 15, S. 112. DOI: 10.1186/1745-6215-15-112.
- Gensichen, J.; Korff, M. von; Peitz, M.; Muth, C.; Beyer, M.; Gütthlin, C. et al. (2009): Case management for depression by health care assistants in small primary care practices: a cluster randomized

trial. In: *Annals of internal medicine* 151 (6), S. 369–378. DOI: 10.7326/0003-4819-151-6-200909150-00001.

Gensichen, J.; Peitz, M.; Torge, M.; Mosig-Frey, J.; Wendt-Hermainski, H.; Rosemann, T. et al. (2006): Die "Depressions-Monitoring-Liste (De MoL)" mit integriertem PHQ-D-Rationale und Entwicklung eines Instruments für das hausärztliche Case Management bei Depression. In: *Zeitschrift für ärztliche Fortbildung und Qualitätssicherung* 100 (5), S. 375–382.

Gensichen, J.; Rosemann, T. (2007): Das Chronic Care Modell: Elemente einer umfassenden Behandlung für Patienten mit chronischen Krankheiten DEGAM-Serie: Betreuung von Patienten mit chronischen Krankheiten. In: *ZFA - Zeitschrift für Allgemeinmedizin* 83 (12), S. 483–486. DOI: 10.1055/s-2007-1004522.

Gensichen, J.; Serras, A.; Paulitsch, M. A.; Rosemann, T.; König, J.; Gerlach, F. M.; Petersen, J. J. (2011): The Patient Assessment of Chronic Illness Care questionnaire: evaluation in patients with mental disorders in primary care. In: *Community mental health journal* 47 (4), S. 447–453. DOI: 10.1007/s10597-010-9340-2.

Gensichen, J.; Torge, M.; Peitz, M.; Wendt-Hermainski, H.; Beyer, M.; Rosemann, T. et al. (2005): Case management for the treatment of patients with major depression in general practices--rationale, design and conduct of a cluster randomized controlled trial--PRoMPT (PRimary care Monitoring for depressive Patient's Trial) ISRCTN66386086--study protocol. In: *BMC public health* 5, S. 101. DOI: 10.1186/1471-2458-5-101.

Gilbody, S.; Richards, D.; Brealey, S.; Hewitt, C. (2007): Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. In: *Journal of general internal medicine* 22 (11), S. 1596–1602. DOI: 10.1007/s11606-007-0333-y.

Glasgow, R. E.; Wagner, E. H.; Schaefer, J.; Mahoney, L. D.; Reid, R. J.; Greene, S. M. (2005a): Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). In: *Medical care* 43 (5), S. 436–444. DOI: 10.1097/01.mlr.0000160375.47920.8c.

Glasgow, R. E.; Whitesides, H.; Nelson, C. C.; King, D. K. (2005b): Use of the Patient Assessment of Chronic Illness Care (PACIC) with diabetic patients: relationship to patient characteristics, receipt of care, and self-management. In: *Diabetes care* 28 (11), S. 2655–2661. DOI: 10.2337/dia-care.28.11.2655.

Goetz, K.; Freund, T.; Gensichen, J.; Miksch, A.; Szecsenyi, J.; Steinhäuser, J. (2012): Adaptation and psychometric properties of the PACIC short form. In: *The American journal of managed care* 18 (2), e55-60. Online verfügbar unter <https://pubmed.ncbi.nlm.nih.gov/22435885/>.

Goldstein, M. G.; Whitlock, E. P.; DePue, J. (2004): Multiple behavioral risk factor interventions in primary care. Summary of research evidence. In: *American journal of preventive medicine* 27 (2 Suppl), S. 61–79. DOI: 10.1016/j.amepre.2004.04.023.

Gräfe, K.; Zipfel, S.; Herzog, W.; Löwe, B. (2004): Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)". In: *Diagnostica* 50 (4), S. 171–181. DOI: 10.1026/0012-1924.50.4.171.

Greenleaf, E. A. (1992): Measuring Extreme Response Style. In: *Public Opinion Quarterly* 56 (3), S. 328. DOI: 10.1086/269326.

Gugiu, C.; Coryn, C. L. S.; Applegate, B. (2010): Structure and measurement properties of the Patient Assessment of Chronic Illness Care instrument. In: *Journal of evaluation in clinical practice* 16 (3), S. 509–516. DOI: 10.1111/j.1365-2753.2009.01151.x.

- Gugiu, P. C.; Coryn, C.; Clark, R.; Kuehn, A. (2009): Development and evaluation of the short version of the Patient Assessment of Chronic Illness Care instrument. In: *Chronic illness* 5 (4), S. 268–276. DOI: 10.1177/1742395309348072.
- Hall, J. A.; Dornan, M. C. (1988): What patients like about their medical care and how often they are asked: a meta-analysis of the satisfaction literature. In: *Social science & medicine* (1982) 27 (9), S. 935–939. DOI: 10.1016/0277-9536(88)90284-5.
- Hall, J. A.; Dornan, M. C. (1990): Patient sociodemographic characteristics as predictors of satisfaction with medical care: a meta-analysis. In: *Social science & medicine* (1982) 30 (7), S. 811–818. DOI: 10.1016/0277-9536(90)90205-7.
- Hiller, T. S.; Freytag, A.; Breitbart, J.; Teismann, T.; Schöne, E.; Blank, W. et al. (2018): Die Jena Angst-Monitoring-Liste (JAMoL) - ein Instrument zur evidenzbasierten Behandlung von Panikstörung mit oder ohne Agoraphobie in der Hausarztpraxis. In: *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 131-132, S. 28–37. DOI: 10.1016/j.zefq.2018.02.003.
- Hsieh, M. O.; Kagle, J. D. (1991): Understanding patient satisfaction and dissatisfaction with health care. In: *Health & social work* 16 (4), S. 281–290. DOI: 10.1093/hsw/16.4.281.
- Jacob, G.; Bengel, J. (2000): Das Konstrukt Patientenzufriedenheit: Eine kritische Bestandsaufnahme. In: *Zeitschrift für Klinische Psychologie und Psychotherapie* 48, 2000 (48), 280,301.
- Jöreskog, K. G. (1994): On the estimation of polychoric correlations and their asymptotic covariance matrix. In: *Psychometrika* 59 (3), S. 381–389. DOI: 10.1007/BF02296131.
- Kalton, G.; Roberts, J. S.; Holt, D. (1980): The Effects of Offering a Middle Response Option with Opinion Questions. In: *The Statistician*. Online verfügbar unter <https://www.semanticscholar.org/paper/The-Effects-of-Offering-a-Middle-Response-Option-Kalton-Roberts/635770433f11be0aedf2670ab7d9859e93b98c1d#related-papers>.
- Koley, M.; Saha, S.; Ghosh, S.; Nag, G.; Kundu, M.; Mondal, R. et al. (2016): Patient-Assessed Chronic Illness Care (PACIC) scenario in an Indian homeopathic hospital. In: *Journal of Traditional and Complementary Medicine* 6 (1), S. 72–77. DOI: 10.1016/j.jtcme.2014.11.020.
- Korff, M. von; Gruman, J.; Schaefer, J.; Curry, S. J.; Wagner, E. H. (1997): Collaborative management of chronic illness. In: *Annals of internal medicine* 127 (12), S. 1097–1102. DOI: 10.7326/0003-4819-127-12-199712150-00008.
- Kroenke, K.; Spitzer, R. L. (2002): The PHQ-9: A New Depression Diagnostic and Severity Measure. In: *Psychiatric Annals* 32 (9), S. 509–515. DOI: 10.3928/0048-5713-20020901-06.
- Kroenke, K.; Spitzer, R. L.; Williams, J. B. (2001): The PHQ-9: validity of a brief depression severity measure. In: *Journal of general internal medicine* 16 (9), S. 606–613. DOI: 10.1046/j.1525-1497.2001.016009606.x.
- Kroenke, K.; Spitzer, R. L.; Williams, J. B. W.; Löwe, B. (2010): The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. In: *General hospital psychiatry* 32 (4), S. 345–359. DOI: 10.1016/j.genhosppsych.2010.03.006.
- Krucien, N.; Le Vaillant, M.; Pelletier-Fleury, N. (2014): Adaptation and validation of the patient assessment of chronic illness care in the French context. In: *BMC health services research* 14, S. 269. DOI: 10.1186/1472-6963-14-269.



- Löwe, B.; Kroenke, K.; Herzog, W.; Gräfe, K. (2004a): Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). In: *Journal of affective disorders* 81 (1), S. 61–66. DOI: 10.1016/S0165-0327(03)00198-8.
- Löwe, B.; Spitzer, R. L.; Zipfel, S.; Herzog, W. (2002): PHQ-D: Gesundheitsfragebogen für Patienten. Online verfügbar unter [https://www.klinikum.uni-heidelberg.de/fileadmin/Psychosomatische\\_Klinik/download/PHQ\\_Manual1.pdf](https://www.klinikum.uni-heidelberg.de/fileadmin/Psychosomatische_Klinik/download/PHQ_Manual1.pdf), zuletzt geprüft am 09.01.2023.
- Löwe, B.; Unützer, J.; Callahan, C. M.; Perkins, A. J.; Kroenke, K. (2004b): Monitoring depression treatment outcomes with the patient health questionnaire-9. In: *Medical care* 42 (12), S. 1194–1201. DOI: 10.1097/00005650-200412000-00006.
- Ludt, S.; van Lieshout, J.; Campbell, S. M.; Rochon, J.; Ose, D.; Freund, T. et al. (2012): Identifying factors associated with experiences of coronary heart disease patients receiving structured chronic care and counselling in European primary care. In: *BMC health services research* 12 (1), S. 221. DOI: 10.1186/1472-6963-12-221.
- Mack, C.; Su, Z.; Westreich, D. (2018): Managing Missing Data in Patient Registries: Addendum to Registries for Evaluating Patient Outcomes: A User's Guide, Third Edition [Internet]. Agency for Healthcare Research and Quality (US), Rockville (MD). Online verfügbar unter <https://www.ncbi.nlm.nih.gov/books/NBK493614/?report=reader>, zuletzt aktualisiert am 17.01.2023.
- Maindal, H. T.; Sokolowski, I.; Vedsted, P. (2012): Adaptation, data quality and confirmatory factor analysis of the Danish version of the PACIC questionnaire. In: *European journal of public health* 22 (1), S. 31–36. DOI: 10.1093/eurpub/ckq188.
- Malliarou, M.; Bakola, E.; Nikolentzos, A.; Sarafis, P. (2020): Reliability and validity of the Greek translation of the patient assessment of chronic illness care + (PACIC-PLUS GR) survey. In: *BMC Family Practice* 21 (1), S. 122. DOI: 10.1186/s12875-020-01192-z.
- Neugebauer, B.; Porst, R. (2001): Patientenzufriedenheit: ein Literaturbericht 2001/07, S. 28. Online verfügbar unter [https://www.ssoar.info/ssoar/bitstream/document/48754/1/ssoar-2001-neugebauer\\_et\\_al-Patientenzufriedenheit\\_ein\\_Literaturbericht.pdf](https://www.ssoar.info/ssoar/bitstream/document/48754/1/ssoar-2001-neugebauer_et_al-Patientenzufriedenheit_ein_Literaturbericht.pdf).
- Peltzer, S.; Müller, H.; Köstler, U.; Schulz-Nieswandt, F.; Jessen, F.; Albus, C. (2020): Detection and treatment of mental disorders in patients with coronary heart disease (MenDis-CHD): A cross-sectional study. In: *PLoS one* 15 (12), e0243800. DOI: 10.1371/journal.pone.0243800.
- Petersen, J. J.; Paulitsch, M. A.; Mergenthal, K.; Gensichen, J.; Hansen, H.; Weyerer, S. et al. (2014): Implementation of chronic illness care in German primary care practices--how do multimorbid older patients view routine care? A cross-sectional study using multilevel hierarchical modeling. In: *BMC health services research* 14 (1), S. 336. DOI: 10.1186/1472-6963-14-336.
- Pilipovic-Broceta, N.; Vasiljevic, N.; Marinkovic, J.; Todorovic, N.; Jankovic, J.; Ostric, I. et al. (2018): Validation of the Patient Assessment of Chronic Illness Care (PACIC) Instrument in the Republic of Srpska, Bosnia and Herzegovina. In: *European Journal for Person Centered Healthcare* 6 (1), S. 12. DOI: 10.5750/ejpch.v6i1.1368.
- Rick, J.; Rowe, K.; Hann, M.; Sibbald, B.; Reeves, D.; Roland, M.; Bower, P. (2012): Psychometric properties of the Patient Assessment Of Chronic Illness Care measure: acceptability, reliability and validity in United Kingdom patients with long-term conditions. In: *BMC health services research* 12, S. 293. DOI: 10.1186/1472-6963-12-293.
- Rosemann, T.; Laux, G.; Drossemeyer, S.; Gensichen, J.; Szecsenyi, J. (2007): Evaluation of a culturally adapted German version of the Patient Assessment of Chronic Illness Care (PACIC 5A)

- questionnaire in a sample of osteoarthritis patients. In: *Journal of evaluation in clinical practice* 13 (5), S. 806–813. DOI: 10.1111/j.1365-2753.2007.00786.x.
- Rosemann, T.; Laux, G.; Szecsenyi, J.; Grol, R. (2008): The Chronic Care Model: congruency and predictors among primary care patients with osteoarthritis. In: *Quality & safety in health care* 17 (6), S. 442–446. DOI: 10.1136/qshc.2007.022822.
- Schwenke, M.; Welzel, F. D.; Luck-Sikorski, C.; Pabst, A.; Kersting, A.; Blüher, M. et al. (2019): Psychometric properties of the Patient Assessment of Chronic Illness Care measure (PACIC-5A) among patients with obesity. In: *BMC health services research* 19 (1), S. 61. DOI: 10.1186/s12913-019-3871-1.
- Sharma, H. (2022): How short or long should be a questionnaire for any research? Researchers dilemma in deciding the appropriate questionnaire length. In: *Saudi journal of anaesthesia* 16 (1), S. 65–68. DOI: 10.4103/sja.sja\_163\_21.
- Simonsen, N.; Koponen, A. M.; Suominen, S. (2018): Patients' assessment of chronic illness care: a validation study among patients with type 2 diabetes in Finland. In: *BMC health services research* 18 (1), S. 412. DOI: 10.1186/s12913-018-3206-7.
- Smith, P. B.; Vignoles, V. L.; Becker, M.; Owe, E.; Easterbrook, M. J.; Brown, R. et al. (2016): Individual and culture-level components of survey response styles: A multi-level analysis using cultural models of selfhood. In: *International journal of psychology* 51 (6), S. 453–463. DOI: 10.1002/ijop.12293.
- Spitzer, R. L.; Kroenke, K.; Williams, J. B. (1999): Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. In: *Journal of the American Medical Association* 282 (18), S. 1737–1744. DOI: 10.1001/jama.282.18.1737.
- Spitzer, R. L.; Williams, J. B.; Kroenke, K.; Linzer, M.; deGruy, F. V.; Hahn, S. R. et al. (1994): Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. In: *Journal of the American Medical Association* 272 (22), S. 1749–1756. Online verfügbar unter <https://pubmed.ncbi.nlm.nih.gov/7966923/>.
- Szecsenyi, J.; Goetz, K.; Campbell, S.; Broge, B.; Reuschenbach, B.; Wensing, M. (2011): Is the job satisfaction of primary care team members associated with patient satisfaction? In: *BMJ quality & safety* 20 (6), S. 508–514. DOI: 10.1136/bmjqs.2009.038166.
- Taggart, J.; Chan, B.; Jayasinghe, U. W.; Christl, B.; Proudfoot, J.; Crookes, P. et al. (2010): Patients Assessment of Chronic Illness Care (PACIC) in two Australian studies: structure and utility. In: *Journal of evaluation in clinical practice* 17 (2), S. 215–221. Online verfügbar unter [https://www.academia.edu/32975221/Patients\\_Assessment\\_of\\_Chronic\\_Illness\\_Care\\_PACIC\\_in\\_two\\_Australian\\_studies\\_structure\\_and\\_utility](https://www.academia.edu/32975221/Patients_Assessment_of_Chronic_Illness_Care_PACIC_in_two_Australian_studies_structure_and_utility).
- Teismann, T.; Lukaschek, K.; Hiller, T. S.; Breitbart, J.; Brettschneider, C.; Schumacher, U. et al. (2018): Suicidal ideation in primary care patients suffering from panic disorder with or without agoraphobia. In: *BMC psychiatry* 18 (1), S. 305. DOI: 10.1186/s12888-018-1894-5.
- Thom, D. H.; Hessler, D.; Willard-Grace, R.; DeVore, D.; Prado, C.; Bodenheimer, T.; Chen, E. H. (2015): Health coaching by medical assistants improves patients' chronic care experience. In: *The American journal of managed care* 21 (10), S. 685–691. Online verfügbar unter <https://pubmed.ncbi.nlm.nih.gov/26633093/>.

- Törő, V.; Kósa, Z.; Takács, P.; Széll, R.; Radó, S.; Szelesné, A. Á. et al. (2022): Patients' Assessment of Chronic Illness Care (PACIC): Validation and Evaluation of PACIC Scale among Patients with Type 2 Diabetes in Hungary. In: *International journal of integrated care* 22 (3), S. 5. DOI: 10.5334/ijic.6010.
- Tsai, A. C.; Morton, S. C.; Mangione, C. M.; Keeler, E. B. (2005): A meta-analysis of interventions to improve care for chronic illnesses. In: *The American journal of managed care* 11 (8), S. 478–488. Online verfügbar unter <https://pubmed.ncbi.nlm.nih.gov/16095434/>.
- Victoor, A.; Delnoij, D. M. J.; Friele, R. D.; Rademakers, J. J. D. J. M. (2012): Determinants of patient choice of healthcare providers: a scoping review. In: *BMC health services research* 12 (1), S. 272. DOI: 10.1186/1472-6963-12-272.
- Wagner, E. H. (1997): Managed care and chronic illness: health services research needs. In: *Health Services Research* 32 (5), S. 702–714.
- Wagner, E. H. (1998): Chronic disease management: what will it take to improve care for chronic illness? In: *Effective clinical practice* 1 (1), S. 2–4.
- Wagner, E. H. (2019): Organizing Care for Patients With Chronic Illness Revisited. In: *The Milbank quarterly* 97 (3), S. 659–664. DOI: 10.1111/1468-0009.12416.
- Wagner, E. H.; Austin, B. T.; Davis, C.; Hindmarsh, M.; Schaefer, J.; Bonomi, A. (2001): Improving chronic illness care: translating evidence into action. In: *Health affairs (Project Hope)* 20 (6), S. 64–78. DOI: 10.1377/hlthaff.20.6.64.
- Wagner, E. H.; Austin, B. T.; Korff, M. von (1996): Organizing care for patients with chronic illness. In: *The Milbank quarterly* 74 (4), S. 511–544.
- Wagner, E. H.; Davis, C.; Schaefer, J.; Korff, M. von; Austin, B. (1999): A survey of leading chronic disease management programs: are they consistent with the literature? In: *Managed care quarterly* 7 (3), S. 56–66.
- Wensing, M.; van Lieshout, J.; Jung, H. P.; Hermesen, J.; Rosemann, T. (2008): The Patients Assessment Chronic Illness Care (PACIC) questionnaire in The Netherlands: a validation study in rural general practice. In: *BMC health services research* 8 (1), S. 182. DOI: 10.1186/1472-6963-8-182.
- Williams, S. J.; Calnan, M. (1991): Convergence and divergence: assessing criteria of consumer satisfaction across general practice, dental and hospital care settings. In: *Social science & medicine (1982)* 33 (6), S. 707–716. DOI: 10.1016/0277-9536(91)90025-8.
- Zeugfang, D.; Wisetborisut, A.; Angkurawaranon, C.; Aramrattana, A.; Wensing, M.; Szecsenyi, J.; Krug, K. (2018): Translation and validation of the PACIC+ questionnaire: the Thai version. In: *BMC Family Practice* 19 (1), S. 123. DOI: 10.1186/s12875-018-0801-y.