

Inauguraldissertation

---

Zur Erlangung der Doktorwürde

im Fach Psychologie, Fachbereich I, Universität Trier

# Komorbidität in der ambulanten Psychotherapie.

Eine Untersuchung ihres Einflusses auf Status, Verlauf und Ergebnis im Rahmen  
eines Modellprojekts zum Qualitätsmonitoring.

Vorgelegt von Katharina Köck, 2012

Gutachter

---

Prof. Dr. Wolfgang Lutz

Prof. Dr. Henning Schöttke

Dissertationsort: Trier

Tag der mündlichen Prüfung: 16.05.2012

## Dank

Mein großer Dank gilt all jenen, die auf unterschiedliche Art und Weise zur Entstehung dieser Arbeit beigetragen haben:

Meinem Doktorvater Prof. Dr. Wolfgang Lutz für seine Betreuung während des Entstehungsprozesses dieser Arbeit und für die vielen Chancen, Herausforderungen und Lernerfahrungen der vergangenen Jahre.

Dr. Thomas Ruprecht von der Techniker Krankenkasse, der für die Projektkoordination des Modellvorhabens verantwortlich war und mit dem immer eine sehr gute und produktive Zusammenarbeit gelang. Weiterhin danke ich ausdrücklich allen teilnehmenden Patientinnen und Patienten, den Therapeutinnen und Therapeuten.

Dr. Niklaus Stulz und Jan Rasmus Böhnke für die gute Einarbeitung in Projekt und Methodik und die immer bereitwillige Beantwortung meiner Fragen. Daneben danke ich auch meinen weiteren lieben Kolleginnen und Kollegen in Trier für die schöne und in vielerlei Hinsicht gewinnbringende Zeit.

Meiner Familie, die immer da ist und an mich glaubt. Ganz besonders meinem Mann Vincent Schlatter, für die unverzichtbare emotionale Unterstützung, das Begleiten durch Höhen und Tiefen und die vielen geteilten Freuden.

# Inhaltsverzeichnis

1	EINLEITUNG .....	10
2	THEORETISCHER HINTERGRUND .....	14
2.1	Qualitätssicherung und naturalistische Forschung in der Psychotherapie	14
2.1.1	Qualitätssicherung .....	14
2.1.2	Naturalistische Psychotherapieforschung .....	17
2.2	Komorbidity .....	28
2.2.1	Einführung und Begriffsklärung .....	28
2.2.2	Erklärungsmodelle für Komorbidity.....	29
2.2.3	Debatte zum Konzept der Komorbidity .....	31
2.2.4	Wert des Komorbiditykonzeptes .....	33
2.2.5	Offene Fragen zum Bereich der Komorbidity .....	34
2.3	Zusammenhang zwischen der Art und Anzahl diagnostizierter Störungen und der Methode der Diagnosestellung .....	36
2.3.1	Befunde .....	36
2.3.2	Diagnosechecklisten - ein Kompromiss? .....	38
2.4	Befunde zur diagnostischen Komorbidity.....	40
2.4.1	Affektive Störungen .....	41
2.4.2	Angststörungen.....	41
2.4.3	Somatoforme Störungen .....	43
2.4.4	Persönlichkeitsstörungen.....	44
2.4.5	Belastungs- und Anpassungsstörungen .....	44
2.4.6	Essstörungen .....	45
2.4.7	Zusammenfassung .....	45
2.5	Befunde zu Komorbidity und Belastungsschwere zu Therapiebeginn .....	48
2.5.1	Allgemeine Symptombelastung / Körperliche und geistige Einschränkungen / gesundheitsbezogene Lebensqualität .....	48

2.5.2	Schwere der Hauptsymptomatik .....	49
2.5.3	Dauer der Hauptsymptomatik.....	50
2.5.4	Suizidalität .....	51
2.5.5	Psychosoziale Beeinträchtigung und allgemeine Leistungsfähigkeit .....	51
2.5.6	Lebensqualität, Hilfesuchverhalten, Behandlungsbereitschaft .....	52
2.5.7	Zusammenfassung .....	52
2.6	Einfluss von Komorbidität auf Therapieergebnis und –verlauf .....	54
2.6.1	Komorbidität und Therapieverläufe .....	54
2.6.2	Komorbidität und Therapieergebnis beziehungsweise -dauer .....	56
3	FRAGESTELLUNGEN UND HYPOTHESEN .....	62
3.1	Veränderung des Diagnoseverhaltens .....	62
3.2	Art und Muster der Komorbidität.....	62
3.3	Konsequenzen der Komorbidität für den Status zu Beginn der Therapie...	63
3.4	Konsequenzen der Komorbidität für Therapieverlauf und -ergebnis.....	63
4	METHODEN .....	67
4.1	Beschreibung der Stichprobe .....	67
4.1.1	Modellprojekt Psychotherapie der TK - Hintergrund und Ziele .....	67
4.1.2	Instrumente und Messplan .....	68
4.1.3	Teilnehmer – Therapeuten und Patienten im TK-Projekt .....	71
4.1.4	Teilstichprobe der vorliegenden Arbeit .....	73
4.2	Methoden zur Beantwortung der Fragestellungen (Status) .....	75
4.2.1	Zur Beantwortung der Hypothese 1 .....	75
4.2.2	Zur Beantwortung der Hypothesen 2 und 3 .....	79
4.2.3	Zur Beantwortung der Hypothese 4.....	82
4.3	Methoden zur Beantwortung der Fragestellungen (Verlauf und Ergebnis).....	83
4.3.1	Latent Growth Curve Models (LGM) .....	83
4.3.2	Growth Mixture Modeling (GMM) .....	87

4.3.3	Regressionsanalysen zur Vorhersage des Therapieergebnisses und der Anzahl Sitzungen unter Einbezug von Komorbidität und Klassenzugehörigkeit.....	90
5	ERGEBNISSE.....	91
5.1	Zu den Hypothesen 1 bis 4 .....	91
5.1.1	Ergebnisse zur Hypothese 1 .....	91
5.1.2	Ergebnisse zu den Hypothesen 2 und 3 .....	95
5.1.3	Ergebnisse zur Hypothese 4 .....	102
5.2	Zu den Hypothesen 5 bis 10 .....	106
5.2.1	Verlaufsanalysen - BSI als abhängige Variable.....	106
5.2.2	Regressionsanalysen erster Block: zur Vorhersage von Therapieergebnis und –dauer unter Berücksichtigung von Komorbidität und Klassenzugehörigkeit.....	111
5.2.3	Verlaufsanalysen – IIP als abhängige Variable.....	114
5.2.4	Regressionsanalysen zweiter Block: zur Vorhersage von Therapieergebnis und –dauer unter Berücksichtigung von Komorbidität.....	116
6	DISKUSSION .....	122
6.1	Reflexion der Ergebnisse.....	122
6.1.1	Zur Veränderung des Diagnoseverhaltens unter Anwendung von IDCL.....	122
6.1.2	Zum gemeinsamen Auftreten psychischer Störungen.....	124
6.1.3	Zur Identifikation latenter Klassen von Diagnosekombinationen.....	127
6.1.4	Zum Status komorbider Patienten zu Therapiebeginn .....	128
6.1.5	Zu den Verlaufsfragestellungen.....	130
6.1.6	Zur Vorhersage des Therapieergebnisses und der Anzahl Sitzungen.....	136
6.2	Kritische Betrachtung der Untersuchung .....	142
6.2.1	Probleme des Studiendesigns.....	142
6.2.2	Diagnosestellung und Zusammenfassung zu Diagnosegruppen .....	144
6.3	Allgemeine Diskussion und Ausblick.....	145
7	LITERATURVERZEICHNIS.....	150
8	ANHANG.....	170

## Abbildungsverzeichnis

Abbildung 1: Schematische Übersicht über die Messzeitpunkte (Selbstberichte) im TK-Modellvorhaben.....	70
Abbildung 2: 5-Klassen-Lösung der Latent Class Analysen .....	100
Abbildung 3: Latente Verlaufsklassen, abhängige Variable BSI.....	109

## Tabellenverzeichnis

Tabelle 1: Demografische Angaben sowie Mittelwerte der Eingangsmessungen in der Gesamtstichprobe und den zur Datenanalyse verwendeten Teilstichproben mit den zugehörigen Hypothesen.....	74
Tabelle 2: Einteilung der Störungsdiagnosen in 16 Kategorien plus „Andere Störungen“ .....	77
Tabelle 3: Therapiesitzung, in welcher die im Messplan vorgegebenen Eingangs- und Zwischenmessungen tatsächlich stattgefunden haben (IG, VT).....	85
Tabelle 4: Anzahl vergebener Diagnosen pro Patient .....	92
Tabelle 5: Häufigkeiten und relative Anteile der Haupt- und Nebendiagnosen zu Therapiebeginn .....	92
Tabelle 6: Ergebnisse der logistischen Regressionsanalyse zur Vorhersage von Komorbidität .....	93
Tabelle 7: Anzahl Patienten mit mindestens einer Nennung der Diagnose .....	94
Tabelle 8: Häufigkeiten und relative Anteile der bivariaten Störungskombinationen ...	96
Tabelle 9: Tetrachorische Korrelationskoeffizienten der bivariaten Diagnosekombinationen (Standardfehler) .....	97
Tabelle 10: Ergebnisse der Latent Class Analysen .....	98
Tabelle 11: Anteil der Latenten Klassen und bedingte Wahrscheinlichkeiten.....	99
Tabelle 12: Verteilung der Patienten in IG und KG auf die Latenten Diagnoseklassen .....	101
Tabelle 13: Deskriptive Statistik zur zweifaktoriellen Varianzanalyse mit den Faktoren Gruppenzugehörigkeit und Latente Diagnoseklasse.....	103

Tabelle 14: Deskriptive Statistik zur zweifaktoriellen Varianzanalyse mit den Faktoren Gruppenzugehörigkeit und Komorbiditätsmuster.....	104
Tabelle 15: Selbst- und fremdberichtete Beeinträchtigung für die Kombination aus rezidivierender depressiver Störung und Persönlichkeitsstörung.....	105
Tabelle 16: Ergebnisse der LGM-Analysen für die abhängige Variable BSI .....	107
Tabelle 17: Ergebnisse der Regression der Wachstumsfaktoren auf Komorbidität (BSI) .....	108
Tabelle 18: Anzahl (relativer Anteil) Patienten in den identifizierten Klassen, in Abhängigkeit des Komorbiditätsmusters (listenweiser Fallausschluss, N=321) . .	110
Tabelle 19: Regressionsanalysen zur Vorhersage von Therapieergebnis und –dauer unter Berücksichtigung von Komorbidität und Klassenzugehörigkeit .....	113
Tabelle 20: Ergebnisse der LGM-Analysen für die abhängige Variable IIP .....	114
Tabelle 21: Ergebnisse der Regression der Wachstumsfaktoren auf Komorbidität (IIP)	115
Tabelle 22: Regressionsanalysen (schrittweise) zur Vorhersage von Therapieergebnis und –dauer unter Berücksichtigung von Komorbidität .....	121

## Zusammenfassung

Bislang liegen nur sehr wenige naturalistisch angelegte Studien vor, die die Auswirkung von Komorbidität in der psychotherapeutischen Praxis umfassend untersuchen. Daneben bestehen Diskussionen dazu, ob Komorbidität ausschließlich als zwangsläufiges Produkt der bestehenden Diagnosesysteme aufzufassen ist oder ob darüber hinaus ein klinischer Nutzen des Konzeptes für die Planung von Psychotherapien und deren Durchführung besteht. Es stellt sich die Frage, ob und inwieweit Mehrfachdiagnosen im Vergleich zu einfachen Diagnosen Einfluss auf den Status zu Beginn einer Therapie, den Therapieverlauf und das Ergebnis ausüben und ob daraus Ableitungen für eine differenzielle Anpassung therapeutischer Interventionen getroffen werden können. Insofern stellt Komorbidität einen möglicherweise differenzierenden Faktor für das Ansprechen auf eine Behandlung dar, weshalb ihre Untersuchung auch im Rahmen einer patientenorientierten Forschung bedeutsam ist. Ein größeres Wissen um Einflussvariablen auf Therapieverläufe und -ergebnisse kann dazu beitragen, bestehende Vorhersage- und Rückmeldesysteme weiterzuentwickeln und eine Qualitätssicherung am Einzelfall umzusetzen. Die vorliegende Arbeit verbindet die Konzepte Komorbidität und naturalistische Forschung, indem hier Mehrfachdiagnosen in einer großen und längsschnittlich angelegten Studie zur Qualitätssicherung in der ambulanten Psychotherapie untersucht werden.

Im Spezifischen wird in dieser Arbeit geprüft, inwieweit die Anzahl gestellter Diagnosen sowie die Häufigkeit bestimmter Diagnosegruppen in Abhängigkeit der diagnostischen Herangehensweise variieren und ob sich spezifische Muster häufig gemeinsam vergebener Störungsdiagnosen identifizieren lassen. Daneben werden Zusammenhänge zwischen dem Vorliegen von Komorbidität und psychometrisch erhobener Belastung untersucht. Bezüglich des Therapieverlaufs wird geprüft, ob sich komorbide Patienten in ihren Verbesserungsraten in allgemeiner psychischer und interpersonaler Belastung von monomorbiden Patienten unterscheiden. Daneben wird getestet, ob sich im Verlauf nicht direkt beobachtbare, das heißt latente Gruppen von Patienten identifizieren lassen und wie diese mit Mehrfachdiagnosen assoziiert sind. Zuletzt wird untersucht, ob sich Komorbidität als Prädiktor

zur Vorhersage des Therapieergebnisses nutzen lässt und ob sich daraus Rückschlüsse auf die Dauer einer Therapie ziehen lassen.

Die in dieser Arbeit analysierten Daten stammen aus dem Modellprojekt „Qualitätsmonitoring in der ambulanten Psychotherapie“ der Techniker Krankenkasse, in welchem regelmäßige psychometrische Erhebungen und eine strukturierte Diagnostik mit Hilfe von Diagnosechecklisten umgesetzt sind. Sie umfassen Eingangsinformationen von 1154 verhaltenstherapeutisch behandelten ambulanten Psychotherapiepatienten. Zur Überprüfung der Fragestellungen kommen regressions- und korrelationsanalytische Verfahren, Latente Wachstumsmodelle zur Analyse von Längsschnittdaten (Latent Growth Curve Models, LGM), sowie statistische Verfahren zur Klassifikation von Personen in latente Subgruppen (Latent Class Analysen, LCA; Growth Mixture Modeling, GMM) zur Anwendung.

Es resultieren höhere Komorbiditätsraten unter strukturierter Diagnostik sowie ein überzufällig häufiges gemeinsames Auftreten von rezidivierender depressiver Störung und Persönlichkeitsstörungen. Patienten mit Mehrfachdiagnosen erweisen sich zu Therapiebeginn als deutlich belasteter als Patienten mit nur einer Störungsdiagnose, jedoch ergeben sich nur für eine Untergruppe komorbider Patienten Unterschiede hinsichtlich ihrer Verbesserungsrate im Therapieverlauf. Patienten mit mehreren Diagnosen profitieren in vergleichbarem Ausmaß wie solche mit nur einer Diagnose, weisen aufgrund der höheren Ausgangsbelastung jedoch ein schlechteres absolutes Therapieergebnis auf. Bezüglich der Assoziation von Komorbidität mit latenten Veränderungsmustern zeigt sich kein eindeutiges Ergebnis. Während zur Vorhersage des Therapieergebnisses Mehrfachdiagnosen nach Berücksichtigung der Eingangsbelastung keinen besonderen Stellenwert einnehmen, ist diese Variable bedeutsam für die Prädiktion der Sitzungsanzahl, indem komorbide Patienten und insbesondere solche mit Persönlichkeitsstörungen länger in der Therapie verbleiben.

Die sich auf mehreren Ebenen manifestierenden Besonderheiten komorbider im Vergleich zu monomorbiden Patienten weist darauf hin, dass das Konzept nicht ausschließlich als Artefakt bestehender Diagnosesysteme gesehen werden kann, sondern dass es von unmittelbarer klinischer Relevanz ist. Obwohl Psychotherapie auch unter Komorbidität große Effekte erzielt, muss mit teils komplizierteren und längeren Verläufen gerechnet werden. Wichtig sind hierbei aber zum einen die Differenzierung zwischen verschiedenen Untergruppen von Komorbidität und zum anderen eine gute diagnostische Grundlage, wobei Hinweise darauf bestehen, dass die Verwendung von Checklisten zu einer genaueren Di-

agnostik führt. Gerade bei Mehrfachdiagnosen erscheint eine individuelle Anpassung des therapeutischen Vorgehens an die spezifischen Bedürfnisse des Patienten wichtig und die Tatsache, dass komorbide Patienten länger in der Therapie verbleiben als solche mit nur einer Diagnose weist auf ein ebensolches differenzielles Vorgehen seitens der Therapeuten hin. Dieses könnte allerdings durch individualisierte Rückmeldungen noch unterstützt werden, in denen von vornherein Abschätzungen für spezifische Subgruppen von Patienten vorgenommen werden und in welchen Komorbidität als ein Indikator zu besserer Ressourcensteuerung in der Psychotherapie genutzt werden könnte.

# 1 EINLEITUNG

Das *Komorbiditätsprinzip*, nämlich einer untersuchten Person so viele Diagnosen zuzuweisen wie notwendig sind um ihre Problembereiche umfassend zu beschreiben, ist in den aktuellen Diagnosesystemen ICD-10 (Weltgesundheitsorganisation, 2005) und DSM-IV (American Psychiatric Association, 1994) explizit umgesetzt. Dementsprechend stellt die Beobachtung von Mehrfachdiagnosen unter psychischen Störungen heute eher die Regel als die Ausnahme dar (u. a. Clarkin & Kendall, 1992). Gegenstand einer regen Diskussion betrifft die Frage, ob Komorbidität damit letztlich nur als künstliches Resultat der standardisierten operationalisierten Diagnostik aufzufassen ist (u. a. Brieger & Marneros, 2000). Wenn dem so wäre, dann würde dieses Phänomen – wenn überhaupt – lediglich auf quantitative Unterschiede, im Sinne erfüllter Diagnosekriterien und gegebenenfalls auch einer größeren Ausprägung psychischer Belastung der betreffenden Personen, hinweisen. Demgegenüber hat das Konzept möglicherweise einen eigenständigen Wert, indem das Vorhandensein mehrerer Diagnosen für die Fallkonzeptionalisierung, Therapieplanung und -durchführung einen qualitativen Unterschied darstellt. Für diese Annahme spricht, dass die existierenden Behandlungsmanuale meist auf eine einzelne Störung bezogen sind, wodurch ihre Generalisierbarkeit unter Komorbidität stark eingeschränkt ist. Stattdessen müssen bei Mehrfachdiagnosen am Einzelfall orientierte Entscheidungen über den Behandlungsfokus, die sequenzielle Bearbeitung verschiedener Problembereiche und möglicherweise bestehende Interaktionen zwischen den Symptombereichen getroffen werden (Clarkin & Kendall, 1992).

Aus der Forschung zum Bereich psychischer Komorbidität bestehen Hinweise auf eine oft größere Gesamtbelastung im Vergleich zu Patienten mit nur einer Diagnose. Bezüglich weiterer Kriterien, wie das Ansprechen auf Behandlung, Dropout, Therapiedauer und –erfolg herrscht hingegen eine uneindeutige Befundlage. Studien, die sich mit der Untersuchung von Komorbidität in klinischen Stichproben befassen, greifen meist nur einen kleinen Bereich möglicher Diagnosen heraus. Demgegenüber stehen epidemiologische Untersuchungen, die mögliche Störungen sehr breit erfassen (Jacobi, Wittchen, et al., 2004; Kessler, Chiu, Demler, & Walters, 2004). Jedoch existieren nur sehr wenige naturalistisch angelegte Studien, die ein großes Spektrum an Diagnosen und Diagnosekombinationen

erfassen und aus denen letztlich Ableitungen darüber getroffen werden könnten, inwieweit Komorbidität in der psychotherapeutischen Praxis tatsächlich als zu beachtendes Merkmal oder Komplikation auch für den Verlauf angesehen werden muss. Gerade das Wissen um mögliche längerfristige Auswirkungen und Besonderheiten komorbider Patienten ist jedoch wichtig, um gegebenenfalls frühzeitig spezifische therapeutische Anpassungen auf der Grundlage diagnostischer Information vornehmen zu können.

Die naturalistische Forschung, die somit neben der Komorbidität einen weiteren Forschungsstrang darstellt in welchen diese Arbeit eingebettet ist, kann als Gegenpol zur experimentellen Psychotherapieforschung verstanden werden. Der größte Teil des heute bestehenden störungsspezifischen Wissens, auf dem empirisch fundierte Behandlungsmanuale beruhen, entstammt randomisierten kontrollierten Studien (randomized clinical trial; RCT<sup>1</sup>). In RCT-Studien findet unter anderem eine Vorselektion der Patienten statt, indem diese lediglich die Kriterien für die untersuchte Zieldiagnose erfüllen dürfen; Komorbidität wird explizit ausgeschlossen. Als wichtige Ergänzung dieses Ansatzes ist auf der anderen Seite die Überprüfung der Wirksamkeit psychotherapeutischer Behandlungsmethoden unter klinischen Routinebedingungen (effectiveness-Studien) und – für den Praktiker besonders wichtig – beim einzelnen Patienten (patientenorientierte Forschung; Howard, Moras, Brill, Martinovich, & Lutz, 1996) zu sehen. In der Praxis zeichnet sich die Qualität einer Psychotherapie nämlich neben ihrer Wirksamkeit unter experimentellen Gegebenheiten auch durch ihre Passung und Erfolgsaussichten im klinischen Alltag und ihre Möglichkeiten zu einer flexiblen Anpassung an die spezifischen Bedürfnisse des Patienten aus (sog. practice-based evidence). Mit letzteren Aspekten befasst sich die genannte *naturalistische Psychotherapieforschung*. Hier spielen Qualitätssicherungssysteme mit Beobachtungen des therapeutischen Fortschritts und dessen Rückmeldung in die Routine eine große Rolle (etwa Lambert, 2007). Daraus resultieren große Datenmengen über Psychotherapien unter realen Bedingungen, die durch eine in vielen Aspekten größere Heterogenität gekennzeichnet sind als Therapien unter kontrollierten Bedingungen. In Studien, welche in einem naturalistischen Kontext angesiedelt sind, werden Patienten mit mehreren Diagnosen ausdrücklich mit berücksichtigt.

---

<sup>1</sup> Üblicherweise wird im Rahmen von RCT-Designs die Wirksamkeit einer Behandlung unter spezifischen experimentellen Bedingungen untersucht (*efficacy-Studien*). Dazu wird ein therapeutisches Verfahren mit einer bereits bewährten Alternative oder einer Kontrollgruppe verglichen. Die Patienten werden einer der Gruppen zufällig zugeordnet. Die interne Validität sollte hoch ausfallen, während die externe Validität eingeschränkt ist.

Diese Arbeit setzt sich nun zum Ziel, psychische Komorbidität in einer groß angelegten naturalistischen Studie zu untersuchen. Die Integration der Bereiche Komorbidität und naturalistische Forschung beziehungsweise patientenorientierte Forschung stellt einen vergleichsweise neuen Ansatz dar. Bei näherer Betrachtung dieser beiden Forschungsfelder wird allerdings offensichtlich, dass eine Zusammenführung für beide Seiten Gewinn bringend sein kann. So bieten naturalistisch gewonnene Datensätze die Möglichkeit, das Auftreten und die Auswirkungen von Mehrfachdiagnosen auf einer breiten Datenbasis und auch über den Therapieverlauf hinweg zu untersuchen. Auf der anderen Seite stellt Komorbidität einen möglicherweise differenzierenden Faktor dar, der für die Verbesserung von Vorhersage- und Rückmeldemodellen im Rahmen einer patientenorientierten Forschung genutzt werden kann.

Konkret sollen im Hinblick auf die beiden Forschungsstränge folgende Fragen beantwortet werden:

- Welche Störungen treten gehäuft gemeinsam auf?
- Inwieweit steht die Beobachtung von Komorbidität in Zusammenhang mit der gewählten diagnostischen Methode?
- Wie beeinträchtigend sind Mehrfachdiagnosen im Vergleich zu nur einer Diagnose?
- Wie wirkt sich Komorbidität auf Therapieverläufe und –ergebnisse aus?

Das Modellvorhaben „Qualitätsmonitoring in der ambulanten Psychotherapie“ der Techniker Krankenkasse (Wittmann et al., 2011), welches im Sommer 2011 abgeschlossen wurde, ist eine naturalistische Studie, in der ambulante Psychotherapien der sogenannten „Richtlinienverfahren“ begleitet wurden. Als Maßnahmen zur Verbesserung der Effizienz und Effektivität ambulanter Therapien fanden in einer Interventionsgruppe regelmäßige psychometrische Erhebungen mit Rückmeldung an den Therapeuten bis zur Beendigung der Behandlung und zur 1-Jahres-Katamnese statt. Aufgrund dieses längsschnittlichen Designs und gleichzeitig der Festlegung breiter diagnostischer Einschlusskriterien mit expliziter Berücksichtigung von Mehrfachdiagnosen bietet sich im Rahmen dieser Studie die Möglichkeit, das Konzept der Komorbidität im psychotherapeutischen Alltag genauer zu untersuchen, und zwar auf den Ebenen Status zu Beginn der Therapie, Verlauf und Ergeb-

nis. Das damit verbundene Potenzial liegt in der Identifikation von Variablen, die das oft beobachtete unterschiedliche Ansprechen von Patienten auf Behandlung erklären könnten. Im Sinne einer wachsenden Patientenorientierung ist es wichtig, zu einer größeren Disaggregation von Daten zu gelangen, die es erlaubt, neben Aussagen für größere Patientengruppen im Mittel auch subgruppenspezifische Charakteristika zu identifizieren. Durch die Entwicklung eines besseren Verständnisses für solche Patientenunterschiede kann zu dem Ziel beigetragen werden, differenzielle Indikationskriterien zu entwickeln und damit letztlich die psychotherapeutische Versorgung im Einzelfall zu verbessern.

### *Aufbau der Arbeit*

Im ersten Teil dieser Arbeit erfolgt eine Übersicht zu den theoretischen Grundlagen. Dabei wird zum einen genauer auf die Themenbereiche Qualitätssicherung und naturalistische Psychotherapieforschung eingegangen (Kap. 2.1), zum anderen das Konzept der Komorbidität einschließlich seiner Erklärungsansätze und offenen Fragen eingeführt (Kap. 2.2). Im Anschluss werden Forschungsergebnisse zum Bereich komorbider Störungen dargestellt (Kap. 2.3 bis Kap. 2.6) und schließlich die Fragestellungen für die vorliegende Arbeit abgeleitet (Kap. 3). Im Kapitel 4 wird das methodische Vorgehen zur Beantwortung der Fragestellungen beschrieben und die Ergebnisse im nachfolgenden Kapitel 5 dargestellt. In der abschließenden Diskussion (Kap. 6) werden die Resultate in Bezug zu den Hypothesen und zum Forschungskontext gesetzt sowie Verbesserungsmöglichkeiten aufgezeigt.

## **2 THEORETISCHER HINTERGRUND**

In diesem Kapitel soll eine Einbettung in den Forschungshintergrund und Hinführung zu den Fragestellungen geschehen. Da die Arbeit im Kontext naturalistischer Psychotherapieforschung angesiedelt ist, wird dieser Bereich im ersten Abschnitt gemeinsam mit dem Begriff der Qualitätssicherung eingeführt. Darüber hinaus erfolgt eine Vorstellung wichtiger Konzepte, die im späteren Verlauf der Arbeit aufgegriffen werden. Im zweiten Abschnitt wird das Konzept Komorbidität umfassend betrachtet. Es werden spezifische Unter Aspekte zur Komorbidität sowie der entsprechende Stand der Forschung dargelegt, woraus schließlich die für diese Arbeit relevanten Fragestellungen und Hypothesen abgeleitet werden.

### **2.1 Qualitätssicherung und naturalistische Forschung in der Psychotherapie**

Neben einer Begriffsbestimmung werden zentrale Elemente des Modellprojekts der Techniker Krankenkasse (TK) zum Qualitätsmonitoring in der ambulanten Psychotherapie vorgestellt, aus welchem die dieser Arbeit zugrunde liegenden Daten stammen. Im Anschluss wird genauer auf den Bereich der naturalistischen Psychotherapieforschung eingegangen und grundlegende Entwicklungen der Verlaufs- und Ergebnisforschung dargelegt.

#### **2.1.1 Qualitätssicherung**

Maßnahmen zur Qualitätssicherung wurden in den vergangenen Jahren für alle Bereiche des Gesundheitssystems, einschließlich der Psychotherapie, thematisiert und umgesetzt und erhielten aufgrund ihrer gesetzlichen Verankerung im Rahmen der Gesundheitsreform des Jahres 2000 zusätzliche Bedeutsamkeit (u. a. Helou, Schwartz, & Ollenschläger, 2002). Die damit anvisierten Ziele sind auf mehreren Ebenen angesiedelt: zum einen wird eine Optimierung angebotener Behandlungen und ihrer Ergebnisse angestrebt, zum anderen soll angesichts des erhöhten Kostendrucks im Gesundheitswesen durch qualitätssichernde Maßnahmen eine verbesserte Wirtschaftlichkeit erreicht werden. Insbesondere im Bereich der Psychotherapie entstand ein wachsender Legitimationsdruck, die angebotenen Leistun-

gen, ihre Zweckmäßigkeit, Wirksamkeit und Effizienz zu dokumentieren. Darüber hinaus werden als Zielkriterien der Qualitätssicherung eine erhöhte Transparenz bezüglich ergriffener Maßnahmen sowie eine größere Patientenzufriedenheit angeführt (u. a. Kastenholz & Both, 2002).

Qualitätssichernde Maßnahmen beruhen somit einerseits auf politischem Druck zu Kosteneinsparungen, bieten jedoch auch vielfältige Anregungen und neue Möglichkeiten für die Psychotherapieforschung. Sie erleichtern die Implementierung von Vorgehensweisen, welche sich in der Forschung als nützlich erwiesen haben und deren Erprobung in der Praxis; gleichzeitig kann die Forschung profitieren durch eine Anpassung ihrer Forschungsfragen hinsichtlich ihrer klinischen Nützlichkeit und Praktikabilität. Die Erhebung von Patientendaten unter Praxisbedingungen, in der Regel an vielen Standorten, die enge Anbindung an Forschungseinrichtungen und die Rückspeisung der Daten und daraus gewonnenen Erkenntnisse in die klinische Routine stellen eine Möglichkeit dar, die oft beklagte sogenannte Wissenschaft-Praxis-Kluft zu reduzieren (u. a. Audin et al., 2001; Kazdin, 2008). Dementsprechend wird die Umsetzung qualitätssichernder Bausteine und Fortschrittsrückmeldungen in die Routineversorgung als ein wichtiger aktueller Trend der Psychotherapieforschung beschrieben (Lutz & Bittermann, 2010). Entgegen der Prominenz des Qualitätssicherungsbegriffes liegen bis zum heutigen Zeitpunkt keine übergreifenden Qualitätssicherungsprogramme für die ambulante Psychotherapie vor (Schulz, Barghaan, Harfst, & Koch, 2008) und es existieren nur wenige Untersuchungen, die im realen Versorgungskontext angesiedelt sind. So wurden beispielsweise – meist an Universitätsambulanzen und in stationären Einrichtungen – Qualitätssicherungsmethoden, insbesondere in Form von computergestützten Warn- und Erfolgssignalen, erprobt (u. a. Berking, Orth, & Lutz, 2006; Borkovec, Echemendia, Ragusea, & Ruiz, 2001; Kraft, Koellner, & Lutz, 2008; Lambert et al., 2003). Jedoch liegt bislang nur wenig Wissen dazu vor, ob und wie eine Umsetzung in Institutionen gelingen kann, welche meist eine weniger enge Verzahnung zu Forschungseinrichtungen aufweisen, nämlich bei niedergelassenen ärztlichen und psychologischen Psychotherapeuten, die einen Großteil der ambulanten psychotherapeutischen Versorgung in Deutschland leisten (Schulz et al., 2008). In diese Lücke versucht das Modellprojekt zum Qualitätsmonitoring in der Psychotherapie der Techniker Krankenkasse (Fydrich, Nagel, Lutz, & Richter, 2003; Lutz, Köck, & Böhnke, 2009; Wittmann et al., 2011) vorzustoßen, welches im Folgenden vorgestellt wird.

### **2.1.1.1 Umsetzung eines Qualitätsmonitoring-Systems in der ambulanten Psychotherapie – das TK-Modellvorhaben**

Für kassenfinanzierte Psychotherapien in Deutschland ist derzeit mit dem Gutachterverfahren ein Qualitätssicherungssystem umgesetzt, welches aufgrund von Vorwürfen mangelnder Objektivität, Reliabilität und Validität in der Kritik steht (Schulz et al., 2008). Mit dem Ziel der Erprobung eines alternativen Qualitätssicherungsverfahrens wurde durch die Techniker Krankenkasse (TK) ein Modellprojekt (im folgenden TK-Modell) angeregt und in den Jahren 2005 bis 2011 durchgeführt (u. a. Lutz, Köck, & Böhnke, 2009). Als zentrale Fragestellungen des Modellvorhabens wurden formuliert (Wittmann et al., 2011): (1) Führt der Einsatz von Qualitätsmonitoring-Instrumenten des TK-Modells in ambulanten Psychotherapien zu einer höheren Ergebnisqualität (Effektivität) im Vergleich zur traditionellen Richtlinienpsychotherapie? (2) Kann das TK-Modell die Wirtschaftlichkeit (Effizienz) der ambulanten Psychotherapie nachhaltig verbessern? Auf Design und verwendete Messinstrumente des TK-Modells wird in Kapitel 4.1 dieser Arbeit genauer eingegangen. Eine ausführliche Darstellung des Evaluationsplans, der Projektrealisation sowie der Ergebnisse zu Haupt- und Nebenfragestellungen ist dem Projekt-Abschlussbericht (Wittmann et al., 2011) zu entnehmen.

Von insgesamt 1708 ambulanten Psychotherapiepatienten wurden Eingangs-, Verlaufs- und Ergebnisinformationen gesammelt. Rund 400 Therapeuten aus drei Regionen der Kassenärztlichen Vereinigungen Hessen, Südbaden und Westfalen-Lippe nahmen auf freiwilliger Basis an der Studie teil und wurden in einem clusterrandomisierten Design auf Interventionsgruppe (IG) und Kontrollgruppe (KG) verteilt. In der IG wurden im Behandlungsverlauf mehrfach Selbstauskünfte der Patienten hinsichtlich unterschiedlicher Problembereiche erfasst. Diese wurden zeitnah an den Therapeuten zurückgemeldet, welcher die Information nutzen konnte um Anpassungen in der Behandlungsplanung und -durchführung vorzunehmen. In der KG erfolgten lediglich Messungen zu Beginn und Ende der Therapie sowie in beiden Gruppen zur 1-Jahres-Katamnese. Im Rahmen einer internen Qualitätssicherung sollte damit die Frage beantwortet werden, ob die wiederholte Anwendung anerkannter psychometrischer Testverfahren und deren Rückmeldung an die Therapeuten (therapiebegleitendes Monitoring) dazu beitragen kann, das therapeutische Handeln zu optimieren und so die Versorgung insgesamt zu verbessern. Neben den Rückmeldungen im Therapieverlauf wurden psychometrische Instrumente genutzt, um Entscheidungen über

die Bewilligung der Aufnahme beziehungsweise Fortführung einer Therapie in der IG zu unterstützen. Das Antrags-, Verlängerungs- und Gutachterverfahren wurde hier um die Sicht des Patienten bezüglich allgemeiner Belastung, störungsspezifischer Belastung und Therapiebeziehung ergänzt, um eine Optimierung der externen Qualitätssicherung zu erreichen. Daneben bestand eine Variation zwischen IG und KG in dem Vorgehen bei der kategorialen Diagnostik zu Therapiebeginn: während den Therapeuten der KG keine diesbezüglichen Vorgaben gemacht wurden, erfolgte in der IG eine strukturierte Eingangsdagnostik mit der Anwendung von Diagnosechecklisten (IDCL; Hiller, Zaudig & Mombour, 2004) als Kernstück. Dass strukturierte Verfahren zur Diagnosestellung zu valideren Diagnosen als ein unstrukturiertes Vorgehen führen, konnte an vielen Stellen gezeigt werden (u. a. Shear et al., 2000; s. a. Kapitel 2.3). Dieses Wissen ist insofern bedeutsam, als dass die Psychotherapie heute über einen großen Schatz an therapeutischem Wissen, störungsspezifischen Therapiemanualen und therapeutischen Leitlinien verfügt (u. a. Margraf, 2003), deren Potenzial jedoch nur ausgeschöpft werden kann, wenn ihrer Anwendung eine umfassende, saubere und verlässliche Diagnosestellung zu Grunde liegt. Somit muss auch die strukturierte Diagnostik als wichtiger Bestandteil qualitätssichernder Maßnahmen betrachtet werden.

Zusammengefasst sind im TK-Modellvorhaben zwei wesentliche Elemente – nämlich regelmäßige psychometrische Erhebungen und strukturierte Diagnostik – umgesetzt, die zwar zum unabdingbaren Standard der psychotherapeutischen Forschung gehören und deren Bedeutsamkeit in wissenschaftlichen Kreisen unbestritten ist, die jedoch in der Praxis bislang (noch) nicht konsequent umgesetzt werden (Miller, Dasher, Collins, Griffiths, & Brown, 2001). Insofern wird auch im Rahmen des TK-Modells ein Versuch unternommen, in die Lücke zwischen Forschung und Praxis vorzudringen und dem Defizit an empirischen Ergebnissen aus der Versorgungsforschung entgegenzuwirken, um so eine Basis für weiterführende qualitätssichernde Maßnahmen zu erstellen.

### **2.1.2 Naturalistische Psychotherapieforschung**

Grundlegend für die Auswahl, Implementierung und Erfolgskontrolle qualitätssichernder Maßnahmen sind Forschungsmethoden, die möglichst nahe an der psychotherapeutischen Realität angesiedelt sind. Auf wesentliche Überlegungen zu einer naturalistischen Psycho-

therapieforschung, die damit verfolgten Ziele und dahinterstehenden Konzepte soll in den folgenden Abschnitten genauer eingegangen werden.

### 2.1.2.1 Begriffsabgrenzung

In der Beurteilung des Nutzens und der Wirkung von Psychotherapie dienen klinisch-experimentelle Forschungsansätze der Beantwortung von Fragen zur generellen Wirksamkeit einer Behandlung im Vergleich zu Kontrollbedingungen (*efficacy*-Forschung; Howard et al., 1996). Aus verschiedenen Gründen sind die Rahmenbedingungen dieser Wirksamkeitsstudien jedoch nicht deckungsgleich mit denen der klinischen Praxis (u. a. Seligman, 1995), weshalb ergänzend eine sogenannte naturalistische empirische Forschung gefordert wird, die unter realen psychotherapeutischen Bedingungen stattfindet und deren Ergebnisse daher eine größere Praxisrelevanz als kontrollierte Studien haben.

Die Untersuchung der Effektivität einer Behandlung unter klinischen Alltagsbedingungen und damit auch bei einer größeren Heterogenität eingeschlossener Patienten- und Diagnosegruppen im Rahmen von quasi-experimentellen Designs wird als *effectiveness*-Forschung bezeichnet (u. a. Howard et al., 1996). Die Ansätze der *efficacy*- und *effectiveness*-Forschung ergänzen sich, indem in Ersterem auf eine Maximierung interner, in Letzterem auf eine Maximierung externer Validität hin gearbeitet wird – so sollten zur wissenschaftlichen Fundierung neuer Behandlungsmethoden immer sowohl sorgfältig geplante experimentelle als auch naturalistische Studien umgesetzt werden (Barkham & Mellor-Clark, 2003; Lutz, Rafaeli, Howard, & Martinovich, 2002). Wenn auch die Anwendung von *effectiveness*-Studien eine Annäherung an den klinischen Alltag bedeutet, so stellt die Übertragung gewonnener Erkenntnisse auf die konkrete Praxis weiterhin eine Schwierigkeit dar (Lambert, Hansen, & Finch, 2001). Dies liegt vor allem darin begründet, dass in beiden bisher angeführten Forschungsansätzen Aussagen auf Grundlage aggregierter Daten aus großen Patientstichproben getroffen werden (etwa über den Vergleich von mittleren Prä-Post-Veränderungen zwischen zwei Behandlungsbedingungen). Von einem praktisch-klinischen Standpunkt aus ist jedoch die Frage nach der Wirksamkeit einer Behandlung für einen bestimmten Patienten („Ist diese Therapie für diesen spezifischen Patienten wirksam?“) mindestens ebenso bedeutsam wie das Wissen um die mittlere Wirksamkeit einer Behandlung für eine Gruppe von Patienten. Der Behandlungsverlauf eines einzelnen Pati-

enten kann deutlich vom durchschnittlichen Trend abweichen. Hier setzt die patientenorientierte (Versorgungs- oder Psychotherapie-) Forschung (*patient-focused research*, Howard et al., 1996) als weiteres Element naturalistischer Forschung an, die den Blick auf die Ebene des Einzelfalls lenkt. In diesem Rahmen kommen Konzepte und Methoden zum Einsatz, die eine direkte Anwendung von Forschungsergebnissen in die Praxis erlauben und damit die zu Beginn benannte Wissenschaft-Praxis-Kluft zu überwinden versuchen. Etwa sollen durch die Anwendung empirisch begründeter Hilfen klinische Entscheidungen des Therapeuten im Therapieverlauf unterstützt werden. Auf Grundlage von Vergleichen zwischen erwarteten und tatsächlichen Therapieverläufen wird das Ansprechen eines Patienten auf die Behandlung überprüft und dem Behandler zeitnah zurück gemeldet, so dass individuelle Anpassungen im Behandlungsprozess vorgenommen werden können, um dem Ziel einer bestmöglichen Behandlung für jeden Patienten näher zu kommen (Howard et al., 1996; Lambert, Whipple, et al., 2001; Lutz, Martinovich, & Howard, 1999). Durch die Überwachung des Fortschritts jedes einzelnen Patienten lässt sich eine Verbindung zur Einzelfallforschung ziehen und das Vorgehen kann insofern als dem klinischen Alltag ähnlich angesehen werden, als dass fortlaufend spezifische Anpassungen therapeutischer Behandlungsstrategien vorgenommen werden.

Im Rahmen einer patientenorientierten Forschung sind somit die Vorhersage des erwarteten und die Rückmeldung des individuellen beobachteten Fortschritts von entscheidender Bedeutung. Um nah an den Entwicklungen des Einzelpatienten zu sein ist es notwendig, regelmäßige Erhebungen im Verlauf der Therapie durchzuführen (etwa Lambert, 2007; Lutz, Köck, & Böhnke, 2009; Lutz et al., 2005; Stiles et al., 2003). Dies hat den Vorteil, dass aus patientenorientierten Forschungsprojekten umfangreiche Datensätze zur Psychotherapie unter Routinebedingungen resultieren, die eine größere Disaggregation von Daten, das heißt die Analyse spezifischer Patientensubgruppen und Einzelpersonen sowie deren Veränderung im Verlauf der Therapie, erlauben als dies etwa bei Daten aus klassischen kontrollierten Designs (RCT) möglich wäre (Krause, Howard, & Lutz, 1998; Lutz, 2002). Sie bieten somit eine gute Möglichkeit, die Auswirkungen von Komorbidität im klinischen Alltag umfassend zu untersuchen.

### 2.1.2.2 Konzepte der Veränderung in der Psychotherapie

Basis für die Umsetzung von Vorhersage- und Rückmeldemodellen in der therapeutischen Versorgung sind zum einen methodische Verfahren zur Identifikation bedeutsamer Veränderungen, zum anderen Modelle und Konzepte, die der Beschreibung von Veränderungsprozessen in der Psychotherapie dienen. Im Folgenden werden die wichtigsten Konzepte zur Veränderung in der Psychotherapie vorgestellt. Dabei wird zunächst auf das Konzept klinisch signifikanter Veränderung sowie das Dosis-Wirkungs-Modell der Psychotherapie (Howard, Kopta, Krause, & Orlinsky, 1986) und das Phasenmodell der Veränderung (Howard, Lueger, Maling, & Martinovich, 1993) zur Beschreibung kontinuierlicher Therapieverläufe eingegangen. Im Anschluss werden konkrete Möglichkeiten einer Umsetzung von Rückmeldesystemen vorgestellt und Forschungsergebnisse zu diskontinuierlichen Verläufen darstellt.

#### *Klinisch signifikante Veränderung*

Wie können tatsächliche relevante Veränderungen von aufgrund mangelnder Reliabilität der Erhebungsinstrumente zufällig zustande gekommenen Schwankungen im Therapieverlauf unterschieden werden? Und wie statistische Signifikanz einer Veränderung von klinischer Relevanz? Eine in der Psychotherapieforschung häufig angewendete Methode zur Bestimmung der Bedeutsamkeit einer Veränderung ist das Konzept *klinisch signifikanter Veränderung* (Jacobson & Truax, 1991), welches zwei Komponenten beinhaltet. Um als zwischen zwei Erhebungen klinisch signifikant gebessert zu gelten, muss ein Patient aufgrund seiner Messwerte zum zweiten Zeitpunkt mit größerer Wahrscheinlichkeit einer funktionalen als einer dysfunktionalen Stichprobe zuzuordnen sein, das heißt den Cutoffpunkt, der diese beiden Stichproben trennt, unterschreiten. Darüber hinaus muss die erreichte Veränderung reliabel sein, also größer als allein aufgrund des Messfehlers zu erwarten und damit statistisch signifikant. Zwar werden einige Kritikpunkte an diesem Konzept genannt, etwa dass es sowohl für chronisch kranke Personen als auch für nur leicht belastete Patienten schwer ist, beide Kriterien zu erfüllen (etwa Hansen, Lambert, & Forman, 2002). Durch seine Anwendbarkeit sowohl auf große Stichproben als auch auf Veränderungen am Einzelfall ist das Konzept dennoch von unbestreitbarem Nutzen für Forschung und Praxis und findet breite Anwendung.

*Modelle zur Beschreibung kontinuierlicher Therapieverläufe*

Die *Dosis-Wirkungs-* (*dose-response-*) *Forschung* befasst sich mit dem Zusammenhang zwischen Therapieaufwand („Dosis“; entspricht in der Regel der Anzahl durchgeführter Therapiesitzungen) und Therapieergebnis („Wirkung“). Die Tatsache, dass in der Psychotherapieforschung immer wieder ein positiver Zusammenhang zwischen Behandlungslänge und –ergebnis gefunden wird (eine Übersicht dazu findet sich bei Orlinsky, Grawe, & Parks, 1994), verleitet implizit zu der Annahme eines linearen und additiven Zusammenhangs zwischen Aufwand und Wirkung (sog. Medikamentenmetapher). Als treffendere Heuristik zur Beschreibung und Untersuchung von Therapieverläufen wird jedoch immer wieder das *Dosis-Wirkungs-Modell* der Psychotherapie (Howard et al., 1986) herangezogen. In einer Metaanalyse aus 15 Studien haben Howard et al. (1986) an insgesamt 2400 Personen aus verschiedenen klinischen Populationen die Beziehung zwischen Verbesserung und unterschiedlich langer Behandlungsdauer untersucht. Sie kamen zu dem Ergebnis eines loglinearen, also positiven und dabei negativ beschleunigten Zusammenhangs zwischen der Anzahl durchgeführter Therapiesitzungen und der Veränderung in einem standardisierten Ergebnismaß. Das Dosis-Wirkungs-Modell sagt somit aus, dass zwar, wie durch die Medikamentenmetapher postuliert, eine Erhöhung der Dosis mit einer besseren Wirkung einhergeht, dass jedoch der Effekt von Psychotherapie in den ersten Sitzungen am größten ist und bei fortschreitender Therapie mehr und mehr Aufwand betrieben werden muss, um eine weitere Veränderung in die gewünschte Richtung zu erreichen (Howard u. a., 1996; Kopta, 2003). Das Modell hat nachfolgende Forschung angeregt, die vielfach Hinweise auf unterschiedliche Ansprechensraten auf eine Behandlung für verschiedene Diagnosen und Symptombereiche liefern konnte (s. u.).

Eine Erweiterung erfuhr das Dosis-Wirkungs-Modell durch das sog. *Phasenmodell* der Psychotherapie (Howard et al., 1993). Darin wird angenommen, dass sich der oben beschriebene loglineare Zusammenhang zwischen Aufwand und Wirkung aufgrund der Tatsache ergibt, dass im Verlauf einer Therapie die Zielbereiche der Veränderung variieren und da die Schwierigkeit der Zielerreichung dabei immer höher wird. Es werden drei Phasen der Veränderung vorgeschlagen, die in einer Therapie schrittweise durchlaufen werden, wobei eine Besserung in der vorangehenden Phase notwendig ist, um in die jeweils nächste Phase überzugehen. Die Phasen sind nach den jeweils dominierenden Zielbereichen benannt. (1) In der Phase der *Remoralisierung* geht es zunächst um die Schaffung

einer Hoffnung auf Besserung und den Aufbau einer positiven Erwartung im Hinblick auf die Behandlung, was mit einer Verbesserung des allgemeinen Wohlbefindens des Patienten einhergeht. (2) Die sich anschließende Phase der *Remediation* fokussiert auf eine Reduktion von Symptomen, indem der Patient vorhandene Kompetenzen nutzt und neue Bewältigungsmöglichkeiten aufgebaut werden. (3) Schließlich sollen in einer *Rehabilitationsphase* neue, funktionalere Verhaltensweisen eingeübt werden, die auch langfristig zu einer Verbesserung des allgemeinen Funktionierens im Alltag beitragen. Werden die drei Zielbereiche in einem globalen Ergebnismaß gemessen, so resultiert durch eine frühe Veränderung im Wohlbefinden eine zunächst sehr steile Kurve. Die größere Schwierigkeit der Erreichung von Zielen der Remediations- und Rehabilitationsphase bewirkt schließlich den abgeflachten Verlauf. An großen Patientenstichproben konnte das sequentielle Durchlaufen der Phasen (Howard et al., 1993) sowie das unterschiedlich schnelle Ansprechen verschiedener Problem- und Symptombereiche (u. a. Barkham, Rees, Stiles, Hardy, & Shapiro, 2002; Kopta, Howard, Lowry, & Beutler, 1994; Maling, Gurtman, & Howard, 1995) bestätigt werden.

Kritik am Dosis-Wirkungs-Modell bezieht sich darauf, dass es in seinen Ursprüngen (Howard et al., 1986) auf Studien beruht, die zum großen Teil lediglich Prä-Post-Erhebungen einbeziehen und nicht die Therapie in ihrem Verlauf betrachten. Jedoch konnte eine negativ beschleunigte Verbesserungskurve auch in sitzungsweise erhobenen Daten gefunden werden (u. a. Anderson & Lambert, 2001; Kadera, Lambert, & Andrews, 1996; Stulz, Lutz, Kopta, & Saunders, in Vorbereitung). Darüber hinaus betrachten Vertreter eines linearen Verbesserungsmodells die negativ beschleunigte Kurve als Artefakt, welches dadurch zustande komme, dass Therapien, zumindest in der klinischen Routine und wenn sie nicht zeitlich reglementiert werden, ihr natürliches Ende dann finden, wenn Patient und Therapeut ein Problem als ausreichend und gewinnbringend behandelt ansehen (sog. *good enough level* (GEL); Barkham et al., 2006, 1996). Statt eines loglinearen Verlaufs wird in dieser Konzeption eine stetige Verbesserung angenommen, deren Steigung abhängig ist von der Art und der Anzahl der Probleme, genauso wie von Therapeuten- und Patientenmerkmalen. Zu Beginn wenig belastete und schnell verbesserte Patienten scheiden nach dieser Annahme früher aus, was bei einer Mittelung über viele Therapien zu einem im Durchschnitt loglinearen Verlauf führe (Barkham et al., 2006; Shapiro et al., 2003). Bezüglich des Phasenmodells existieren Studien, die seine Gültigkeit nicht für alle Gruppen von

Patienten untermauern können (Joyce, Ogrodniczuk, Piper, & McCallum, 2002; Stulz & Lutz, 2007). Dennoch erfahren Dosis-Wirkungs- und Phasenmodell, unter anderem aufgrund ihrer theoretischen Schlüssigkeit und intuitiven Nachvollziehbarkeit, breite Zustimmung in der Welt der Psychotherapieforschung als allgemeines Veränderungsmodell, welches auch als Richtschnur für Behörden und Versicherungsunternehmen empfohlen wird (Lambert & Ogles, 2004).

#### *Der Zusammenhang zwischen Therapielänge und -Ergebnis*

Die Übertragung des Dosis-Wirkungs-Zusammenhangs auf die Psychotherapie lässt nicht nur Fragen nach der Form der Veränderung aufkommen, sondern auch danach, wie viel Therapie im Allgemeinen notwendig und ausreichend ist, damit Patienten geholfen ist. Mit einer Erhebung von Psychotherapiedaten auf Sitzungsebene, wie sie mit Aufkommen der Patientenorientierten Forschung gefordert wird (u. a. Kadera et al., 1996; Lambert, 2001), ist es möglich, bedeutsame Veränderungen direkt zu beobachten anstatt aus Prä-Post-Daten zu interpolieren (Hansen et al., 2002). In der Folge soll ein kurzer Überblick zum Stand der Forschung bezüglich der benötigten Therapiedauer bei verschiedenen Problembereichen gegeben werden. In der erwähnten Metaanalyse fanden Howard et al. (1986) Besserungsraten von 30, 53, beziehungsweise 74 Prozent nach 2, 8 beziehungsweise 26 Therapiesitzungen. Darüber hinaus konnten zum Teil stark unterschiedliche Besserungsraten verschiedener diagnostischer Gruppen (Howard et al., 1986; Pilkonis & Frank, 1988), der aus dem Phasenmodell abgeleiteten Veränderungsbereiche (Kopta et al., 1994) sowie interpersonaler Probleme (Maling, Gurtman, & Howard, 1995) identifiziert werden. Barkham et al. (1996) untersuchten die Symptomverbesserung in einer Gruppe depressiver Patienten bei variierender Sitzungsanzahl. Sie fanden Belege für eine negativ beschleunigte Kurve, indem die doppelte Sitzungsanzahl zwar mit einem besserem, nicht aber mit einem doppelt so guten Ergebnis einherging. Darüber hinaus variierte die Besserungsrate in Abhängigkeit vom Zielbereich (schnelleres Ansprechen in Maßen der Symptombelastung, langsames Ansprechen im Bereich interpersonaler Probleme). Ein noch genaueres Bild des Dosis-Wirkungs-Zusammenhangs konnte mit Hilfe von Überlebenskurven auf der Basis von Daten auf Sitzungsebene gewonnen werden (Anderson & Lambert, 2001; Hansen & Lambert, 2003; Kadera et al., 1996).

Diese Studien zusammenfassend kann, relativ unabhängig von Behandlungsmethode und Diagnose, davon ausgegangen werden, dass etwa zwischen 13 und 18 Therapiesitzungen benötigt werden, um bei 50% der Patienten eine klinisch bedeutsame Besserung zu erreichen (Hansen et al., 2002), wobei von einem Zusatznutzen bei mehr Sitzungen zur weiteren Symptomreduktion, insbesondere im interpersonalen Bereich, auszugehen ist. Vergleichbare Behandlungslängen bei noch etwas besserem Ergebnis wurden auch in einer Analyse von randomisierten klinischen Studien gefunden (Hansen et al., 2002). Auf Grundlage der geschilderten Befunde ist, bei allen Belegen für das Vorliegen eines Zusammenhangs zwischen Dosis und Wirkung in der Psychotherapie, jedoch davon auszugehen, dass sich nicht alle Patienten gleich schnell und gut verbessern. Vielmehr spricht der Stand der Forschung dafür, dass unterschiedliche Patienten unterschiedlich viele Sitzungen bis zu einer Besserung benötigen, und zwar in Abhängigkeit vom Schweregrad der Einschränkungen zu Beginn der Behandlung (u. a. Anderson & Lambert, 2001), aber auch von der gestellten Diagnose (Howard et al., 1986) und dem Vorliegen komorbider Störungen (etwa Joormann, Kosfelder, & Schulte, 2005; Pilkonis & Frank, 1988). So fassen auch Lambert & Ogles (2004, S. 155) zusammen: „Thus, rather than a general recovery curve [...], it might be best to rely on more specific curves based on variables of particular interest such as diagnoses“.

#### *Vorhersage- und Rückmeldesysteme als Umsetzung einer Patientenorientierten Forschung*

Wie schon angedeutet erklären das Dosis-Wirkungs- und das Phasenmodell Veränderungen in der Psychotherapie für den durchschnittlichen Patienten. Die Verläufe einzelner Personen können dabei beträchtlich vom allgemeinen Trend abweichen (etwa Krause et al., 1998). Um dieser Individualität Rechnung zu tragen, und damit eine wirklich patientenorientierte Forschung umzusetzen, wurden Methoden entwickelt, um Therapieverläufe einzelner Patienten vorherzusagen und in den Behandlungsprozess zurückzumelden. Dabei können unterschiedliche Entscheidungsregeln zur Bestimmung der erwarteten Veränderung eines Patienten zum Einsatz kommen, in die meist die Annahme eines loglinearen Genesungsverlaufes des Dosis-Wirkungs-Modells (Howard et al., 1986) einfließt. *Rationale Entscheidungsregeln* gründen sich auf Expertenurteilen, die festlegen wie sehr sich ein Patient verändern muss um in einer Therapie als erfolgreich zu gelten (u. a. Lutz, Stulz, Martinovich, Leon, & Saunders, 2009). Darin wird auch Wissen zu frühen Veränderungen

in der Therapie und zu klinisch signifikanten Veränderungen einbezogen (Lambert, 2007; Lambert et al., 2001). Rationale Entscheidungsregeln erwiesen sich vor allem als hilfreich zur frühzeitigen Identifikation negativer Therapieverläufe (u. a. Lambert, 2007). Sie wurden auch zur Generierung von Rückmeldungen im Rahmen des TK-Modellvorhabens genutzt. Demgegenüber beruhen *empirische Entscheidungsregeln* zum therapeutischen Fortschritt auf statistisch bestimmten erwarteten Genesungsverläufen, an denen der tatsächliche Fortschritt eines Patienten gemessen wird. Dazu werden aus bereits vorhandenen Therapieverlaufsdaten mittels hierarchisch linearer Modellierung (HLM; z. B. Raudenbush & Bryk, 2002) Verlaufsmodelle geschätzt, aufgrund derer dann Vorhersagen für neue Patienten vorgenommen werden können (Howard et al., 1996; Lutz et al., 2005, 1999, 2006). Man spricht hier auch von Kurven für den erwarteten Verlauf, oder *expected treatment response* (ETR)-Ansätzen (Lueger et al., 2001). Damit ist es möglich, auf Grundlage von Eingangsinformationen in verschiedenen Patientenmerkmalen und unter Annahme einer loglinearen Veränderungsfunktion individuelle Unterschiede im Behandlungsverlauf vorherzusagen (etwa Lutz et al., 1999). Solche Modelle wurden bei verschiedenen diagnostischen Gruppen (Lueger, Lutz, & Howard, 2000) und unter Heranziehung unterschiedlicher Ergebnismaße (Lutz, Lowry, Kopta, Einstein, & Howard, 2001; Lutz et al., 1999) erfolgreich angewendet und können noch verbessert werden, wenn Information über den Verlauf in frühen Sitzungen für die Vorhersage späterer Veränderungen genutzt wird (Haas, Hill, Lambert, & Morell, 2002; Lutz, Rafaeli, Howard, & Martinovich, 2002). Ob und inwieweit Information über das Vorliegen von Einfach- oder Mehrfachdiagnosen die Vorhersagegenauigkeit bezüglich des Therapieergebnisses einzelner Patienten steigern kann, ist bislang noch wenig untersucht.

Eine weitere Verfeinerung der ETR-Methode gelingt durch die sogenannte Technik der nächsten Nachbarn (*nearest neighbors*, NN): unter der Annahme, dass bestimmte Prädiktoren besonders für bestimmte Untergruppen gelten, werden hier aus einer Gruppe bereits behandelter Patienten die dem Zielpatienten ähnlichsten ausgewählt um für diesen spezifische Vorhersagen zu treffen (etwa Lutz, Leach, et al., 2005).

### *Modelle zur Beschreibung diskontinuierlicher Therapieverläufe*

Neben der Deskription kontinuierlicher Therapieverläufe beschäftigt sich ein weiterer Strang der Psychotherapieforschung mit der systematischen Untersuchung von Diskontinu-

itäten, das heißt wiederholten Schwankungen oder plötzlichen und einmaligen starken Veränderungen im Psychotherapieverlauf. Dabei wird häufig eine Gruppe von Patienten beschrieben, die sehr früh in der Therapie eine beträchtliche Besserung der Symptomatik aufweisen – sogenannte *rapid (early) responder* (u. a. Haas et al., 2002; Stulz & Lutz, 2007). Bislang existiert keine allgemeingültige Definition für frühe Veränderungen, weshalb auch zur Identifikation solcher Gruppen von Patienten sehr unterschiedliche Methoden angewendet werden: beispielsweise werden dazu die Differenzen zwischen beobachteten und erwarteten Veränderungswerten (ermittelt über Wachstumsanalysen) genutzt (Haas et al., 2002) oder aber es wird ein Abfall um 50% im Symptomaß innerhalb der ersten zwei Therapiesitzungen gefordert (Renaud et al., 1998). Neuere datenanalytische Methoden machen es darüber hinaus möglich, Subgruppen von Patienten auf Grundlage geteilter Veränderungsmuster in einer im Verlauf mehrfach erhobenen Ergebnisvariablen zu identifizieren. Über das sogenannte *Growth Mixture Modeling* (GMM; Muthén & Muthén, 2000) konnten in unterschiedlichen Stichproben nicht direkt beobachtbare Gruppen von Patienten, darunter auch „rapid (early) responder“ identifiziert werden (Lutz, Stulz, & Köck, 2009; Lutz, Stulz, Smart, & Lambert, 2007; Stulz & Lutz, 2007; Stulz, Lutz, Leach, Lucock, & Barkham, 2007). Des Weiteren konnten Zusammenhänge zwischen der Zugehörigkeit in frühen Veränderungsgruppen und dem Therapieergebnis (Lutz, Stulz, et al., 2009, 2007; Stulz & Lutz, 2007) beziehungsweise der Behandlungslänge (Stulz, Lutz, et al., 2007) hergestellt werden. Zum Zustandekommen früher Veränderungen in der Therapie werden unterschiedliche Erklärungsansätze diskutiert (u. a. Haas et al., 2002). Möglicherweise sind frühe Verbesserungen Ausdruck einer guten Passung zwischen Therapeut und Patient oder der Erfahrung des Therapeuten und damit seiner Fähigkeit, zum richtigen Zeitpunkt die richtigen Interventionen zu wählen. Oder aber sie weisen auf die hohe Motivation und Veränderungsbereitschaft eines Patienten hin, der in den ersten Sitzungen besonders von allgemeinen Faktoren und weniger von spezifischen Techniken profitiert. Stulz, Lutz, et al. (2007) konnten Depressivität und Ängstlichkeit zu Beginn einer Therapie als Prädiktoren für die Zugehörigkeit in bestimmten frühen Veränderungsgruppen ausmachen. Noch offen ist, inwieweit eine Belastung, die durch das gleichzeitige Vorhandensein mehrerer Störungsbilder charakterisiert ist, in Zusammenhang mit der Veränderung in frühen Therapiephasen steht.

Neben frühen Veränderungen werden in zahlreichen Arbeiten auch plötzliche Gewinne und Verluste beziehungsweise anhaltende Diskontinuitäten im Verlauf einer Behandlung untersucht. Unter anderem über die Methode des GMM konnte eine Gruppe von Patienten mit stark diskontinuierlichen Verläufen identifiziert werden, die auch eine erhöhte Wahrscheinlichkeit für reliable Verschlechterungen zum Behandlungsende aufweist (Stulz, Lutz, et al., 2007). Die Annahme, dass psychotherapeutische Veränderung nicht schleichend und schrittweise geschehen muss, sondern durch deutliche Sprünge gekennzeichnet sein kann, schlägt sich auch im Konzept der *sudden gains* (Tang & DeRubeis, 1999) nieder. Dieses bezeichnet eine substanzielle und beständige Verbesserung der Symptomatik zwischen zwei aufeinander folgenden Therapiesitzungen und steht in positiver Relation zum Therapieergebnis (Stiles et al., 2003; Tang & DeRubeis, 1999). Das Konzept fand eine Erweiterung durch die Einführung von *sudden losses*, also plötzlichen Verschlechterungen im Therapieverlauf (Lutz & Tschitsaz, 2007; Tschitsaz-Stucki & Lutz, 2009). Dabei zeigte sich, dass insbesondere das Auftreten beider Phänomene, also *sudden gains* und *sudden losses*, mit längeren Therapien und schlechteren Ergebnissen in Verbindung steht.

## 2.2 Komorbidität

Nachdem im vorangehenden Unterkapitel der für diese Arbeit gegebene Rahmen gesetzt und relevante Konzepte vorgestellt wurden, soll dieses Unterkapitel einen Überblick zur Komorbidität psychischer Störungen geben. Es geht zum einen auf Geschichte und Begrifflichkeiten, zum anderen aber auch auf Erklärungsmodelle bezüglich des Komorbiditätsprinzips bei psychischen Störungen ein. Darüber hinaus werden die Argumentationsstränge hinsichtlich des Für und Wider komorbider Diagnosen skizziert, der Wert des Komorbiditätsprinzips besprochen, sowie offene Fragen beleuchtet.

### 2.2.1 Einführung und Begriffsklärung

Der Begriff der *Komorbidität* als Bezeichnung für Patienten mit zwei oder mehr Diagnosen hielt in der empirischen Literatur durch Feinstein (1970) Einzug. Er definiert co-morbidity als „[...] any distinct additional clinical entity that has existed or that may occur during the clinical course of a patient who has the index disease under study“ (Feinstein, 1970, S. 456f). Dabei schließt er explizit nicht nur Krankheiten, sondern alle körperlichen Zustände ein, die Einfluss auf eine Erkrankung haben können. In seinem Artikel weist Feinstein bereits ausführlich auf die Bedeutung und mögliche Auswirkungen komorbider Zustände hin, etwa für Diagnostik, Behandlung und Prognose der Haupterkrankung. Ausgehend von dieser zunächst rein organmedizinischen Begriffsdefinition hat das Konzept der Komorbidität Einzug gehalten in die Bereiche der Psychiatrie und Klinischen Psychologie, wo es seit Ende der 1980er- beziehungsweise Anfang der 1990er-Jahre zunehmend an Präsenz gewonnen hat (Tholen, 2006). Die seither stetig anwachsende Zahl an Diskussionen und empirischen Veröffentlichungen zum Thema Komorbidität ist in engem Zusammenhang mit der Weiterentwicklung der Diagnosesysteme zu sehen. Explizite und weitgehend ätiologiefreie Operationalisierungskriterien für psychische Störungen wurden erstmals in der dritten Version des DSM festgelegt, in der später erschienenen Revision (DSM-III-R; American Psychiatric Association, 1987) wurden darüber hinaus bis dahin vielfach geltende Ausschlusskriterien weitestgehend abgeschafft (Krueger & Markon, 2006). Erst damit wurden multiple Diagnosen zugelassen und zu ihrer Vergabe ausdrücklich aufgefordert, was auch als „Paradigmenwechsel“ in der Diagnostik bezeichnet wird (Klerman, 1990). Heute ist das Komorbiditätsprinzip zentrales Merkmal der Klassifikationssysteme psychischer Stö-

rungen. Mit der Sichtbarmachung von Komorbidität wurden in Theorie und Forschung vermehrt auch die Zusammenhänge zwischen Störungen thematisiert. Entsprechend existieren zahlreiche Begriffsdefinitionen, von denen sich einige direkt auf psychische Störungen beziehen. So versteht etwa Klerman (1990) unter Komorbidität „the joint occurrence of two or more mental disorders occurring with each other and/or with medical conditions“.

In die vorangehende Definition gehen auch medizinische Krankheitsfaktoren ein. Dies verweist auf die Problematik der jeweiligen Beurteilungskriterien von Komorbidität, in deren Abhängigkeit berichtete Raten divergieren. Neben der Frage, ob medizinische und psychische Krankheitsfaktoren zusammen oder getrennt behandelt werden sollen spielt vor allem auch das betrachtete Zeitfenster eine Rolle: es kann zwischen aktueller (querschnittlicher) und Lebenszeit- (längsschnittlicher) Komorbidität unterschieden werden (u. a. Wittchen, 1996). Auf der Ebene der Untersuchungseinheiten müssen nicht nur die Art und Anzahl untersuchter Störungsgruppen bestimmt werden, sondern auch, ob lediglich Störungsdiagnosen nach DSM-IV (American Psychiatric Association, 1994) beziehungsweise ICD-10 (Weltgesundheitsorganisation, 2005) oder bereits einzelne Symptome Zielbereich der Betrachtung darstellen (Moggi, 2002). Darüber hinaus spielt für die Untersuchung von psychischer Komorbidität die Selektion von Versuchspersonen eine große Rolle, wobei in stationären Einrichtungen mit höheren Komorbiditätsraten als im ambulanten Setting und dort wiederum mit höheren Raten als in repräsentativen Bevölkerungsstichproben gerechnet werden muss (Moggi, 2002). In Anlehnung an die Definition von Burke, Wittchen, Regier, & Sartorius (1990; „Comorbidity can be defined as the presence of more than one specific disorder in a person in a defined period of time“) werden für die vorliegende Arbeit folgende Festlegungen getroffen: Komorbidität wird hier betrachtet als das Vorhandensein von (1) mehr als einer Diagnose (2) zu einem spezifischen Zeitpunkt (3) im Gültigkeitsbereich psychischer Störungen nach ICD-10 (Weltgesundheitsorganisation, 2005) (4) in einer Stichprobe ambulanter Psychotherapiepatienten.

### **2.2.2 Erklärungsmodelle für Komorbidität**

Hinsichtlich der Gründe für und Ursachen von Komorbidität gibt es unterschiedliche Erklärungsansätze. Bereits Kaplan & Feinstein (1974) unterscheiden vier Komorbiditätstypen, die gleichzeitig Hinweise auf die vielfältigen denkbaren Entstehungsfaktoren geben. (1) Von *diagnostischer Komorbidität* sprechen sie, wenn sich die Diagnosekriterien zweier

Störungen überschneiden. Das Argument, dass Komorbidität aufgrund von überlappenden Diagnosekriterien zustande kommen muss und folglich nicht immer Aussagen über inhaltliche Zusammenhänge erlaubt, sondern vielmehr als ein Artefakt betrachtet werden kann, wird auch in der aktuelleren Literatur immer wieder gefunden (u. a. Frances, Widiger, & Fyer, 1990; Klein & Riso, 1993; Krueger & Markon, 2006; s. u.). Darüber hinaus werden Probleme in der Grenzziehung zwischen Störungskategorien diskutiert: in diesen Argumentationsstrang fällt, dass Komorbidität als spezifische Unterform von Störungen aufgefasst werden könnte beziehungsweise eine distinkte Einheit bildet, die als unabhängige Störung neben den Einzeldiagnosen klassifiziert werden müsste (Klein & Riso, 1993; Krueger & Markon, 2006). (2) *Prognostische Komorbidität* liegt nach Kaplan & Feinstein (1974) dann vor, wenn das Vorhandensein einer Erkrankung zu weiteren Erkrankungen prädisponiert. Dieser Typ von Komorbidität wird auch im sogenannten „causation model“ (Krueger & Markon, 2006) beschrieben. (3) Dagegen benennt *pathogenetische Komorbidität* nach Feinstein (1970) das Vorhandensein eines gemeinsamen Vulnerabilitätsfaktors für zwei Erkrankungen. (4) *Therapeutische Komorbidität* nach Kaplan & Feinstein (1974) schließlich bezeichnet den möglichen Einfluss einer komorbiden Erkrankung auf den Behandlungserfolg. Zusätzlich wird Komorbidität als ein möglicher Effekt der Stichprobenziehung und Basisraten psychischer Störungen diskutiert und könnte damit rein zufällig beziehungsweise als Effekt von Stichprobenverzerrungen zustande kommen (Frances et al., 1990; Krueger & Markon, 2006; Rhee, Hewitt, Corley, Willcutt, & Pennington, 2005; s. u.). Auf inhaltlicher Ebene sind außerdem in Abhängigkeit vom jeweils zugrunde liegenden psychopathologischen Problemverständnis unterschiedliche Interpretationen zum Zustandekommen von mehr als einer Diagnose zu einem Zeitpunkt denkbar. So könnte etwa unter einer behavioralen Betrachtungsweise die zufällige gemeinsame Verstärkung verschiedener Verhaltensweisen ursächlich sein, während aus psychodynamischer Perspektive unspezifische Abwehrmechanismen wirksam sein könnten (Cloninger, Martin, Guze, & Clayton, 1990).

Zusammenfassend ist festzuhalten, dass eine Vielzahl von Erklärungsmodellen existieren, die zum Teil wahre Zusammenhänge zwischen Störungen annehmen, und zwar im Sinne einer geteilten Ätiologie oder spezifischer Prädispositionen. Demgegenüber stehen Modelle, die Komorbidität vorrangig als methodisches und klassifikatorisches Artefakt einordnen. Fest steht, dass es alleine auf der Grundlage deskriptiver Information unmöglich er-

scheint zu entscheiden, welche dieser vielfältigen Interpretationen die jeweils angemessene ist (Frances et al., 1990).

### **2.2.3 Debatte zum Konzept der Komorbidität**

Die in den vergangenen zwanzig Jahren stetig steigende Präsenz des Themas Komorbidität in der psychiatrischen und klinisch-psychologischen Literatur (Tholen, 2006) spiegelt nicht nur den möglicherweise gegebenen klinischen Nutzen dieses Konzeptes wider, sondern umso mehr noch seine kontroverse Natur. In Teilen wurde die Diskussion um Komorbidität im vorangehenden Abschnitt angedeutet; im Folgenden soll nun eine umfassende Darstellung dazu erfolgen.

Kritische Stimmen beziehen sich vor allem auf die Untrennbarkeit des Komorbiditätskonzeptes von den jeweils gültigen Klassifikationsgrundlagen. Insofern vermischen sich Beiträge zur Komorbiditätsdebatte mit solchen, die Kritik etwa an kategorialen Klassifikationsansätzen üben. Komorbidität wird an einigen Stellen der Wert als eigenständiges Phänomen abgesprochen, da sie lediglich ein künstliches Produkt der standardisierten operationalisierten Diagnostik sei (Brieger & Marneros, 2000). Tatsächlich sind die Raten von Mehrfachdiagnosen immer in Abhängigkeit bestehender diagnostischer Systeme zu sehen. Zum einen führte die genaue Operationalisierung von Diagnosekriterien und die Entwicklung von diagnostischen Interviews zu einer gestiegenen Augenscheinlichkeit bestimmter Diagnosen (Frances et al., 1990), zum anderen ist die Anzahl diagnostischer Kategorien für psychische Störungen mit den Jahren immer weiter angewachsen. Dies wird unter anderem begründet mit der oben erwähnten atheoretischen Natur aktueller Diagnosesysteme, aber auch mit der Tendenz zu immer größer definierten Geltungsbereichen psychischer Störungen (indem zum Beispiel Schlafstörungen einbezogen werden). Vor allen Dingen wird jedoch die Suche nach möglichst homogenen diagnostischen Kategorien beanstandet, die zu einer Aufsplittung von Diagnosen in immer enger gefasste, möglicherweise willkürlich gewählte Einheiten geführt habe, die dann zwangsläufig in gewissem Maße untereinander korrelieren müssten (Frances et al., 1990). Eine Ausnahme zu dieser Tendenz bilden wenige „Doppelkategorien“, wie Angst und depressive Störung gemischt (F41.2 im ICD-10).

Ein anderer Kritikpunkt bezieht sich auf die in aktuellen Diagnosesystemen weitreichend umgesetzten Schwellenwerte für die zur Diagnosestellung notwendige Anzahl an Kriterien.

Die Leichtigkeit der Überschreitung dieser Schwellenwerte beeinflusst die Diagnosehäufigkeit jeder einzelnen Störung und kann somit die Wahrscheinlichkeit für Mehrfachdiagnosen erhöhen (Klein & Riso, 1993). Darüber hinaus werden die meisten Diagnosekriterien als gleichwertig behandelt und eine Diagnose wird über die Aufsummierung gegebener Symptomgruppen gestellt. Dieses Vorgehen führt zu großer Heterogenität innerhalb einer Störungskategorie und könnte aussagekräftige Schlüsse zum regelhaften gemeinsamen Auftreten von Störungen verkomplizieren (Cooper & Balsis, 2009). Schließlich besteht eine nicht geringe Kriterienüberlappung zwischen Diagnosen, das heißt auch Kernmerkmale einer Störung müssen nicht für diese spezifisch sein (Maser & Cloninger, 1990). Dies kann zu erhöhten Komorbiditätsraten führen, wobei schwer zu trennen ist zwischen einer künstlichen, deskriptiven Ähnlichkeit und einer möglicherweise zugrunde liegenden ätiologisch oder biologisch begründeten Nähe (Frances et al., 1990). Jedoch wurde auch bei Nichtbeachtung überlappender Symptomgruppen zweier Diagnosen Komorbidität gefunden, was gegen das Argument von Komorbidität als einfaches Resultat sich überschneidender Störungsklassen spricht (Angold, Costello, & Erkanli, 1999). Als Schluss aus den genannten Argumenten kann festgehalten werden, dass Komorbidität immer vor dem Hintergrund der geltenden Klassifikationssysteme gesehen werden muss und damit Komorbiditätsstudien nicht ohne weiteres vergleichbar sind.

Weitere Beanstandung an der Validität des Komorbiditätskonzeptes bezieht sich auf Selektionseffekte und Verzerrungen aufgrund der Basisraten psychischer Störungen. Zum Problem der Basisraten ist festzuhalten, dass rein aufgrund von gegebenen Wahrscheinlichkeiten solche Merkmale, die häufig in einer Stichprobe auftauchen, auch häufig gemeinsam diagnostiziert werden. Damit kann berichtete Komorbidität auf wahren Ursachen beruhen oder aber auf dem Zufallswege zustande gekommen sein. Umgekehrt kann nicht davon ausgegangen werden, dass keine Verbindung zwischen zwei Störungen besteht, wenn eine der beiden nur selten oder gar nicht in einer Stichprobe auftaucht. Komorbidität zweier Störungsdiagnosen muss folglich immer in Bezug dazu gesetzt werden, was anhand der relativen Häufigkeit der beiden Störungen in einer Stichprobe zu erwarten wäre (etwa Krueger & Markon, 2006). Bezüglich des Einflusses von Selektionseffekten ist vor allem das Hilfesuchverhalten zu nennen. Personen mit mehr als einer Störung scheinen sich häufiger in Behandlung zu begeben (Galbaud du Fort, Newman, & Bland, 1993), was zu einer Überschätzung von Mehrfachdiagnosen bei alleiniger Betrachtung von klinischen Stich-

proben führen würde (Clark, Watson, & Reynolds, 1995). Auch wenn diese Argumente in vielerlei Hinsicht schlüssig sind, so können genannte Biases sicher nicht für jegliches Zustandekommen von Doppeldiagnosen herangezogen werden, was auch durch zahlreiche empirische Arbeiten untermauert wird (zusammenfassend bei Angold, Costello, & Erkanli, 1999; Clark, Watson, & Reynolds, 1995).

Den genannten Problemen und Schwierigkeiten in Zusammenhang mit dem Komorbiditätskonzept wird auf unterschiedliche Art und Weise begegnet. Auf einer übergeordneten theoretischen Ebene gibt es heftige Diskussionen um eine Neuordnung der derzeitigen Diagnosesysteme, wobei ein Schwerpunkt auf der Debatte um kategoriale und dimensionale Diagnostik liegt (etwa Brown & Barlow, 2005; Klein & Riso, 1993; Widiger & Samuel, 2005). Einige Forschungsanstrengungen wurden auch dahingehend unternommen, das gehäufte gemeinsame Auftreten bestimmter Störungen durch zugrundeliegende gemeinsame Faktoren zu erklären und einige psychische Störungen damit in einer dimensional statt kategorialen Form zu repräsentieren (etwa Clark & Watson, 1991; Krueger & Markon, 2006; McGlinchey & Zimmerman, 2007).

#### **2.2.4 Wert des Komorbiditätskonzeptes**

Auf Grundlage der dargestellten Kontroversen um Entstehungsmodelle für Komorbidität und um deren Validität sowie die Abhängigkeit von diagnostischen Systemen lässt sich die Frage nach dem grundlegenden Wert des Konzeptes anschließen. Argumente für die Bedeutsamkeit des Komorbiditätsprinzips finden sich auf drei Ebenen: Erstens spricht für das inhaltliche Gewicht des Konzeptes die vielfach beschriebene Beobachtung, dass sich Personen mit nur einer Diagnose von solchen mit mehreren Diagnosen hinsichtlich vieler Aspekte unterscheiden. Es gibt zahlreiche Hinweise auf einen Zusammenhang zwischen dem Vorliegen von mehr als einer Diagnose und einem größeren Schweregrad der psychischen Belastung (Clark et al., 1995), worauf in Kapitel 2.5 noch genauer eingegangen wird. Daraus ergibt sich ein zweites Argument, nämlich das der klinischen Nützlichkeit des Komorbiditätskonzeptes. Information über Mehrfachdiagnosen stellt einen informativen Mehrwert für den Kliniker dar, indem sie prognostische Implikation hat für das Ansprechen auf Behandlung, erwarteten Verlauf, Rückfallrate und Behandlungskosten (u. a. Blashfield, 1990; Brieger & Marneros, 2000; Clark et al., 1995; Maser & Cloninger, 1990). Als dritter Beleg für die Wichtigkeit des Konzeptes wird sein Potenzial zur Anregung von Forschung

genannt: häufig komorbide auftretende Störungen können einen Ausgangspunkt darstellen für die Untersuchung von möglicherweise geteilten Aspekten in Ätiologie und Pathogenese der betreffenden Störungen (u. a. Vella, Aragona, & Alliani, 2000) sowie von allgemeinen Veränderungsmechanismen, die relativ störungsunspezifisch wirken (Grawe, 2004; Widiger & Samuel, 2005). Genauso stimuliert die häufige Beobachtung von Komorbidität Diskussionen und Überprüfungen zur Unabhängigkeit und diskriminanten Validität psychischer Störungen (Maser & Cloninger, 1990) und trägt damit zu einer Weiterentwicklung aktueller Klassifikationssysteme bei (Angold, Costello, et al., 1999; Maj, 2005).

### **2.2.5 Offene Fragen zum Bereich der Komorbidität**

Zwar liegen viele Forschungsarbeiten zum Konzept, zu möglichen Entstehungsbedingungen, Korrelaten und Auswirkungen von Komorbidität vor; dennoch ist es schwierig, eindeutige Schlussfolgerungen zu ziehen. Gründe dafür liegen vor allem – bei einer insgesamt ungemein großen Menge an Literatur – in deren Kleinteiligkeit und Lückenhaftigkeit. Nahezu alle Studien zu diesem Zielbereich untersuchen Komorbidität an einer Auswahl von Störungen, wobei häufig gestellte Diagnosen und Diagnosekombinationen den größten Niederschlag finden. Meist wird eine „Zielstörung“ herausgegriffen, zu der dann mögliche Zusatzdiagnosen untersucht werden. Weitere Diagnosen werden nicht betrachtet oder dienen als Ausschlusskriterium. Ein genauer Überblick zu Befunden diagnostischer Komorbidität wird in Kapitel 2.4 gegeben. Ohne diesem vorzugreifen lässt sich schon an dieser Stelle festhalten, dass beim Zusammensetzen dieser vielen Informationsteile ein Bild entsteht, welches am besten mit der Aussage „alles ist möglich“ charakterisierbar ist: für den überwiegenden Teil psychischer Störungen lassen sich Belege für ein (überzufällig gehäuftes) gemeinsames Auftreten mit bestimmten anderen Störungen finden – jedoch oft auch Belege dagegen. Ganzheitliche Betrachtungen, das heißt unter Einschluss so vieler Diagnosen wie möglich, stammen vor allem aus der epidemiologischen Forschung (etwa Kessler et al., 2005; Jacobi, Klose & Wittchen, 2004). Umfangreiche Betrachtungen psychischer Komorbidität innerhalb definierter Stichproben, die unter Routinebedingungen behandelt wird, fehlen. Daher soll diese Arbeit unter anderem die Frage beantworten, welche psychischen Störungen bei einer breiten Palette an Einschlussdiagnosen in einer Stichprobe ambulanter Psychotherapiepatienten gehäuft gemeinsam diagnostiziert werden und ob sich

bestimmte Muster paralleler Diagnosen identifizieren lassen (Hypothesen 2 und 3, Kap. 2.4).

Daneben stellt sich nicht nur die Frage, inwieweit, wie oben dargestellt, Komorbidität in Abhängigkeit vom zugrundeliegenden Klassifikationssystem variiert, sondern auch, ob und inwiefern Raten von Mehrfachdiagnosen sich in Abhängigkeit vom gewählten diagnostischen Ansatz in ein und demselben Klassifikationssystem unterscheiden. Wie in Kapitel 2.3 weiter ausgeführt, existieren zu dieser Thematik einige Untersuchungen, die jedoch ausschließlich an den zwei Extrempunkten der diagnostischen Strukturiertheit ansetzen: strukturiertes klinisches Interview und klinisches Urteil ohne Vorgaben. Praktikable Alternativen zur Unterstützung der Diagnostik im klinischen Alltag, wie Diagnosechecklisten, wurden bislang nicht hinsichtlich der Auswirkungen auf Art und Anzahl gestellter Diagnosen mit klinischem Routinevorgehen verglichen (Hypothese 1, Kap. 2.3).

Eine weitere nicht gänzlich geklärte Problematik betrifft den Zusammenhang zwischen Komorbidität und Belastungsschwere. Wie schon angedeutet, lassen sich dazu eine Reihe von Korrelationen finden, woraus sich die Frage ergibt ob Komorbidität gleichzusetzen ist mit einer großen allgemeinen psychischen Belastetheit. Die hierzu existierende Literatur ist aufgrund unterschiedlichster Stichproben, Störungsgruppen und Ergebnismaße nur schwer vergleichbar. Eine differenzierte Betrachtung der selbst- und fremdberichteten Belastung bei verschiedenen reinen und komorbiden Störungsbildern in einer großen ambulanten Stichprobe fehlt bislang (Hypothese 4, Kap. 2.5).

Während die Literatur zum Zusammentreffen bestimmter Störungskombinationen und zum Status komorbider Patienten vor oder zu Beginn einer Behandlung nur schwer zu überblicken ist, ist die Forschung zu den Auswirkung von Mehrfachdiagnosen auf den Verlauf von Psychotherapien vergleichsweise lückenhaft. Wenn jedoch der Wert des Komorbiditätsprinzips gerade in der Unterstützung individueller Therapieplanung liegen soll, so ist die Frage naheliegend, ob Therapien komorbider Patienten anders verlaufen als solche von Patienten mit nur einer Diagnose. Darüber hinaus lässt sich fragen, ob komorbide Patienten sich nicht nur gegebenenfalls in anderer Art und Weise verändern, sondern auch, ob sie insgesamt in gleichem Ausmaß von einer Psychotherapie profitieren können wie nicht-komorbide Patienten – ist also das Therapieergebnis von Komorbidität beeinflusst? Einige Studien dazu werden in Kapitel 2.6 vorgestellt, welches auf die Hypothesen 5 bis 10 dieser Arbeit hinleitet.

## **2.3 Zusammenhang zwischen der Art und Anzahl diagnostizierter Störungen und der Methode der Diagnosestellung**

Wie in Kapitel 2.2.3 diskutiert wurde, ist Komorbidität zwangsläufig beeinflusst von Aufbau, Struktur und Ausdifferenziertheit der zum Erhebungszeitpunkt bestehenden Diagnosesysteme. Darüber hinaus resultieren sehr stark unterschiedliche Ergebnisse hinsichtlich der Diagnosehäufigkeit psychischer Störungen in Abhängigkeit von Art und Weise der Diagnosestellung. Um psychische Störungen zu diagnostizieren reicht das Spektrum von unstrukturierten klinischen Erhebungen, mehr oder weniger orientiert an den in gegenwärtigen Diagnosesystemen definierten Kriterien, über Selbstbericht-Fragebögen bis hin zu standardisierten und computergestützt erhobenen Interviews.

### **2.3.1 Befunde**

Interviewverfahren zur Unterstützung der Diagnosestellung sind in der Forschung weit verbreitet. Als Vorteile voll- und teilstrukturierter Interviewverfahren werden die höhere Interrater-Reliabilität hervorgehoben, sowie die sehr viel umfassendere Abfrage psychischer Symptome (Sher & Trull, 1996; Zimmerman, 1994). In der Praxis ist es dennoch ein häufiges und oft wenig hinterfragtes Vorgehen, Diagnosen anhand von unstrukturierten Befragungen zu vergeben (Miller et al., 2001). In der unterschiedlichen Diagnosestellung liegt ein Grund für die divergenten und teilweise widersprüchlichen Prävalenz- und Komorbiditätszahlen vieler Störungen (Clark et al., 1995). Einige Studien gehen gezielt der Frage nach, inwieweit unterschiedliche diagnostische Methoden zu abweichenden Ergebnissen gelangen beziehungsweise für welche Methode mit den verlässlichsten und validesten Diagnosen zu rechnen ist.

Beispielsweise untersuchten Zimmerman und Mattia (1999a) den Einfluss von Erhebungsmethoden auf die Häufigkeit der Diagnose von Borderline-Persönlichkeitsstörungen in der ambulanten Praxis. Es wurden 909 Patienten verglichen, deren Diagnosen entweder anhand eines unstrukturierten klinischen Interviews oder anhand eines halbstrukturierten Interviews (SCID-P; First, Spitzer, Gibbon, & Williams, 1994) gewonnen wurden. Der Anteil solcher Personen mit Diagnose einer Borderline-Persönlichkeitsstörung war in der Gruppe der Personen, mit denen ein strukturiertes Interview durchgeführt wurde, signifikant erhöht. Die Raten stiegen auch in der Gruppe mit unstrukturierten Befragungen an,

wenn den Klinikern zusätzliche Information aus strukturierten Interviews zur Verfügung gestellt wurde - ein Hinweis darauf, dass Daten aus strukturierten Interviews als vollständiger und valider einzustufen sind und dass im klinischen Routinevorgehen schlichtweg übersehene Information in weniger Diagnosen resultiert. Zu einem ähnlichen Schluss kommen auch Zimmerman, Rothschild und Chelminski (2005) in einem Übersichtsartikel.

Die Beobachtung einer deutlich häufigeren Diagnose bestimmter Störungen und Störungskombinationen unter strukturierter Diagnostik wird nicht nur bei Persönlichkeitsstörungen, sondern auch bei Störungen der Achse I gemacht. Hier wird bei klinischer Routinediagnostik mit unstandardisiertem Vorgehen von durchschnittlich einem Viertel der Patienten mit komorbiden Störungen berichtet; demgegenüber stehen etwa ein Drittel bis die Hälfte komorbider Patienten nach strukturierten Interviews (Wittchen, 1996; Zimmerman & Mattia, 1999b). In zwei Gruppen ambulanter nicht-psychotischer Patienten wurden die in den Patientenakten nach klinischem Routinevorgehen festgehaltenen Diagnosen mit strukturiert erhobenen Diagnosen (SKID) für dieselben Patienten verglichen (Shear et al., 2000). Die Übereinstimmung zwischen den beiden diagnostischen Methoden ist mit kappa-Werten zwischen 0.1 und 0.3 als mangelhaft einzustufen. Vor allem Angststörungen wurden mit Hilfe des SKID-Interviews bedeutend häufiger diagnostiziert, wohingegen die Diagnose einer Anpassungsstörung häufiger nach dem unstrukturierten klinischen Vorgehen gestellt wurde. In einer weiteren Untersuchung verglichen Zimmerman und Mattia (1999b) die Komorbiditätsraten von ambulanten Patienten, die entweder anhand eines unstrukturierten klinischen Interviews oder anhand eines halbstrukturierten Interviews diagnostiziert wurden. Die Wahrscheinlichkeit für diagnostische Komorbidität war unter unstrukturierten klinischen Erhebungen deutlich reduziert, wobei hier die phobischen Störungen und nicht näher bezeichnete depressive Störungen in der unstrukturierten Diagnostik am häufigsten übersehen wurden. Im stationären Kontext wurden ebenfalls Diagnosemethoden verglichen, wobei sich auch hier die strukturierten Methoden als überlegen erwiesen (Miller et al., 2001).

Zusammenfassend lässt sich festhalten, dass die Art und die Anzahl diagnostizierter Störungen je nach diagnostischer Methode variieren, dass im Allgemeinen unter Verwendung strukturierter Interviews mit reliableren und mehr Diagnosen zu rechnen ist als unter unstrukturiertem Routinevorgehen und dass die Diagnosewahrscheinlichkeit für bestimmte

Gruppen von Störungen, insbesondere Angststörungen, unter Verwendung strukturierter Interviews steigt.

### **2.3.2 Diagnosechecklisten - ein Kompromiss?**

Obwohl sie in der Forschung weit verbreitet sind, liegt das Problem diagnostischer Interviews in der Praxis in der oft begrenzten Zeit, die dem Kliniker zur Verfügung steht. Angesichts der fraglichen Validität von Diagnosen, die nach unstandardisiertem Routinevorgehen gestellt wurden (Costello, 1993) besteht ein alternatives Vorgehen in der Anwendung von Screening-Instrumenten und Checklisten zur Unterstützung der Störungsdiagnostik.

Dass auch Screeningfragebögen zu ähnlich hohen Komorbiditätsraten wie halbstrukturierte Interviews führen können zeigte eine Untersuchung von Rush, Zimmerman und Kollegen (2005). Auf die trotz guter Reliabilität bestehenden Nachteile von strukturierten Interviews gehen Hiller, Zaudig, Mombour und Bronisch (1993) ein: sie nennen neben dem schon erwähnten Zeitaufwand auch die relative Umständlichkeit in der Anwendung und die Schwierigkeit der Durchführung bei Patienten mit sehr schwerer oder akuter Symptomatik, bei hirngestörten oder schlichtweg wenig kommunikativen Patienten. Als Alternative zu Interviews haben sie diagnostische Checklisten entwickelt, die eine systematische Begutachtung von Patienten anhand der in ICD-10 beziehungsweise DSM-IV definierten Kriterien erleichtern sollen (Bronisch, Hiller, Mombour, & Zaudig, 1995; Hiller, Zaudig, & Mombour, 2004). Bronisch und Mombour (1994) verglichen die Diagnosechecklisten für Persönlichkeitsstörungen (IDCL-P) mit einem strukturierten Interview zur Erhebung von Persönlichkeitsstörungen nach DSM-III-R und ICD-10. Die Übereinstimmung zwischen diesen Diagnosemethoden lag für spezifische Persönlichkeitsstörungen in einer ähnlichen Größenordnung wie in Studien, die zwei strukturierte Interviews miteinander verglichen; für die Identifikation irgendeiner Persönlichkeitsstörung wurde eine größere Übereinstimmung als in Vergleichsstudien erreicht.

Inwieweit unterscheiden sich nun die nach IDCL erhobenen Diagnosen von solchen Diagnosen, die nach klinischen Routinevorgehen gestellt werden? Aus den bisherigen Untersuchungen zu den Diagnosechecklisten lässt sich schließen, dass diese in Sensitivität und Spezifität strukturierten Interviews nahe kommen. Jedoch konnte keine Studie gefunden

werden, die die Prävalenz und Komorbidität psychischer Störungen in einer Stichprobe untersucht, die sich hinsichtlich der Diagnosemethode - per IDCL-Checklisten versus klinisches Routinevorgehen - unterscheidet.

## 2.4 Befunde zur diagnostischen Komorbidität

Wie stellen sich nun Prävalenz und Komorbidität psychischer Erkrankungen dar? Einen guten ersten Überblick bieten Studien, in denen die Auftretenshäufigkeit psychischer Störungen in der Allgemeinbevölkerung untersucht wird. Ausgehend vom Zusatzsurvey "Psychische Störungen" des Bundesgesundheits surveys (Bellach, Knopf, & Thefeld, 1998) liegt die 12-Monats-Prävalenz für psychische Störungen in Deutschland bei 31%, die Querschnittprävalenz (4 Wochen) bei ca. 20% der erwachsenen Bevölkerung zwischen 18 und 65 Jahren. Rund 64% der Personen mit aktueller psychischer Störung (4-Wochen-Prävalenz) erfüllen die Kriterien nur einer Störung, wohingegen 36% unter zwei oder mehr Störungen leiden (Jacobi, Wittchen, et al., 2004). In einer vergleichbaren repräsentativen epidemiologischen Studie in den USA, dem National Comorbidity Survey, NCS (z. B. Kessler et al., 2004) ergab sich im 12-Monats-Zeitraum unter den Personen, die die Kriterien für mindestens eine psychische Störung erfüllen ein Anteil von rund 45% mit weiteren Störungen.

Unter dieser zunächst noch sehr allgemeinen Betrachtungsweise ist festzuhalten, dass bei Vorliegen einer psychischen Störung die Wahrscheinlichkeit für weitere psychische Störungen etwa im Bereich ein Drittel bis die Hälfte liegt. Im Folgenden soll nun ein differenzierteres Bild gezeichnet werden, indem die diagnostische Komorbidität für verschiedene Gruppen psychischer Störungen getrennt betrachtet wird. Zum gemeinsamen Auftreten psychischer Störungen besteht eine Fülle an Literatur. Einige Arbeiten befassen sich schwerpunktmäßig mit diesem Thema, häufig tauchen Prävalenz- und Komorbiditätsangaben aber auch lediglich in Stichprobenbeschreibungen oder als Ergebnis von Nebenfragestellungen auf. Ziel an dieser Stelle ist somit nicht eine erschöpfende Darstellung von Komorbiditätsbefunden zu den verschiedenen Störungsbildern. Vielmehr soll ein Überblick zu Strömungen und vielfach replizierten Zusammenhängen anhand beispielhafter Studien gegeben werden. Die Ergebnisse stammen dabei sowohl aus epidemiologischen als auch aus behandelten Stichproben und sowohl aus naturalistischen als auch kontrollierten Behandlungssettings.

### 2.4.1 Affektive Störungen

Bei Betrachtung von 12-Monats-Prävalenzen sind in epidemiologischen Studien für ungefähr 60% der Personen mit unipolarer Depression die Kriterien für mindestens eine weitere Störung erfüllt, dies besonders häufig aus der Gruppe der Angststörungen und hier vor allem der Generalisierten Angststörung (Jacobi, Klose, & Wittchen, 2004; Kessler et al., 1996; Kessler, DuPont, Berglund, & Wittchen, 1999). Ähnlich hohe Überschneidungsraten von sowohl Major Depression als auch Dysthymia insbesondere mit Angststörungen ergeben sich auch aus Untersuchungen an ambulanten Stichproben (Rush, Zimmerman, et al., 2005; Sanderson, Beck, & Beck, 1990). Daneben ergeben sich sehr hohe Raten (zwischen 30% und 40%) gemeinsamen Auftretens zwischen Depression und Persönlichkeitsstörungen (Diguer, Barber, & Luborsky, 1993; Hardy, Barkham, et al., 1995; Hirschfeld, 1999; Pilkonis & Frank, 1988; Shea et al., 1990; Shea, Widiger, & Klein, 1992).

Bei Personen mit bipolarer Erkrankung wurde im Bundesgesundheitsurvey für nur etwa ein Viertel der Erkrankten im Zwölfmonatszeitraum eine reine Diagnose gestellt, ein großer Anteil (rund 35%) erhält dagegen sogar drei oder mehr Zusatzdiagnosen (Jacobi, Wittchen, et al., 2004). Mit Abstand am häufigsten werden Zusammenhänge mit Angststörungen berichtet (u. a. Freeman, Freeman, & McElroy, 2002), insbesondere mit Panikstörung (Angst, 1998; Chen & Dilsalver, 1995; Feske et al., 2000; Young, Cooke, Robb, Levitt, & Joffe, 1993) und Generalisierter Angststörung (Young et al., 1993) aber auch Zwangsstörungen (Chen & Dilsaver, 1995; Kruger, Cooke, Hasey, Jorna, & Persad, 1995).

### 2.4.2 Angststörungen

Gut 60% der Patienten mit einer Angststörung erfüllen für die zurückliegenden zwölf Monate die Kriterien für mindestens eine weitere Diagnose (Jacobi, Wittchen, et al., 2004). Angststörungen sind oft untereinander komorbide, werden aber auch sehr häufig parallel zu affektiven Störungen, substanzinduzierten Störungen oder Persönlichkeitsstörungen diagnostiziert. Im Folgenden wird genauer auf die Befunde zur Komorbidität verschiedener Angststörungen eingegangen.

### *Generalisierte Angststörung*

Für die Generalisierte Angststörung (GAS) werden häufig sehr hohe Komorbiditätsraten, insbesondere mit anderen Angststörungen<sup>2</sup> (z. B. Brawman-Mintzer et al., 1993) und Depression (z. B. Kessler, DuPont, et al., 1999), berichtet. Jedoch wird gerade für dieses Störungsbild ein großer Hilfesuch-Bias angenommen (Kessler, Keller, & Wittchen, 2001), das heißt, dass Personen mit Generalisierter Angststörung sich oft erst dann in Behandlung begeben, wenn weitere psychische Probleme bestehen. So liegen in klinischen Untersuchungen die Komorbiditätsraten zwischen 40% und 80%, darunter neben depressiven Störungen (z. B. Borkovec, Abel, & Newman, 1995; Newman, Przeworski, Fisher, & Borkovec, 2010) unter Anwendung des DSM auch sehr häufig soziale Phobie gefolgt von spezifischer Phobie (Borkovec, Abel, et al., 1995; Brawman-Mintzer et al., 1993; Newman et al., 2010). Aber auch in der Allgemeinbevölkerung erfüllt ein Großteil der Patienten mit Generalisierter Angststörung die Kriterien einer Depression (12-Monats-Prävalenz zwischen 58% und 70%; Jacobi, Wittchen, et al., 2004; Kessler, DuPont, et al., 1999).

### *Panik/Agoraphobie*

Die Raten für eine oder mehrere komorbide Diagnosen bei Lifetime-Panikstörung mit oder ohne Agoraphobie liegen in der National Comorbidity Study über 80% (Kessler et al., 2006), in klinischen Stichproben zwischen 50% und 80% (Brown, Antony, & Barlow, 1995; Chambless, Renneberg, Gracely, Goldstein, & Fydrich, 2000). Hohe Überschneidungsraten finden sich hier vor allem mit anderen Angststörungen und dies besonders dann, wenn auch eine Agoraphobie besteht. Hohe Komorbidität besteht aber auch für affektive Störungen, Impulskontrollstörungen, Substanzmissbrauch sowie Persönlichkeitsstörungen (Noyes et al., 1990).

### *Spezifische Phobie*

Zur Häufigkeit des Vorkommens anderer Störungen bei Vorliegen einer spezifischen Phobie existiert nur wenig Literatur. In der Allgemeinbevölkerung finden sich deutliche Zusammenhänge mit anderen Angststörungen, affektiven Störungen, Impulskontrollstörungen, Somatoformen Störungen und Alkoholabhängigkeit (Choy, Fyer, & Goodwin, 2007;

---

<sup>2</sup> In diesem Zusammenhang muss allerdings darauf hingewiesen werden, dass im ICD-10, im Gegensatz zum DSM-IV, für die Generalisierte Angststörung einige andere Angststörungen explizit als Ausschlussdiagnosen definiert sind, nämlich Panikstörung (F41.0), phobische Störung (F40) und Zwangsstörung (F42).

Depla, Have, Balkom, & Graaf, 2007; Essau, Conradt, & Petermann, 2000; Kessler et al., 2005).

### *Soziale Phobie*

Komorbidität bei sozialer Phobie findet sich hauptsächlich für andere Angststörungen, darunter vor allem Panik, Agoraphobie und GAS (Andrews, Slade, & Issakidis, 2002; Schneier, Johnson, Hornig, Liebowitz, & Weissman, 1992), und affektive Störungen (Essau, Conradt, & Petermann, 1999; Kessler, Stang, Wittchen, Stein, & Walters, 1999; Kessler, Stein, & Berglund, 1998; Schneier et al., 1992; Stein & Kean, 2000). Bei Vorliegen einer komplexen im Vergleich zu einer einfachen sozialen Phobie sind die Raten für weitere Störungen aus dem Bereich der Angststörungen und affektiven Störungen noch einmal deutlich erhöht (Kessler et al., 1998). In klinischen Stichproben stellen soziale Phobien die häufigste Zweitdiagnose bei anderen Angststörungen dar (Brunello et al., 2000).

### *Zwangsstörung*

Nach epidemiologischen Studien weisen 60% der Personen mit einer Zwangsstörung mindestens zwei (Torres et al., 2006) oder sogar drei oder mehr zusätzliche Diagnosen auf (Jacobi, Wittchen, et al., 2004). Vergleichbar hohe Raten finden sich in klinischen Studien (Abramowitz, Franklin, Street, Kozak, & Foa, 2000; Tükel, Polat, Özdemir, Aksüt, & Türksoy, 2002). Mit einer Zwangsstörung gehäuft parallel diagnostiziert werden depressive Störungen (Kessler et al., 2005; Steketee, Chambless, & Tran, 2001; Tükel et al., 2002), andere Angststörungen (Steketee et al., 2001) und bipolare Erkrankungen (Angst et al., 2004; Perugi et al., 1997). Dagegen liegen zum gehäuften Auftreten von substanzinduzierten Störungen widersprüchliche Befunde vor (Angst et al., 2004; Kessler et al., 2005).

## **2.4.3 Somatoforme Störungen**

Im Vergleich zu anderen Störungsgruppen treten Somatoforme Störungen in Bevölkerungsstichproben relativ häufig ohne Zusatzdiagnosen auf (45% Einzeldiagnosen im Zwölfmonats-Zeitraum; Jacobi, Wittchen, et al., 2004). Jedoch werden in Hausarztstudien (de Waal, Arnold, Eekhof, & van Hemert, 2004; Mergl et al., 2007) und einer stationären klinischen Stichprobe (Rief, Schäfer, Hiller, & Fichter, 1992) hohe Überschneidungsraten

mit Angst- oder depressiven Störungen, aber auch Essstörungen und substanzinduzierten Störungen gefunden. Daneben wird über eine sehr hohe Komorbidität mit Persönlichkeitsstörungen berichtet (Bass & Murphy, 1995).

#### **2.4.4 Persönlichkeitsstörungen**

Hohe Komorbiditätsraten werden sowohl innerhalb der Gruppe der Persönlichkeitsstörungen, als auch mit anderen psychischen Störungen berichtet (z. B. Fiedler, 1998). Beispielsweise finden sich parallele Persönlichkeitsstörungsdiagnosen häufig mit depressiven Störungen, Angststörungen, Somatoformen Störungen, Essstörungen und Zwangsstörungen, wobei die Komorbidität mit Achse-I-Störungen in Abhängigkeit des Persönlichkeitsstörungs-Clusters<sup>3</sup> variiert (Skodol et al., 1995; Zimmerman et al., 2005). Darüber hinaus ist bei Vorliegen einer Persönlichkeitsstörung die Wahrscheinlichkeit für eine weitere Diagnose aus diesem Bereich deutlich erhöht (Stuart et al., 1998; Zimmerman et al., 2005).

#### **2.4.5 Belastungs- und Anpassungsstörungen**

Zur psychischen Komorbidität von Anpassungsstörungen existiert nur wenig Literatur. Strain und Kollegen (1998) berichten im stationären Kontext nur für einen geringen Anteil (7.4%) der Personen mit diagnostizierter Anpassungsstörung Komorbidität mit anderen Achse-I- oder Achse-II-Störungen, darunter am häufigsten organische psychische Störungen, Persönlichkeitsstörungen oder substanzinduzierte Störungen. Laut ICD-10 wird eine Anpassungsstörung bei Symptomen vergeben, die ähnlich derer bei affektiven Störungen, neurotischen, Belastungs- und somatoformen Störungen und Störungen des Sozialverhaltens sind, jedoch nur wenn die Kriterien der einzelnen Störungen nicht erfüllt werden. Somit sollte mit diesen Störungsbereichen keine Komorbidität bestehen. Von größerer Bedeutung als psychische Komorbidität scheint hier die somatische Komorbidität zu sein. Für die Posttraumatische Belastungsstörung (PTSD) werden dagegen sehr hohe und mit ansteigender Anzahl von PTSD-Symptomen größer werdende (Marshall et al., 2001) Komorbiditätsraten berichtet. Im National Comorbidity Survey ergibt sich für die vergangenen zwölf

---

<sup>3</sup> Im DSM-IV sind die Persönlichkeitsstörungen nach sogenannten Clustern aufgeteilt. Cluster A (sonderbar, exzentrisch) beinhaltet die paranoide, schizoide und schizotypische Persönlichkeitsstörung. Cluster B (dramatisch, emotional) beinhaltet die emotional instabile, histrionische, narzisstische und dissoziale Persönlichkeitsstörung. Cluster C (ängstlich, vermeidend) beinhaltet die selbstunsichere, abhängige, zwanghafte und passiv-aggressive Persönlichkeitsstörung.

Monate ein Anteil von rund 80% mit zusätzlicher Diagnose bei PTSD, darunter besonders häufig Angststörungen, Depression, Impulskontrollstörungen und Substanzmissbrauch (Kessler et al., 2005). Auch klinische Studien kommen zu hohen Überschneidungsraten (um die 40%) besonders mit Depression (Ferrada-Noli, Asberg, Ormstad, Lundin, & Sundbom, 1998; O'Donnell, Creamer, & Pattison, 2004; Shalev et al., 1998), aber auch mit Angststörungen (Deering, Glover, Ready, Eddleman, & Alarcon, 1996; Ferrada-Noli et al., 1998).

#### **2.4.6 Essstörungen**

Im Bundesgesundheitssurvey erfüllen nur 35% der Personen mit diagnostizierter Essstörung nicht die Kriterien für mindestens eine weitere Diagnose im Zwölfmonats-Intervall (Jacobi, Wittchen, et al., 2004). In Übersichtsarbeiten finden sich kontinuierlich hohe Überschneidungsraten mit affektiven und Angststörungen (Godard et al., 2007; Kaye, Bulik, Thornton, Barbarich, & Masters, 2004; O'Brien & Vincent, 2003; Pallister & Waller, 2008). Personen mit diagnostizierter Anorexia nervosa weisen besonders häufig zusätzlich eine affektive oder substanzinduzierte Störung, Personen mit diagnostizierter Bulimia nervosa darüber hinaus häufig Impulskontroll- und Angststörungen auf (Hudson, Hiripi, Pope, & Kessler, 2007). Zur Komorbidität mit Zwangsstörungen besteht eine gemischte Befundlage (Kaye et al., 2004; O'Brien & Vincent, 2003; Pallister & Waller, 2008). Während Personen mit Bulimia nervosa hohe Raten an Persönlichkeitsstörungen des Clusters B und dort besonders der Borderline Persönlichkeitsstörung aufweisen, sind für Personen mit der Diagnose Anorexia nervosa die Raten für Störungen des Clusters C erhöht (O'Brien & Vincent, 2003).

#### **2.4.7 Zusammenfassung**

Zur Komorbidität psychischer Störungen existieren zahllose Studien, von denen hier nur ein Bruchteil vorgestellt werden konnte. Die vorangehenden Abschnitte geben dennoch einen guten Überblick über die Vielfalt und Vielgestaltigkeit der Forschung zur Auftrenshäufigkeit psychischer Störungen und ihrer Überschneidungen.

Als wichtigsten Punkt lässt sich fraglos festhalten, dass die parallele Diagnose mehrerer Störungen die Regel darstellt, und zwar über viele Störungsgruppen hinweg. Eine Aus-

nahme stellt lediglich die Anpassungsstörung dar, für die hier Mehrfachdiagnosen unter 10% berichtet werden. Allerdings existiert zu den Anpassungsstörungen im Vergleich zu anderen Störungen nur sehr wenig Literatur. Es kann vermutet werden, dass das Verzeichnen von Anpassungsstörungen bei Vorliegen weiterer Störungen häufig unterbleibt und nach den Diagnosesystemen auch in den meisten Fällen nicht möglich ist, da die Symptome einer Anpassungsstörung bereits in der spezifischen Diagnose klassifiziert sind und angenommen wird, dass eine zusätzliche Diagnose hier keinen Mehrwert für Problemverständnis und Therapieplanung schafft (Bengel & Hubert, 2010). Für alle weiteren betrachteten Störungsgruppen ergeben sich hohe bis sehr hohe Überschneidungsraten. Affektive und Angststörungen haben die größten Prävalenzzahlen psychischer Störungen und treten dementsprechend besonders häufig parallel mit weiteren Diagnosen auf. Affektive und Angststörungen zeigen auch untereinander hohe Raten gemeinsamen Auftretens (insbesondere Depression und Generalisierte Angststörung; Borkovec et al., 1995; Kessler, Dupont, et al., 1999; Newman et al., 2010), genauso wie oft mehrere Angststörungen parallel diagnostiziert werden (z. B. Andrews et al., 2002; Borkovec et al., 1995; Chambless et al., 2000). Daneben sind häufige Komorbiditätsmuster solche zwischen Depression und Persönlichkeitsstörungen (u. a. Hirschfeld, 1999; Shea et al., 1990, 1992). Letztere werden vergleichsweise selten alleine diagnostiziert. Viele Studien berichten beeindruckende Prozentangaben zur Überschneidung bestimmter Störungsgruppen. Diese Angaben stehen jedoch in direkter Abhängigkeit zur Prävalenz der betrachteten Störungen, das heißt ihrer jeweiligen Basisraten, und sind daher nur beschränkt geeignet, Aussagen über das Zusammentreffen zweier Störungen zu machen, die über das rein nach Auftretenshäufigkeit der Einzelstörungen zu erwartende Maß hinausgehen. Krueger (2002) betont im Zuge dessen die dringende Notwendigkeit, im Rahmen der Komorbiditätsforschung zwischen den Konzepten „co-occurrence“ (welches ein eher zufälliges Zusammentreffen bezeichnet) und Korrelation (gemeinsames Auftreten, welches jenseits des Zufallsniveaus liegt) zu unterscheiden. Präzisere Aussagen erlauben gegenüber einfachen Prozentangaben solche Verfahren, die statistisch eine Gewichtung vornehmen zwischen der Wahrscheinlichkeit, Gleichzeitigkeitsdiagnosen zu finden versus nicht zu finden, etwa Odds Ratios oder tetrachorische Korrelationen. Als weiteres Fazit aus den zusammengestellten Befunden ist festzuhalten, dass nicht selten mehr als zwei Diagnosen gemeinsam auftreten. Häufig werden neben der Hauptdiagnose zwei, drei oder mehr zusätzliche Diagnosen vergeben, so dass die Zeich-

nung von klaren, das heißt empirisch belegten und in unterschiedlichen Stichproben replizierbaren Mustern gemeinsamen Auftretens erschwert ist.

## **2.5 Befunde zu Komorbidität und Belastungsschwere zu Therapiebeginn**

Wie in Kapitel 2.2 zusammengetragen, gibt es eine rege Debatte über mögliche Ursachen sowie Sinn und Unsinn von Mehrfachdiagnosen. Es stellt sich die Frage, inwieweit die Vergabe von zwei, drei oder noch weiteren Zusatzdiagnosen nicht nur einen quantitativen, sondern auch einen qualitativen Unterschied darstellt. Sind Personen mit mehr als einer psychischen Störung tatsächlich stärker belastet, erleben größere Einschränkung und Gefährdung als solche mit einer Störungsdiagnose? Oder aber beinhalten die gleich- oder nachrangigen parallel diagnostizierten Störungen letztlich keine zusätzliche Information über das Befinden der diagnostizierten Person und angemessene Behandlungsmöglichkeiten?

In den folgenden Abschnitten soll ein Überblick gegeben werden zu bisherigen Forschungsarbeiten, die Mehrfachdiagnosen in Zusammenhang mit verschiedenen Belastungsbereichen bringen. Dabei wird neben der wahrgenommenen allgemeinen Symptombelastung auf Schwere und Dauer der Hauptsymptomatik, auf Suizidalität, allgemeines Funktionieren, sowie Lebensqualität und Hilfesuchverhalten eingegangen. Der Fokus liegt auf unbehandelten Personen beziehungsweise deren Einschränkungen und Beschwerdebereiche zu Beginn einer Intervention. Der Einfluss von Komorbidität auf den Verlauf und das Ergebnis einer Behandlung wird in einem späteren Unterkapitel 2.6 behandelt.

### **2.5.1 Allgemeine Symptombelastung / Körperliche und geistige Einschränkungen / gesundheitsbezogene Lebensqualität**

Epidemiologische Studien und Studien aus der Primärversorgung untersuchen besonders häufig den Einfluss von psychischer Komorbidität auf die wahrgenommene körperliche und geistige Leistungsfähigkeit beziehungsweise auf die allgemeine gesundheitsbezogene Lebensqualität. Zum Einsatz kommt hier oft der Fragebogen zum Gesundheitszustand (in der deutschen Version: Bullinger & Kirchberger, 1998) in seiner Vollversion mit 36 Items oder in einer Kurzversion (12 Items). Bedeutsam erhöhte Belastungswerte bei Vorliegen von mehr als einer Diagnose ergeben sich sowohl in Bevölkerungsstichproben (Andrews et al., 2002; Wittchen, Carter, Pfister, Montgomery, & Kessler, 2000) als auch in einer Hausarztstudie (Schonfeld et al., 1997) und unter ambulant behandelten Patienten mit depressi-

ver Störung (Rush, Zimmerman, et al., 2005; Skodol et al., 2005). Ein weiteres allgemeines Maß für die subjektive Beeinträchtigung durch körperliche und psychische Symptome stellt die Symptom-Check-Liste (SCL-90; Derogatis, 1992; in der deutschen Version Franke, 1995) beziehungsweise in ihrer Kurzform der Brief Symptom Inventory (BSI; Derogatis, 1993; Franke, 2000) dar. In Studien, die sich mit der Komorbidität psychischer Störungen befassen, kommt diese jedoch vergleichsweise selten zum Einsatz. Zwei Studien berichten signifikant erhöhte Werte in allgemeiner Psychopathologie, erfasst über den SCL-90 beziehungsweise SCL-90-R, für Personen mit komorbiden Persönlichkeitsstörungen im Vergleich zu solchen mit reinen Achse-I-Störungen (Dreessen, Hoekstra, & Arntz, September; Hardy et al., 1995). Daneben wird in verschiedenen Untersuchungen auch über erhöhte Raten medizinischer Komorbidität (Newman, Moffitt, Caspi, & Silva, 1998; Rush, Zimmerman, et al., 2005) und über größere wahrgenommene körperliche Einschränkungen (de Waal et al., 2004; Wittchen et al., 2000; Wittchen, Lieb, Wunderlich, & Schuster, 1999) bei Vorliegen mehrerer psychischer Diagnosen berichtet. Komorbide Persönlichkeitsstörungen (Hardy et al., 1995), aber auch Mehrfachdiagnosen auf Achse I des DSM-IV (Shear et al., 2000), gehen darüber hinaus auch mit einer bedeutsam erhöhten interpersonellen Belastung, erfasst über das Inventar Interpersoneller Probleme (IIP-D; Horowitz, Strauss, & Kordy, 2000), einher.

### **2.5.2 Schwere der Hauptsymptomatik**

Zahlreiche Studien weisen darauf hin, dass das Vorliegen komorbider Störungen nicht nur zu größeren allgemeinen psychischen, interpersonellen und körperlichen Einschränkungen führt, sondern dass sich dies auch in der Schwere der Hauptsymptomatik niederschlägt. Die im Folgenden zitierten Studien nehmen eine Trennung zwischen Haupt- und Nebendiagnosen vor, indem sie für den Schweregrad einer definierten Hauptstörung Vergleiche anstellen zwischen Personen, für die allein diese Störung vorliegt und solchen Personen, die darüber hinaus Kriterien für weitere Diagnosen erfüllen.

Zur selbstberichteten depressionsbezogenen Symptombelastung mit und ohne Komorbidität liegen einige Studien vor. Sowohl bei komorbider Persönlichkeitsstörung (Diguier, Barber, & Luborsky, 1993; Hardy et al., 1995), als auch bei verschiedenen komorbiden Achse-I-Störungen (Kessler, Stang, et al., 1999; Rush, Zimmerman, et al., 2005; Sherbourne &

Wells, 1997) berichten Patienten eine schwerwiegendere depressive Symptomatik als bei Vorliegen einer reinen Depression.

Ebenso finden sich deutlich erhöhte Ausprägungen der Hauptsymptomatik bei komorbiden im Vergleich zu einfachen Diagnosen für Panikstörung (Brown et al., 1995; Dreesen, Arntz, Lutjels, & Sallaerts, 1994; Joormann et al., 2005; Kessler et al., 2006; Noyes et al., 1990), Soziale Phobie (Erwin, Heimberg, Juster, & Mindlin, 2002), Generalisierte Angststörung (Newman et al., 2010), Zwangsstörungen (Tükel et al., 2002), Posttraumatische Belastungsstörung (Shalev et al., 1998) und Somatoforme Störungen (de Waal et al., 2004; Hiller, Rief, & Fichter, 1997). Mineka, Watson und Clark (1998) ziehen die Schlussfolgerung "[...] it may be impossible at this stage to disentangle the issues of comorbidity and severity of disorders".

Allerdings bestehen auch dem widersprechende Befunde: Tsao, Lewin und Craske (1998) konnten für Panikpatienten mit und ohne psychische Komorbidität (Achse I) keine bedeutsamen Unterschiede im Schweregrad der Paniksymptomatik feststellen. Ferner fanden Rief, Trenkamp, Auer und Fichter (2000) trotz bedeutsamer Unterschiede in einigen Symptommaßen keine Signifikanzen im Bereich agoraphobischer Vermeidung bei Panikpatienten mit und ohne komorbide Depression. Newman et al. (2010) berichten für einen Teil ihrer Maße zu Beginn der Behandlung keine Differenzen zwischen Generalisierter Angststörung mit und ohne Depression. Ebenso fanden Joormann et al. (2005) vergleichbare Angstwerte bei Patienten mit reiner sozialer Phobie oder sozialer Phobie und gleichzeitiger Depression.

### **2.5.3 Dauer der Hauptsymptomatik**

Nahe liegend ist der Gedanke, dass Komorbidität nicht nur auf die Schwere der Hauptsymptomatik, sondern auch auf deren Dauer wirkt. In der Tat weisen nicht wenige Studien darauf hin, dass unter Mehrfachdiagnosen mit einem chronischen Verlauf zu rechnen ist. Dabei wird sich im Besonderen auf Depression bezogen (Mineka et al., 1998; Rush, Zimmerman, et al., 2005; Viinamäki et al., 2002), aber auch auf die Posttraumatische Belastungsstörung (McFarlane & Papay, 1992) oder psychische Störungen im Allgemeinen (Newman et al., 1998). Für die Bereiche Soziale Angst (Erwin et al., 2002) und Zwangs-

störungen (Tükel et al., 2002) wurden keine Unterschiede bezüglich der Länge der Erkrankung zwischen Personen mit reinen und komorbiden Diagnosen gefunden.

#### **2.5.4 Suizidalität**

Selbsttötungsgedanken, -absichten und -versuche sowie vollendete Suizide sind eine besonders schwere Komplikation psychischer Störungen, die zu erkennen von großer Wichtigkeit ist. Bei depressiven Störungen stellt Suizidalität ein Diagnosekriterium dar und ist damit auch besonders häufig zu erwarten. In der Tat scheint es so zu sein, dass das Vorliegen einer komorbiden Depression das Suizidalitätsrisiko auch bei anderen psychischen Störungen erhöht. So wiesen in einer Studie an ambulanten Patienten Personen mit Panikstörung und komorbider Depression bedeutsam mehr Suizidversuche in der Vergangenheit auf als Personen mit reiner Panikstörung; Unterschiede in suizidalen Gedanken bestanden jedoch nicht (Diaconu & Turecki, 2007). Umgekehrt stellt sich das Bild in einer Untersuchung an traumatisierten Patienten dar: hier wurden bei komorbider Depression mehr suizidale Gedanken, jedoch ein Trend zu weniger Suizidversuchen festgestellt (Ferrada-Noli et al., 1998). Bei depressiver Störung als Indexdiagnose ist sowohl in der Allgemeinbevölkerung (Bronisch & Wittchen, 1994) als auch in einer ambulanten Stichprobe (Rush, Zimmerman, et al., 2005) ein deutlicher Anstieg an Suizidversuchen zu verzeichnen, wenn daneben weitere Störungen bestehen. Besonders häufig ist die Kombination aus Depression und Angststörungen auch im Hinblick auf Suizidalität untersucht worden und hier ergeben sich erhöhte Raten für Doppeldiagnosen im Vergleich zu reinen Störungen, unabhängig davon welche Diagnose als vorrangig angesehen wird (Bronisch & Wittchen, 1994; Mineka et al., 1998; Wittchen et al., 1999). Bei bipolaren Erkrankungen stellt insbesondere eine erhöhte Angstsymptomatik einen Risikofaktor für Suizidalität dar (Tohen et al., 2007; Young, 1993). Keine Zusammenhänge wurden hingegen für die Bereiche Zwangsstörungen (Torres et al., 2006) und soziale Phobie (Kessler, Stang, et al., 1999) gefunden.

#### **2.5.5 Psychosoziale Beeinträchtigung und allgemeine Leistungsfähigkeit**

Mit der allgemeinen Leistungsfähigkeit rücken Aspekte in den Blick, die die Auswirkung psychischer Störungen in verschiedenen Lebensbereichen betreffen. Nach Achse V des DSM-IV werden psychische, soziale und berufliche Funktionen mit Hilfe der Skala zur

Globalen Erfassung des Funktionsniveaus (GAF) auf einem hypothetischen Kontinuum von psychischer Krankheit (Extremwert 1) bis Gesundheit (Extremwert 100) beurteilt. Signifikant verringerte GAF-Werte unter Komorbidität finden sich für verschiedenste Diagnosen, darunter Soziale Phobie (Rodebaugh, Holaway, & Heimberg, 2004), Depression (Markowitz et al., 2007; Wittchen et al., 1999), Posttraumatische Belastungsstörung (Shalev et al., 1998) und Essstörungen (Skodol et al., 1993). Auch unter Anwendung alternativer Maße ergeben sich größere Funktionseinschränkungen unter Komorbidität bei Depression (Kessler, DuPont, et al., 1999; Rush, Zimmerman, et al., 2005; Shea et al., 1990), Angststörungen (Kessler et al., 2006; Kessler, DuPont, et al., 1999; Mineka et al., 1998; Stein & Kean, 2000; Wittchen et al., 2000) und Somatoforme Störungen (de Waal et al., 2004; Hiller et al., 1997). Damit verbunden werden für Personen mit mehreren psychischen Störungen im Vergleich zu einfach Belasteten ein geringeres Jahreseinkommen (Erwin et al., 2002), häufigere Umzüge, weniger soziale Unterstützung, geringere familiäre Bindung und höhere Arbeitslosigkeitsraten (Newman et al., 1998) berichtet.

### **2.5.6 Lebensqualität, Hilfesuchverhalten, Behandlungsbereitschaft**

Entsprechend der dargestellten vielfältigen Belastungen geben Personen mit mehreren psychischen Störungen eine bedeutsam eingeschränkte allgemeine Lebensqualität an (Newman et al., 1998; Rush, Zimmerman, et al., 2005; Wittchen et al., 1999). Der erhöhte Leidensdruck führt zu einer größeren Wahrscheinlichkeit des Aufsuchens professioneller Hilfe (Andrews et al., 2002; Kessler et al., 2003; Mineka et al., 1998; Torres et al., 2006). Bei bestimmten Störungsbildern, wie einigen Persönlichkeitsstörungen oder der Generalisierten Angststörung besteht ein regelrechter Hilfesuchbias bei Komorbidität (Fiedler, 1998; Kessler et al., 2001). Komorbidität scheint außerdem dazu zu führen, dass Hausärzte psychische Störungen mit einer größeren Wahrscheinlichkeit entdecken und dass diese in der Folge angemessen behandelt werden (Sartorius, Üstün, Lecrubier, & Wittchen, 1996).

### **2.5.7 Zusammenfassung**

Über verschiedenste Störungsbilder hinweg, in unterschiedlichen Stichproben und in zahlreichen Zielbereichen, erfasst über diverse Instrumente finden sich Belege für erhöhte Belastungswerte bei Komorbidität. Demgegenüber gibt es aber auch Befunde, die diesem

Bild widersprechen, und zwar insbesondere für die Diagnosen Zwangsstörung und soziale Phobie, die sich in verschiedenen Zielbereichen unter Komorbidität nicht schwerer belastet darstellen. Studien, die keinen Unterschied in den Belastungswerten zwischen komorbiden und monomorbiden Personen berichten, stellen alles in allem jedoch Ausnahmen dar. Über die Gesamtheit der dargestellten Befunde hinweg scheint das Vorliegen mehrerer Diagnosen somit einen qualitativen Unterschied in der selbstberichteten und durch Fremdbeurteiler eingeschätzten Belastung zu bewirken.

Es fällt auf, dass zur Erfassung einer allgemeinen Belastung sehr häufig der Fragebogen zum Gesundheitszustand (SF-36 beziehungsweise SF-12; Bullinger & Kirchberger, 1998) zum Einsatz kommt. Demgegenüber finden sich vergleichsweise wenige Arbeiten, die die ansonsten in der Forschung weit verbreiteten Instrumente SCL-90 (Symptom-Check-Liste; Derogatis, 1992) beziehungsweise BSI (Brief Symptom Inventory; Derogatis, 1993) und IIP (Inventar Interpersonaler Probleme; Horowitz et al., 2000) im Kontext komorbider Störungen anwenden. Die zitierten Studien beschränken sich meist auf spezifische Störungskombinationen und untersuchen jeweils nur ausgesuchte Zielbereiche der durch psychische Störungen hervorgerufenen Belastung. Aufgrund der zahlreichen Einzelbefunde ist es schwierig, zusammenfassende oder auch differenzielle Aussagen über den Zusammenhang von Mehrfachdiagnosen und Belastungswerten zu treffen. Studien, die gleichzeitig ein breites Spektrum psychischer Störungen und damit möglicher Komorbiditäten einschließen und daneben verschiedene Maße zur Erfassung der allgemeinen psychischen Belastung aus einer Selbst- und Fremdbeurteilerperspektive erheben, fehlen bislang.

## 2.6 Einfluss von Komorbidität auf Therapieergebnis und –verlauf

Wie im vorangehenden Kapitel dargestellt, liegen zahlreiche Studien zum Zusammenhang komorbider Störungen mit verschiedenen Belastungsmaßen in epidemiologischen oder in unbehandelten klinischen Stichproben vor. Oft weisen Personen mit mehr als einer Diagnose höhere Eingangsbelastungswerte auf. Welche Auswirkungen haben diese veränderten Startbedingungen auf den Verlauf und das Ergebnis einer Therapie? Ist für Personen mit komorbiden Störungen mit einem komplizierteren, längeren und weniger erfolgreichen Verlauf zu rechnen? Oder haben sie aufgrund der erhöhten Eingangswerte in verschiedenen Symptommaßen mehr „Raum zur Verbesserung“ und profitieren damit möglicherweise sogar mehr als Personen mit einer Diagnose? Diesen Fragen soll im folgenden Abschnitt nachgegangen werden.

### 2.6.1 Komorbidität und Therapieverläufe

Zur Untersuchung des Einflusses von Komorbidität auf Veränderungsprozesse wird mit Messwiederholungen mehrfach im Verlauf und zum Teil auch mehrfach nach Beendigung einer Behandlung gearbeitet. In den nachfolgend vorgestellten Studien kommen dazu Hierarchische Lineare Modelle (auch Mixed-Effect Modelle oder Random Regression Modelle; u. a. Raudenbush & Bryk, 2002) zum Einsatz. Diese eignen sich gut zur Analyse längsschnittlicher Daten, da darin die vorhandene Nestung der Daten (mehrere Erhebungen pro Person in einem Instrument) berücksichtigt wird. Varianzanteile für Schwankungen innerhalb von Personen und der Einfluss systematischer Varianzquellen können genau bestimmt werden. Der individuelle Wert in einem Outcomemaß wird darin beschrieben als Funktion des individuellen Eingangswertes (intercept), der Wachstumskurve (slope) und des Zufallsfehlers. In einem weiteren Schritt können Veränderungsverläufe für die Gesamtstichprobe oder Gruppen innerhalb dieser Stichprobe (etwa komorbide versus reine Diagnose) bestimmt werden. Auch können Daten von Personen einbezogen werden, die nicht zu jedem Messzeitpunkt an der Untersuchung teilgenommen haben.

Markowitz et al. (2007) untersuchten in einer naturalistischen Follow-up-Studie den Einfluss von komorbiden Persönlichkeitsstörungen auf das psychosoziale Funktionsniveau von depressiven Personen über einen Zeitraum von zwei Jahren. Sie verglichen dabei drei Gruppen: Personen mit reiner Depression, Personen mit Depression und komorbider Per-

sönlichkeitsstörung, Personen mit Depression und remittierter Persönlichkeitsstörung. Es resultierten signifikante Effekte für Gruppe und Zeit in allen relevanten Outcomemaßen, wobei die Gruppe mit persistierender Persönlichkeitsstörung am schlechtesten abschnitt. Bedeutsame Gruppe x Zeit-Interaktionen ergaben sich unter anderem für den Global Assessment of Functioning (GAF) –Wert: die größten Verbesserungen wurden innerhalb der ersten sechs Monate erreicht. Der steilste Anstieg war dabei für die Gruppe mit remittierter Persönlichkeitsstörung zu verzeichnen, die zum Ende des Erhebungszeitraumes vergleichbare Werte wie die nicht-komorbide Gruppe erreichte. Das Ergebnis verdeutlicht, wie sehr das Funktionsniveau bei Depression durch das Vorhandensein oder nicht(-mehr)-Vorhandensein von Persönlichkeitsstörungen beeinflusst ist.

Der Einfluss von komorbiden Störungen auf die Veränderung von Patienten mit Generalisierter Angststörung (GAS) bis zum Ende der Therapie und in einem Katamnesezeitraum von zwei Jahren wurde von Newman et al. (2010) überprüft. Es resultierten signifikante Unterschiede in der Verbesserungsrate der GAS-Symptome über die Zeit der Therapie, mit größeren Veränderungsraten für die komorbide Gruppe. Im Gegensatz zum Beginn bestand zum Ende der Behandlung kein bedeutsamer Unterschied in den störungsspezifischen Instrumenten zwischen reiner und komorbider Gruppe. Dies ergab sich sowohl unter Einbezug von Komorbidität als kategoriale Variable als auch unter Berücksichtigung des Schweregrades komorbider Störungen in dimensionaler Form. Komorbidität prädiizierte dagegen keine Unterschiede in Verläufen im Katamnesezeitraum, das heißt erreichte Behandlungseffekte verschlechterten sich zwar etwas im Zwei-Jahres-Zeitraum, blieben aber für beide Gruppen vergleichbar.

Depression als mögliche Komplikation bei Panikstörung und sozialer Phobie wurde von Joormann et al. (2005) untersucht. Es fanden regelmäßige Zwischenerhebungen in störungsspezifischen Instrumenten im Verlauf der Therapie (nach Sitzung 4, 8, 16, 24, 32 usw.) statt, wovon die ersten fünf Erhebungszeitpunkte in hierarchische Regressionsanalysen gingen. Typisches Ergebnis in den verschiedenen Maßen bei Panikstörung waren signifikante Haupteffekte für Zeit und Gruppe. Jedoch wurden Interaktionseffekte Gruppe mal Zeit für panikspezifische Ergebnismaße nicht statistisch bedeutsam, so dass die durchschnittlichen Verbesserungsraten zwischen den Gruppen vergleichbar waren. Für die soziale Phobie ergaben sich in den angstspezifischen Maßen keine bedeutsamen Effekte der Gruppe und keine relevanten Interaktionseffekte. Für beide Hauptdiagnosen resultierten

jedoch im BDI bedeutsame Wechselwirkungseffekte, mit einem schnelleren Abfall depressiver Symptomatik in der komorbiden Gruppe. Auch hier bestand daneben ein signifikanter Haupteffekt für die Gruppe.

Zusammenfassend zu diesen Verlaufsanalysen lässt sich festhalten, dass sich durchgängig ein Haupteffekt der Zeit ergibt, das heißt, es werden Besserungen im Verlauf der Behandlung erreicht, unabhängig von den Diagnosekombinationen. Daneben resultiert in vielen Untersuchungen ein Haupteffekt der Gruppe, das heißt, fast durchgängig stellen sich Personen mit Mehrfachdiagnosen sowohl zu Beginn als auch im Verlauf einer Behandlung belasteter dar als Personen mit einer Diagnose. Für das Vorhandensein von Interaktionseffekten ergeben sich keine eindeutigen Befunde. Häufig erreichen diese Effekte keine statistische Bedeutsamkeit; in einer grafischen Ergebnisdarstellung schlägt sich dies in parallelen Verbesserungsverläufen für komorbide und nicht komorbide Gruppen nieder. Werden jedoch Interaktionseffekte gefunden, so weisen sie für komorbide Patienten mal in eine negative (Markowitz et al., 2007), mal in eine positive Richtung (Newman et al., 2010). Nach den Ergebnissen von Joormann und Kollegen (2005) ist es denkbar, dass das Finden von Interaktionseffekten abhängig ist vom eingesetzten Ergebnismaß und dass es variiert in Abhängigkeit davon, ob die definierte Haupt- oder Nebensymptomatik in störungsspezifischen Instrumenten erfasst wird.

### **2.6.2 Komorbidität und Therapieergebnis beziehungsweise -dauer**

#### *Allgemeine, interpersonelle und störungsspezifische Belastung*

Zu parallel diagnostizierter Persönlichkeitsstörung als prominente Form von Komorbidität liegen einige Übersichtsarbeiten vor, die sich mit deren Einfluss auf das Therapieergebnis auseinandersetzen (Mulder, 2002; Perry, Banon, & Ianni, 1999; Reich, 2003; Shea, Widiger, et al., 1992). Die Mehrzahl kommt dabei zu negativen Schlussfolgerungen. So fassen etwa Reich und Green (1991) nach einer Analyse von naturalistischen und kontrollierten Studien an Patienten mit Depression, Panikstörung oder Zwangsstörung zusammen: „In practically all studies, it is clear that personality pathology is a negative predictor of outcome“ (S. 80). In weiteren Arbeiten werden negative Einflüsse maladaptiver Persönlichkeitszüge oder Persönlichkeitsstörungen auf das Therapieergebnis bei Angst und Depression berichtet (Dreessen et al., 1998; Newton-Howes, Tyrer, & Johnson, 2006; Reich, 2003;

Shea, Widiger, et al., 1992). Ein etwas positiveres Bild zeichnet Mulder (2002), der in seiner Zusammenschau die eingegangenen Studien nach Art der Diagnosemethoden gewichtet hat. Werden vor allem diejenigen Studien betrachtet, in denen Persönlichkeitsstörungen mittels strukturierter Interviews erfasst werden und in denen kontrollierte Behandlungsbedingungen realisiert sind, so schneiden Patienten mit komorbider Persönlichkeitsstörung kaum schlechter ab als Patienten ohne diese Komplikation.

Bezüglich der Ergebnisbereiche allgemeine, interpersonale und störungsspezifische Belastung werden teils divergierende Ergebnisse berichtet, in Abhängigkeit davon, ob der Statuswert im entsprechenden Instrument zum Ende der Therapie als Ergebnismaß herangezogen wird, oder aber ob Differenzwerte zwischen Eingangs- und Endwert berichtet werden. Bezüglich der subjektiv empfundenen Beeinträchtigung durch körperliche und psychische Symptome, erhoben über den SCL-90 (Derogatis, 1992), bestehen uneinheitliche Ergebnisse. Unter Betrachtung des Gesamtwertes allein zu Therapieende schneiden Personen mit komorbiden Störungen schlechter ab als Personen mit nur einer Störung (Viinamäki et al., 2002; Wonderlich, Fullerton, Swift, & Klein, 1994). Werden Differenzwerte im Vergleich zum Therapiebeginn herangezogen, so besteht in der Untersuchung von Wonderlich et al. (1994) kein Unterschied zwischen essgestörten Personen mit und ohne komorbide Persönlichkeitsstörung. Für Persönlichkeitsstörung bei komorbider Angststörung wird sogar signifikant mehr Verbesserung im SCL gegenüber reinen Angststörungen berichtet (Dreessen et al., 1994). Demgegenüber steht ein schlechteres Abschneiden bei Achse-II-Komorbidität im Vergleich zu reiner Depression unter Heranziehung der Prä-Post-Effektgrößen in SCL-90-R und im Inventar Interpersoneller Probleme (IIP) (Hardy et al., 1995).

Bezüglich der Veränderung in den mit der Hauptdiagnose assoziierten Symptommaßen weisen die Ergebnisse ebenfalls in unterschiedliche Richtungen. Bei alleiniger Betrachtung des Status zum Therapieende schneiden komorbide Patienten auch hier schlechter ab (Noyes et al., 1990; Viinamäki et al., 2002; Wonderlich et al., 1994). Wird das Therapieergebnis über Differenzmaße definiert, so findet man für die GAS bei Vorliegen weiterer Diagnosen sogar bedeutsam höhere Verbesserungsraten als bei Vorliegen nur einer Störung (Newman et al., 2010; Wetherell et al., 2005). Für Panikstörung (Tsao et al., 1998) und Agoraphobie (Chambless et al., 2000), genauso wie für Zwangsstörungen (Abramowitz et al., 2000; Gershuny, Baer, Jenike, Minichiello, & Wilhelm, 2002) ist das Bild um-

gekehrt: die Besserungsraten sind hier höher für Gruppen ohne komorbide Störungen. Für Depression fanden Hardy et al. (1995) divergierende Ergebnisraten bei Komorbidität in Abhängigkeit vom Therapieverfahren. In zahlreichen Untersuchungen wurden statt einfacher Differenzwertberechnung auch messwiederholte (Ko-)Varianzanalysen durchgeführt (Diguer, Barber, & Luborsky, 1993; Dreessen et al., 1998; Joormann et al., 2005; Rief et al., 2000; Shea et al., 1990; Wonderlich et al., 1994). Typisches Ergebnis dabei sind signifikante Haupteffekte sowohl für den Faktor Gruppe (komorbide vs. nicht komorbide) als auch für den Faktor Zeit (Prä- vs. Postmessung), jedoch kein signifikanter Interaktionseffekt Gruppe x Zeit. Danach sind Patienten mit komorbider Störung zwar zur Eingangsmessung belasteter, verbessern sich jedoch im Mittel bis zum Therapieende im gleichen Maße wie Patienten mit nur einer Diagnose.

### *Symptomfreiheit/Genesung*

Die häufige Beobachtung erhöhter Eingangsbelastung und vergleichbarer Verbesserungsraten in den Gruppen komorbider vs. nicht komorbider Patienten hat bei gleicher Therapiedauer zwangsläufig auch Auswirkungen auf die Rate symptomfreier beziehungsweise genesener Patienten zum Therapieende. Freiheit von depressiver Symptomatik beziehungsweise das Ausbleiben erneuter depressiver Episoden werden bei gleichzeitig vorliegender Persönlichkeitsstörung weniger häufig erreicht (Markowitz et al., 2007; Shea et al., 1990; Viinamäki et al., 2002). Gleiches gilt bei Vorliegen einer komorbiden Angstsymptomatik sowohl bei unipolaren (Sherbourne & Wells, 1997) als auch bei bipolaren affektiven Störungen (Feske et al., 2000). Für die Genesungsrate bei der GAS liegen widersprüchliche Befunde vor: während Durham und Kollegen (1997) von einem negativen Einfluss komorbider Achse-I-Störungen gegenüber einer Nicht-Beeinflussung durch Achse-II-Störungen berichten, fanden Yonkers, Dyck, Warshaw und Keller (2000) genau entgegengesetzte Effekte. Für andere Angststörungen konnte kein nachteiliger Effekt auf die Remissionsraten durch Komorbidität gefunden werden (Brown et al., 1995; Dreessen et al., 1998).

### *Funktionsniveau*

Für die allgemeine psychische und soziale Funktionsfähigkeit zum Ende einer Behandlung ergeben sich für Patienten mit komorbider Persönlichkeitsstörung bei Depression durchgehend schlechtere Ergebnisse im Vergleich zu nicht-komorbiden Vergleichsgruppen

(Andreoli et al., 1993; Markowitz et al., 2007; Shea et al., 1990). Lediglich im Bereich beruflichen Funktionierens wurden keine Unterschiede gefunden (Shea et al., 1990). Für Patienten mit Zwang oder Panikstörung hatten gleichzeitig bestehende weitere psychische Störungen keinen Einfluss auf das Funktionsniveau zum Therapieende (Steketee et al., 2001).

### *Behandlungsabbrüche*

Zum Einfluss von Komorbidität auf das frühzeitige Ausscheiden aus Behandlungen sind verschiedene Wirkrichtungen denkbar. Durch die erhöhte Eingangsbelastung unterliegen komorbide Patienten möglicherweise einem größeren Veränderungsdruck und weisen eine höhere Therapiemotivation auf, welche sich positiv auf das Verbleiben in der Therapie auswirken kann. Demgegenüber erscheint es ebenso plausibel, dass sich die Behandlung von Patienten mit mehreren Störungen als schwieriger und komplikationsreicher darstellt. Insbesondere komorbide Persönlichkeitsstörungen und damit verbundene interpersonelle Schwierigkeiten, die sich auch in der Therapie manifestieren können, haben einen möglicherweise negativen Einfluss auf das Verbleiben in der Behandlung. So gibt etwa Reich (2003) in einem Übersichtsartikel erhöhte dropout-Raten bei Persönlichkeitsstörungen (komorbid zu affektiven oder Angststörungen) an und diskutiert, dass das oft beobachtete im Mittel schlechtere Abschneiden bei Achse-II-Störungen zum Teil auf die erhöhten Abbrecherraten zurückzuführen sein könnte. Dagegen konnten Dreessen et al. (1998) auch unter Einbezug der Abbrecher keinen Einfluss von Persönlichkeitsstörungen auf das Therapieergebnis bei Zwangsstörungen ausmachen. In Originalarbeiten zur Komorbidität von Persönlichkeitsstörungen bei Depression (Shea et al., 1990) und Angststörungen (Dreessen et al., 1994; Steketee et al., 2001) bestanden keine Unterschiede im Hinblick auf verfrühte Abbrüche. Auch zusätzliche Diagnosen auf Achse I sind meist nicht prädiktiv für drop-out bei Angststörungen (Brown et al., 1995; Steketee et al., 2001). Insbesondere bei komorbider GAS ergaben sich widersprüchliche Ergebnisse bezüglich der Abbruchraten (Newman et al., 2010; Steketee et al., 2001).

### *Vorhandensein weiterer Störungen*

Einige Untersuchungen ziehen die zum Therapieende bestehenden Störungsdiagnosen als Ergebniskriterium heran. Auch hier stellt sich die Frage nach dem Einfluss komorbider

Störungen. Es finden sich zwei Studien, die den Einfluss der Behandlung von GAS auf gleichzeitig bestehende weitere Störungen aus dem Bereich Angst- und affektiven Störungen prüfen (Borkovec, Abel, et al., 1995; Newman et al., 2010) sowie eine weitere Studie, die sich auf die Panikstörung als Indexdiagnose bezieht (Brown et al., 1995). Alle stellen einen deutlichen Rückgang komorbider Diagnosen nach einer Behandlung der Hauptdiagnose fest (Prä-Post), der im Katamnesezeitraum jedoch wieder leicht rückläufig ist (Brown et al., 1995; Newman et al., 2010). Insgesamt scheint die Behandlung von Hauptdiagnosen auf parallel bestehende weitere Problembereiche und komorbide Diagnosen zumindest in Teilen zu generalisieren.

#### *Vorhersage des Therapieergebnisses anhand von Komorbidität*

Kann Information über Komorbidität genutzt werden, um frühzeitig Erfolg beziehungsweise Misserfolg einer Therapie vorherzusagen? Bleiben die Zusammenhänge zum Therapieergebnis im Kontext anderer Variablen bestehen? Um diesen Fragen nachzugehen, wurden in einigen Studien Regressionsanalysen durchgeführt, die Komorbidität zum Therapiebeginn als Prädiktorvariable behandeln. Yonkers und Kollegen (2000) prädizierten Voll- beziehungsweise Teilremission für Personen mit GAS über einen Zeitraum von fünf Jahren. Neben hoher Eingangsbelastung in Maßen der Lebensqualität stellten sich komorbide Persönlichkeitsstörungen als signifikant negativer Prädiktor für eine Remission heraus. Dagegen ohne Einfluss blieben komorbide Diagnosen aus den Bereichen affektive Störungen, Posttraumatische Belastungsstörung, Zwang, Soziale Phobie und Panikstörung.

In einer weiteren Studie an Patienten mit GAS wurden anhaltende Verbesserung beziehungsweise Rückfälle im Rahmen logistischer Regressionen untersucht (Durham et al., 1997). Das Ergebnismaß wurde hier zusammengesetzt aus einem allgemeinen und einem störungsspezifischem Instrument, sowie der selbst eingeschätzten Gesamtverbesserung. Über soziodemografische und sozioökonomische Variablen sowie die Variablen Behandlungsmethode und Allianz hinaus resultierte die Achse-I-Komorbidität als bedeutsamer Prädiktor: Personen mit zusätzlichen Diagnosen auf Achse I und erhöhten Psychotizismuswerten im BSI erlitten mit größerer Wahrscheinlichkeit Rückfälle bis zu 12 Monate nach Beendigung der Behandlung.

Der Einfluss verschiedener Achse-I- und Achse-II-Diagnosen auf symptombezogene Ergebnismaße bei Agoraphobie wurde von Chambless et al. (2000) geprüft. Als Kriterium

dienten die Post- und Follow-up-Werte (6 Monate) in Vermeidungsverhalten und die Häufigkeit von Panikattacken. Als Prädiktoren gingen der jeweilige Eingangswert als Kontrollvariable und das Vorhandensein weiterer Diagnosen (dichotomisiert) ein. Für die Prädiktion des Vermeidungsverhaltens erreichte neben dem Eingangswert nur das gleichzeitige Vorhandensein einer Depression Signifikanz, für die Häufigkeit von Panikattacken waren Komorbidität mit Dysthymia beziehungsweise vermeidender Persönlichkeitsstörung und sozialer Phobie prädiktiv bedeutsam.

#### *Behandlungsdauer / Anzahl Therapiesitzungen*

Der Einfluss von Komorbidität auf die Behandlungsdauer kann nur über naturalistische Therapiestudien oder aber, in indirekter Form, über kontrollierte Studien mit im Design umgesetzter Variation in der Anzahl durchzuführender Sitzungen untersucht werden. In einer naturalistischen Studie an Angstpatienten wurde kein Unterschied in der Anzahl der Sitzungen abhängig vom Vorliegen einer komorbiden Persönlichkeitsstörung gefunden (Dreessen et al., 1994). Eine vergleichbare Anzahl an Therapiesitzungen sowie eine vergleichbare Behandlungsdauer für komorbide und nicht komorbide Patienten fanden auch Reardon, Cukrowicz, Reeves und Joiner (2002). Dagegen benötigten in einer Untersuchung von Pilkonis und Frank (1988) unter standardisierten Bedingungen Patienten länger bis zu einer Remission, wenn zusätzlich zur rezidivierenden depressiven Störung noch eine Persönlichkeitsstörung vorlag. Auch Joormann et al. (2005) berichten, in diesem Fall für komorbide Depression bei Angststörungen, von einer Tendenz zum längeren Verbleib in der Therapie. Ausgehend von der Beobachtung paralleler Verbesserungsverläufe wurde insbesondere für Persönlichkeitsstörungen die Vermutung angestellt, dass komorbide Patienten bei einer erhöhten Sitzungsanzahl beziehungsweise verlängerter Therapiedauer genauso profitieren könnten wie solche mit nur einer Diagnose (Durham et al., 1997; Perry et al., 1999). Dies würde für eine differenzielle Therapieplanung in Abhängigkeit von den spezifischen Bedürfnissen des Patienten, einschließlich der Anzahl gestellter Diagnosen, sprechen (u. a. Barkham et al., 2006).

### 3 Fragestellungen und Hypothesen

Aus den in den vorangehenden Abschnitten dargestellten theoretischen Hintergründen werden im Folgenden die in dieser Arbeit untersuchten Fragestellungen, gegliedert nach vier Untersuchungsbereichen, abgeleitet.

#### 3.1 Veränderung des Diagnoseverhaltens

In Analogie zu den zitierten Unterschieden zwischen strukturierten Interviews und unstrukturierter Diagnostik (Kapitel 2.3) wird angenommen, dass in sonst vergleichbaren Patientenstichproben unter Verwendung von Diagnosechecklisten mehr Störungen diagnostiziert werden als wenn keine Vorgaben hinsichtlich der Diagnostik gemacht werden. Konkrete Hinweise aus der Literatur ergeben sich für eine erhöhte Diagnosewahrscheinlichkeit für Persönlichkeitsstörungen und Angststörungen, sowie für eine verringerte Diagnosewahrscheinlichkeit für Anpassungsstörungen unter Verwendung von strukturierten Diagnoseverfahren.

*Hypothese 1:* In einer Stichprobe ambulanter Psychotherapiepatienten werden im Vergleich zu einer Stichprobe mit unstandardisiertem diagnostischem Vorgehen unter Verwendung der Internationale Diagnosen Checklisten (IDCL)

- a) mehr Diagnosen pro Patient vergeben,
- b) mehr Diagnosen aus der Gruppe der Persönlichkeitsstörungen und Angststörungen vergeben und,
- c) weniger Diagnosen aus dem Bereich der Anpassungsstörungen vergeben.

#### 3.2 Art und Muster der Komorbidität

Nach den vorgestellten Untersuchungen zum gemeinsamen Auftreten psychischer Störungen (Kapitel 2.4) wird erwartet, dass auch unter Berücksichtigung der Basisraten im Rahmen einer bivariaten Betrachtungsweise einige Störungsgruppen überzufällig oft gemeinsam diagnostiziert werden.

*Hypothese 2:* Im Rahmen einer bivariaten Betrachtungsweise von Komorbidität ergeben sich Störungsgruppen, welche häufiger gemeinsam diagnostiziert werden, als unter Annahme der Gleichverteilung zu erwarten.

Daneben stellt sich die Frage, ob über bivariate Zusammenhänge hinaus auch größere Gruppen von Diagnosen, also distinkte Typen von Diagnosekombinationen (sog. Latente Klassen), die mit einer erhöhten Wahrscheinlichkeit gemeinsam auftreten, zu identifizieren sind. Möglicherweise stehen solche Typen von Störungskombinationen in Zusammenhang mit der Belastungsschwere.

*Hypothese 3:* Im Rahmen einer multivariaten Betrachtung lassen sich latente Klassen von Patienten auf der Grundlage ihrer Diagnosekombinationen identifizieren.

### **3.3 Konsequenzen der Komorbidität für den Status zu Beginn der Therapie**

Abgeleitet aus der im Abschnitt 2.5 zitierten Literatur wird erwartet, dass Patienten mit Mehrfachdiagnosen im Mittel höhere psychische Belastungswerte aufweisen als Patienten mit nur einer Diagnose, und dass dies der Fall ist sowohl für Belastungseinschätzungen aus der Selbst- als auch aus der Fremdbeurteilerperspektive. Daneben soll explorativ für die vorab identifizierten latenten Klassen an Diagnosekombinationen überprüft werden, ob diese in Zusammenhang mit Maßen der selbst- und fremdeingeschätzten Belastung stehen.

*Hypothese 4:* Zu Beginn der Therapie bestehen bedeutsame Unterschiede zwischen Personen mit nur einer Diagnose gegenüber Personen mit Komorbidität

- a) hinsichtlich der selbstberichteten Belastungsschwere und
- b) hinsichtlich der durch den Therapeuten eingeschätzten allgemeinen Belastungsschwere.

### **3.4 Konsequenzen der Komorbidität für Therapieverlauf und Therapieergebnis**

Aus den zitierten Arbeiten zur Komorbidität und Therapieverläufen (Kapitel 2.6.1) bleibt offen, wie sich Komorbidität im Verlauf eines allgemeinen Ergebnismaßes darstellt, wel-

ches nicht direkt auf Haupt- oder Nebendiagnosen bezogen ist. Dabei wird zum einen die allgemeine psychische Belastung, erfasst über das BSI (Franke, 2000), zum anderen die interpersonelle Belastung, erfasst über das IIP (Horowitz et al., 2000), als Outcome herangezogen. Nach den dargestellten Befunden sind auch hier Haupteffekte von Zeit und (Komorbiditäts-)Gruppe zu erwarten. Daneben soll untersucht werden, ob und wenn ja, in welcher Richtung sich Unterschiede in den Verbesserungsraten in Abhängigkeit des Vorliegens komorbider Diagnosen sowie in Abhängigkeit der Eingangsbelastung im jeweiligen Maß ergeben. Die genannten Fragestellungen werden in den Hypothesen 5 bis 7 zusammengefasst.

*Hypothese 5:* Im Verlauf der Behandlung wird, unabhängig vom Vorhandensein komorbider Diagnosen, eine bedeutsame Besserung in Maßen selbstberichteter Belastung erreicht, und zwar:

- a) in allgemeiner psychischer Belastung.
- b) in interpersonaler Belastung.

*Hypothese 6:* Patienten unterscheiden sich hinsichtlich ihrer Veränderungsrate in selbsterichteter

- a) allgemeiner psychischer Belastung und
- b) interpersonaler Belastung.

Diese Unterschiede können durch Komorbidität vorhersagesagt werden.

*Hypothese 7:* Es besteht ein Zusammenhang zwischen dem Ausgangswert und der Veränderungsrate

- a) in allgemeiner psychischer Belastung,
- b) in interpersonaler Belastung.

In einigen der dargestellten Studien sowie in Hypothese 6 wurde Komorbidität als potenziell differenzierender Faktor in die Analyse von Verlaufsfunktionen aufgenommen. Möglicherweise wird durch diese a-priori-Festlegung von Gruppen aber wichtige Verände-

rungsinformation außer Acht gelassen. In den vergangenen Jahren wurden zahlreiche Untersuchungen durchgeführt, in denen es gelang, so genannte latente Gruppen der Veränderung in Psychotherapie-Verlaufsdaten zu identifizieren, darunter immer wieder eine Gruppe von Patienten, die sehr früh im Therapieverlauf eine deutliche Besserung aufweisen (s. a. Abschnitt 2.1.2.2; Cuijpers, van Lier, van Straten, & Donker, 2005; Lutz, Stulz, et al., 2009; Stulz & Lutz, 2007; Stulz, Lutz, et al., 2007). Eine bislang ungeklärte Frage ist, ob die Zugehörigkeit zu latenten Veränderungsgruppen und insbesondere zu der Gruppe schnell verbesserter Patienten mit dem Vorliegen von Ein- beziehungsweise Mehrfachdiagnosen assoziiert ist. Dementsprechend soll auch im vorliegenden Datensatz der Versuch der Identifikation latenter Veränderungsgruppen in allgemeinen Ergebnismaßen unternommen werden, welche dann in Zusammenhang mit Komorbidität gebracht werden. Wenn so geartete distinkte Veränderungsgruppen gefunden werden können, schließt sich die Frage an, inwieweit die Information über die Zugehörigkeit zu diesen Mustern und darüber hinaus über Komorbidität zur Vorhersage des Therapieergebnisses und -aufwandes geeignet ist.

*Hypothese 8:*

1. Es lassen sich bezüglich
  - a) der allgemeinen psychischen Belastung und
  - b) der interpersonalen Belastungnicht direkt beobachtbare Gruppen von Patienten identifizieren, die charakterisiert sind durch ihren Eingangswert und ihre Veränderungsrate in den jeweiligen Instrumenten.
2. Diese latenten Gruppen von Patienten stehen in Zusammenhang mit der Anzahl komorbider Störungen beziehungsweise mit der Diagnose-Kombination.
3. Die Informationen über die Zugehörigkeit zu diesen latenten Gruppen von Patienten und darüber hinaus über die Komorbidität sind nutzbar zur Vorhersage des Therapieergebnisses und der Anzahl an Therapiesitzungen.

Studien zum Zusammenhang zwischen Komorbidität und Therapieergebnis weisen in unterschiedliche Richtungen (Kapitel 2.6.2). Während einige Studien vergleichbare Verbes-

serungsraten zwischen komorbiden und nicht komorbiden Patienten finden, stellen andere durchaus einen Zusammenhang zwischen Komorbidität und auch dem am Ausgangswert relativierten Therapieergebnis fest, wobei jedoch die Richtung des Zusammenhangs variiert<sup>4</sup>. Daneben hängt möglicherweise die Stärke des Zusammenhangs davon ab, welche spezifische Diagnose komorbid vorliegt. Die in den genannten Studien eingesetzten Ergebnismaße beziehen sich häufig auf die definierte Hauptsymptomatik. Zu Maßen der allgemeinen körperlichen und psychischen Belastung und insbesondere interpersoneller Schwierigkeiten gibt es hingegen nur wenige und dabei sehr gemischte Befunde. Dementsprechend wird in dieser Arbeit der Frage nachgegangen, ob sich Komorbidität zur Vorhersage des Therapieergebnisses in allgemeiner, interpersonaler und störungsspezifischer Belastung unter Einbeziehung der Ausgangswerte eignet.

*Hypothese 9:*

Komorbidität erweist sich im Kontext der individuellen Eingangsbelastung und der Anzahl durchgeführter Therapiesitzungen als bedeutsamer Prädiktor des Therapieergebnisses zum Abschluss der Behandlung und zur Ein-Jahres-Katamnese

- a) in allgemeiner psychischer Belastung,
- b) in interpersonaler Belastung,
- c) in störungsspezifischer Belastung.

Denkbar ist, dass gerade komorbide Patienten von einer variablen Anpassung der Therapielänge in Abhängigkeit von der Eingangsbelastung und Diagnosekombination profitieren könnten. Jedoch liegen zur Frage des Zusammenhangs zwischen Komorbidität und der Dauer der Behandlung nur wenige Untersuchungen vor, deren Befunde ebenfalls in unterschiedliche Richtungen weisen. Daher soll in dieser Arbeit auch die Prädiktionskraft von Komorbidität für die Anzahl der Therapiesitzungen untersucht werden.

*Hypothese 10:* Komorbidität erweist sich im Kontext der individuellen Eingangsbelastung als bedeutsamer Prädiktor für die Anzahl durchgeführter Therapiesitzungen bis zur Beendigung einer Behandlung.

---

<sup>4</sup> positiver Zusammenhang zwischen Komorbidität und Ergebnis: Dreesen et al. (1998); Newman et al. (2010); Wetherell et al. (2005); negativer Zusammenhang zwischen Komorbidität und Ergebnis: Abramowitz et al. (2000); Chambless et al. (2000); Hardy et al. (1995); Markowitz et al. (2007).

## 4 METHODEN

Im folgenden Kapitel wird zunächst genauer auf das bereits vorgestellte Modellprojekt der Techniker Krankenkasse eingegangen und die in dieser Arbeit verwendete Stichprobe vorgestellt. Daran anschließend wird auf die zur Beantwortung der Fragestellungen angewendeten Methoden eingegangen. Dazu wird unterschieden zwischen Methoden, die der Beantwortung der Status-Fragestellungen (Hypothesen 1 bis 4; Unterkapitel 4.2) dienen und solchen, die der Beantwortung der Verlaufs- und Ergebnisfragestellungen (Hypothesen 5 bis 10; Unterkapitel 4.3) dienen. Alle Datenanalysen erfolgten mit Hilfe der Software PASW Statistics 18.0, G\*Power (Faul, Erdfelder, Lang, & Buchner, 2007) und Mplus 6.1 (Muthén & Muthén, 2010).

### 4.1 Beschreibung der Stichprobe

Wie in den einleitenden Kapiteln angedeutet stammen die für die folgenden Analysen genutzten Daten aus dem Modellprojekt „Qualitätsmonitoring in der ambulanten Psychotherapie“, welches im Jahr 2005 durch die Techniker Krankenkasse ins Leben gerufen und dessen Datenerhebung Ende Juni 2010 abgeschlossen wurde. Im Folgenden sollen Design, verwendete Messinstrumente und resultierende Stichprobe des Projektes, sowie die sich daraus für diese Arbeit ergebende Stichprobe genauer vorgestellt werden.

#### 4.1.1 Modellprojekt Psychotherapie der TK - Hintergrund und Ziele

Ausschlaggebend für die Planung dieser Studie zum Qualitätsmonitoring in der ambulanten Psychotherapie waren – neben der in den vergangenen Jahren allgemein gestiegenen Bedeutung von und auch der gesetzlich umgesetzten Forderung nach Qualitätssicherungsmaßnahmen im deutschen Gesundheitssystem (z. B. Kastenholz & Both, 2002) – zwei für Psychotherapien in Deutschland spezifische Punkte: Zum einen bestehen nach wie vor keine einheitlichen Prinzipien zur Dokumentation und Evaluation von Psychotherapien im ambulanten Setting, und damit nur unzureichendes Wissen über deren Verlaufs- und Ergebnisqualität (Lutz, Köck, & Böhnke, 2009). Zum anderen beruht das System zur Beantwortung und Verlängerung ambulanter Psychotherapien in Deutschland derzeit auf einem

rein expertenbezogenen Gutachterverfahren, welches die Patientenperspektive nicht einbezieht. Ziel des Modellvorhabens war es daher, zum einen ein alternatives Gutachterverfahren zu erproben, welches explizit Patientenangaben berücksichtigt und sich ein Stück von anonymen Expertenurteilen löst. Zum anderen sollten Therapieergebnisse systematisch dokumentiert und auch regelmäßig im Verlauf erhoben und dem Therapeuten beziehungsweise Patienten zurückgemeldet werden, um den Einfluss dieser Intervention auf die Ergebnisqualität zu untersuchen.

#### **4.1.2 Instrumente und Messplan**

In der IG erfolgte eine strukturierte Eingangsdiagnostik; daneben wurden hier mehrfach im Verlauf der Behandlung Selbstauskünfte der Patienten hinsichtlich verschiedener Zielbereiche erfasst, welche in einer Rückmeldung an den Therapeuten zusammengefasst wurden. Außerdem wurde hier das Antrags-, Verlängerungs- und Gutachterverfahren um die Sicht des Patienten hinsichtlich der allgemeinen Symptombelastung, der störungsspezifischen Belastung, interpersonaler Beziehungen und der Therapiebeziehung ergänzt. In der KG wurde auf Maßgaben zur kategorialen Diagnostik bei Therapiebeginn verzichtet. Außerdem erfolgten hier psychometrische Erhebungen lediglich zu Beginn und Ende der Behandlung und zur 12-Monats-Katamnese.

##### **4.1.2.1 Verfahren zur kategorialen Eingangsdiagnostik**

Zu Beginn der Behandlung wurden pro Patient bis zu vier Diagnosen durch den behandelnden Therapeuten gestellt. Zum Zweck der Diagnosestellung waren die Therapeuten der IG angehalten, bestimmte teilstrukturierte diagnostische Verfahren anzuwenden, um eine klinisch valide und reliable Diagnostik zu fördern. Hier kam zunächst ein Screening der Symptomstörungen in Anlehnung an das Strukturierte Klinische Interview für DSM-IV, Achse I (SKID-I; Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997) zum Einsatz. Hauptbestandteil des diagnostischen Prozesses bildete die Diagnosestellung für Symptomstörungen auf der Basis der Internationalen Diagnose Checklisten nach ICD-10 (IDCL; Hiller et al., 2004). Darin werden die Diagnosekriterien der psychischen und Verhaltensstörungen nach ICD-10 (Weltgesundheitsorganisation, 2005) aufgeführt und können vom Therapeuten in strukturierter Form anhand von Entscheidungsbäumen abgeklärt werden.

Ein Screening und die Diagnostik von Persönlichkeitsstörungen erfolgte über die Internationalen Diagnose Checklisten nach ICD-10 für Persönlichkeitsstörungen (IDCL-P; Bronisch et al., 1995). Therapeuten der KG erhielten keinerlei Vorgaben hinsichtlich ihres Vorgehens in der kategorialen Diagnostik (s. a. Kap. 4.2.1).

#### **4.1.2.2 Psychometrische Instrumente**

Die Auswahl der Messinstrumente orientierte sich an dem Ziel, zum einen Informationen zu unterschiedlichen Bereichen psychischen Funktionierens zu erhalten (multidimensionale Erfassung), zum anderen sowohl Patienten- als auch Therapeutenperspektive mit einzu beziehen (multimodale Erfassung). In der IG fanden Therapeutenratings zu Beginn der Therapie, bei einer eventuellen Verlängerung, sowie zum Ende der Behandlung statt. Hier wurden Einschätzungen zur allgemeinen Belastung (BSS; Schepank, 1995), zu Problemrelevanten Lebensbereichen, zur therapeutischen Allianz (HAQ; Bassler, Potratz, & Krauthauser, 1995) und zum allgemeinen Funktionsniveau des Patienten (GAF; Endicott, Spitzer, Fleiss, & Cohen, 1976) vorgenommen. In der KG fanden diese Erhebungen lediglich zum Therapiebeginn und -abschluss statt. Selbstberichte der Patienten wurden in fünf verschiedenen Maßen erhoben. Es wurden Urteile zur therapeutischen Allianz (HAQ; Bassler et al., 1995), zur gesundheitsbezogenen Lebensqualität (SF-12; Bullinger & Kirchberger, 1998), zu interpersonalen Problemen (IIP-D; Horowitz et al., 2000) und der allgemeinen Symptombelastung (BSI; Franke, 2000) erfasst. Zur Dokumentation störungsspezifischer Probleme kam eines von fünf weiteren Instrumenten zum Einsatz, abhängig davon, in welchem Bereich der Therapeut die Hauptbelastung des Patienten lokalisierte (AKV, Ehlers & Margraf, 2001; BDI, Hautzinger, Bailer, Worall, & Keller, 1995; HZI-K, Klepsch, Zaworka, Hand, Lünenschloss, & Jauering, 1993; EDI, Paul & Thiel, 2005; SOMS, Rief, Hiller, & Heuser, 1997). Diese Selbstauskünfte wurden für alle Patienten zu Therapiebeginn und -abschluss sowie zur 1-Jahres-Katamnese erhoben. In der IG fanden darüber hinaus regelmäßige Zwischenerhebungen im Verlauf der Therapie statt, anhand derer Fortschrittsrückmeldungen an den Therapeuten generiert wurden. Da im Modellvorhaben alle drei in Deutschland durch die Krankenkassen zugelassenen Psychotherapieverfahren vertreten sind, erfolgte eine Anpassung der Erhebungsintervalle für die Zwischenmessungen in Abhängigkeit von den jeweiligen Bewilligungszeiträumen. Eine schematische Übersicht zu den Messzeitpunkten der Patienten-Selbstauskünfte findet sich in Abbildung 1.

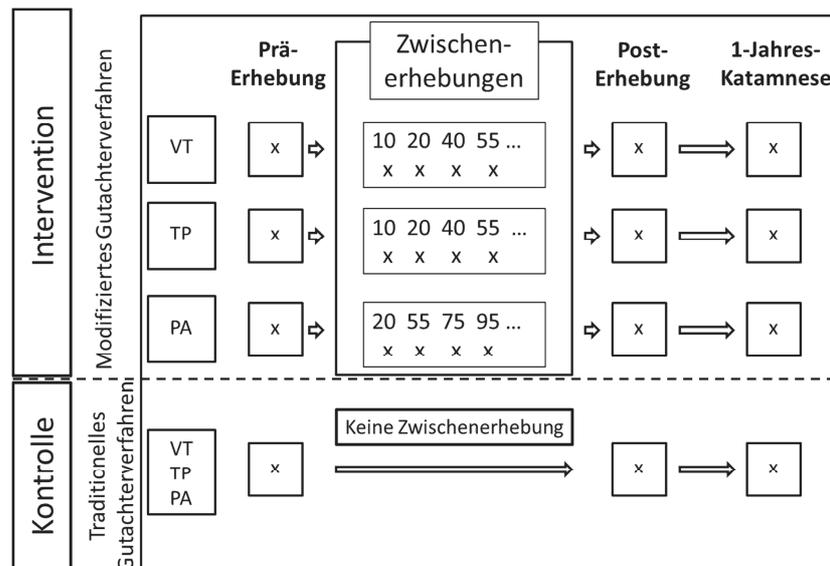


Abbildung 1: Schematische Übersicht über die Messzeitpunkte (Selbstberichte) im TK-Modellvorhaben; Ziffern geben die Sitzungsnummer nach den probatorischen Sitzungen an, zu denen routinemäßig erhoben wird. VT=Verhaltenstherapie; TP=Tiefenpsychologisch fundierte Psychotherapie; PA=Analytische Psychotherapie (aus: Lutz, Böhnke, Köck, & Bittermann, 2011).

Im Rahmen des Modellprojektes kam ein rationales Entscheidungssystem (Lutz & Stulz, 2008) zum Einsatz. Grundlegend für die Beurteilung des Eingangsstatus und Fortschrittes waren dabei drei zentrale psychometrische Instrumente: BSI (Franke, 2000), IIP (Horowitz et al., 2000) und das jeweilige störungsspezifische Instrument, wobei am häufigsten der BDI (Hautzinger et al., 1995) angewendet wurde. Zu Therapiebeginn kam bei Patienten der IG ein alternatives Gutachterverfahren zum Einsatz: sofern ein Patient mindestens eine ICD-10-Diagnose erhielt und in mindestens einem der eingehenden Messinstrumente (BSI, IIP und störungsspezifisches Instrument) klinisch bedeutsame Werte (angelehnt an das Konzept von Jacobson & Truax, 1991) aufwies, wurde die Therapie ohne Einschaltung eines externen Gutachters bewilligt. Eine zweite Art von Entscheidungsregeln bezog sich auf die Rückmeldungen im Therapieverlauf. Hier wurden die gleichen Selbstbericht-Fragebögen wie zu Therapiebeginn verwendet. Für jedes Instrument wurde zu jedem Messzeitpunkt die klinisch bedeutsame Veränderung in Bezug auf den Eingangswert (nach dem Konzept von Jacobson & Truax, 1991) berechnet und der Veränderungsverlauf in einer grafischen Rückmeldung dargestellt. Da BSI, IIP und BDI zentrale psychometrische Instrumente im Rahmen des Modellprojektes darstellen und auch als abhängige Variablen in dieser Arbeit dienen, werden diese in der Folge genauer vorgestellt.

Das Brief Symptom Inventory (BSI, Franke, 2000) stellt eine Kurzform der Symptom Checkliste (Derogatis, 1992; in der deutschen Version Franke, 1995) dar. Mit diesem Instrument lässt sich die subjektive Beeinträchtigung durch körperliche und psychische Symptome erfassen. Das Instrument besteht aus 53 Items, die zu neun Skalen und drei globalen Kennwerten zusammengefasst werden: GSI (Global Severity Index), PSDI (Positive Symptom Distress Index) und der PST (Positive Symptom Total). Im Rahmen dieser Arbeit wird auf den GSI zurückgegriffen, welcher Auskunft über die allgemeine psychische Belastung gibt. Höhere Werte stehen dabei für höhere Belastung. Untersuchungen zur Reliabilität zeigen interne Konsistenzen der Skalen zwischen  $r=.71$  und  $r=.85$  für ambulante Psychotherapiepatienten, für den GSI liegt die Retest-Reliabilität bei  $r=.90$  (Franke, 2000).

Das Inventar zur Erfassung interpersonalen Probleme (IIP-D; Horowitz et al., 2000) dient der Selbsteinschätzung interpersonalen Schwierigkeiten. Es fragt zum einen nach Verhaltensweisen, die der befragten Person schwer fallen, zum anderen nach Verhaltensweisen, die die Person im Übermaß zeigt. Der Fragebogen wurde im Rahmen des TK-Modells in seiner Kurzform (64 Items) eingesetzt. Das IIP-D besteht aus acht theoretisch verankerten, faktorenanalytisch gewonnenen Skalen, die den Oktanten des interpersonalen Kreismodells (Leary, 1957) entsprechen. Daneben kann ein Gesamtwert gebildet werden, auf welchen im Rahmen dieser Arbeit zurückgegriffen wird. Je höher der Gesamtwert, desto größere Schwierigkeiten im Bereich interpersonalen Probleme liegen vor. Die retest-Reliabilitäten liegen zwischen  $r=.81$  und  $r=.90$  (Horowitz et al., 2000).

Das Beck Depressions Inventar (BDI, Hautzinger et al., 1995) ist ein Selbstbeurteilungsinstrument zur Erfassung der depressiven Symptomatik und umfasst 21 Gruppen von Aussagen, die typische depressive Symptome beschreiben. Die Summenwerte liegen zwischen 0 und 63 Punkten. Die interne Konsistenz liegt zwischen  $r=.89$  und  $r=.94$ , die Retest-Reliabilität bei  $r=.93$ . Sowohl SCL-90 beziehungsweise BSI, als auch IIP und BDI sind international verbreitet, gelten als änderungssensitiv und werden daher häufig für die Verlaufs- und Ergebnismessung eingesetzt.

#### **4.1.3 Teilnehmer – Therapeuten und Patienten im TK-Projekt**

Das Projekt erfasst ambulante Psychotherapiepatienten, die bei niedergelassenen Psychotherapeuten behandelt wurden. In den drei Modellregionen der Kassenärztlichen Vereini-

gungen Hessen, Südbaden und Westfalen-Lippe wurden insgesamt 403 Therapeuten in eigener Praxis in allen als Kassenleistung anerkannten Verfahren zur Teilnahme rekrutiert. Mehr als die Hälfte der Therapeuten (213; 52.9%) gibt an, eine verhaltenstherapeutische Ausbildung absolviert zu haben, weitere 13 Therapeuten (3.2%) besitzen zusätzlich noch einen tiefenpsychologischen Ausbildungshintergrund. Die übrigen 177 Therapeuten (43.9%) geben eine rein tiefenpsychologische oder analytische Ausbildung an. Auf Therapeutenebene fand eine Randomisierung nach IG (zu Studienbeginn n=203) und KG (zu Studienbeginn n=200) statt. Für die Therapeuten der IG war vor Projektbeginn die Teilnahme an einer einführenden Schulung verpflichtend. Die teilnehmenden Therapeuten behandelten alle in ihrer Praxis neu beginnenden Patienten nach den Bedingungen des TK-Modellvorhabens, vorausgesetzt die Patienten waren mindestens 18 Jahre alt, Versicherte der Techniker Krankenkasse und erklärten sich nach einer Aufklärung über Zweck und Vorgehen des Modellvorhabens schriftlich zu einer Teilnahme bereit. Als Ausschlussdiagnosen wurden Störungen aus den Gruppen „Psychische und Verhaltensstörungen durch psychotrope Substanzen“ (F10-F19 des ICD-10) sowie „Schizophrenie, schizotype und wahnhaftige Störungen“ (F20-F29 des ICD-10) definiert. Bis zum Abschluss der Datenerhebung im Juni 2010 konnten 1708 Patienten in das Projekt aufgenommen werden. Bei 1154 dieser Patienten (67.7%) gaben die behandelnden Therapeuten einen verhaltenstherapeutischen Ausbildungshintergrund an. Rund zwei Drittel der teilnehmenden Patienten (1162; 68.0%) sind weiblichen Geschlechts. Das Alter zu Projektende (Altersangaben liegen für 1697 Patienten vor) liegt im Mittel bei 44 Jahren (SD=11.5) oder 528 Monaten (SD=138). Der überwiegende Anteil (1571 von 1618; entspricht 97.1%) der Patienten besitzt die deutsche Staatsangehörigkeit. Rund 55% (bezogen auf 1579 Patientenangaben) gaben als höchsten allgemeinbildenden Schulabschluss die Fachhochschul- oder Hochschulreife an, weitere 30% die mittlere Reife. Zu Beginn der Therapie waren gut drei Viertel der Patienten (76.1%; bezogen auf 1577 Angaben) als arbeitsfähig eingestuft, für weitere 5% ist die Arbeitsfähigkeit nicht bekannt. Aus den eingesetzten psychometrischen Instrumenten ergibt sich eine deutliche Beeinträchtigung zu Therapiebeginn. Sowohl im BSI (M=1.23, SD=.67), als auch im IIP (M=1.55, SD=.55) und BDI (M=20.60, SD=9.23) liegen die einbezogenen Patienten im Mittel im Bereich ambulanter Vergleichsstichproben oder darüber (Wittmann et al., 2011). Informationen über einen Therapieabschluss im Rahmen der Laufzeit des TK-Modells liegen für 891 Patienten (52.2%) vor.

#### 4.1.4 Teilstichprobe der vorliegenden Arbeit

Wie im vorherigen Abschnitt angedeutet variieren die Messpläne in Abhängigkeit vom angewendeten psychotherapeutischen Verfahren, mit den größten Abständen zwischen Verlaufserhebungen für psychoanalytisch behandelte Patienten und geringsten Abständen für verhaltenstherapeutisch behandelte Patienten (s. Abbildung 1). Aufgrund der kürzeren Erhebungsabstände konnte im begrenzten Zeitraum der Studie eine größere Anzahl an Zwischenmessungen bei den verhaltenstherapeutisch behandelten Patienten durchgeführt werden. Darüber hinaus geben für rund zwei Drittel der Patienten die Therapeuten einen (unter anderem) verhaltenstherapeutischen Ausbildungshintergrund an. Aus diesem Grund werden in der vorliegenden Arbeit Aussagen zunächst auf die Gruppe der verhaltenstherapeutisch behandelten Patienten begrenzt.

Die Grund-Stichprobe der dieser Arbeit beinhaltet folglich 1154 ambulante Psychotherapiepatienten, deren Therapeuten verhaltenstherapeutisch ausgebildet sind. Sie werden von 153 Therapeuten behandelt (im Mittel 7.5 Patienten pro Therapeut, Range 1 bis 36). Die Patienten sind zu 67.5% (779) weiblichen Geschlechts, im Durchschnitt 44 Jahre alt (SD=11.42) und gehören zu 61% (704) der Interventionsgruppe der TK-Studie an. Es ergeben sich keine Unterschiede der gewählten Patientenstichprobe zur ausgeschlossenen Stichprobe hinsichtlich der Geschlechtsverteilung ( $\chi^2_{(N=1708, df=1)}=.457; p=.50; \text{Phi} = .016$ ), der Zugehörigkeit zu IG und KG des TK-Projektes ( $\chi^2_{(N=1708, df=1)}=.613; p=.43; \text{Phi} = .019$ ), selbstberichteter Belastung zu Therapiebeginn (BSI-GSI:  $t_{(1575)}=-1.302; p=.19; 1-\beta=.99$ ; IIP:  $t_{(1583)}=.639; p=.52; 1-\beta=.99$ ; BDI:  $t_{(1153)}=-.035; p=.97; 1-\beta=.99$ ) und dem fremdeingeschätzten allgemeinen Funktionsniveau zu Therapiebeginn (GAF:  $t_{(1473)}=-1.266; p=.21; 1-\beta=.97$ ). Der Unterschied bezüglich des Alters erreicht hingegen statistische Signifikanz ( $t_{(1695)}=-2.123; p<.05; d=.17$ ), fällt mit im Durchschnitt 15 Monaten jedoch gering aus.

Tabelle 1: Demografische Angaben sowie Mittelwerte der Eingangsmessungen in der Gesamtstichprobe und den zur Datenanalyse verwendeten Teilstichproben mit den zugehörigen Hypothesen

	Hypo- these(n)	Alter (Monate)	Geschlecht (N, % weiblich)	BSI-GSI	IIP Gesamt	BDI Gesamt
TK-Gesamtstichprobe (N=1708)	/	N=1697 M=528 SD=138	1162 (68%)	N=1577 M=1.23 SD=.67	N=1585 M=1.55 SD=.55	N=1155 M=20.60 SD=9.23
Therapeut mit VT-Ausrichtung (N=1154)	1	N=1148 M=533 SD=137	779 (68%)	N=1073 M=1.24 SD=.68	N=1077 M=1.54 SD=.57	N=775 M=20.61 SD=9.33
Therapeut VT, Patient eindeutig Diagnosegruppen zuzuordnen (N=967)	1, 2, 3, 4	N=961 M=536 SD=137	651 (67%)	N=908 M=1.24 SD=.68	N=912 M=1.53 SD=.56	N=656 M=20.74 SD=9.26
Therapeut VT, IG (N=704)	/	N=703 M=530 SD=137	470 (67%)	N=684 M=1.26 SD=.67	N=686 M=1.56 SD=.57	N=486 M=20.98 SD=9.01
Therapeut VT, IG, Diagnoseinfo vorhanden (N=558)	/	N=559 M=532 SD=137	368 (66%)	N=553 M=1.24 SD=.66	N=556 M=1.54 SD=.56	N=391 M=20.76 SD=8.87
Therapeut VT, IG, Diagnoseinfo und Endwert BSI vorhanden (N=314)	9, 10	N=314 M=542 SD=137	201 (64%)	N=310 M=1.22 SD=.63	N=313 M=1.51 SD=.54	N=219 M=20.37 SD=8.67
Therapeut VT, IG, Diagnoseinfo, mind. 2 der ersten 4 Mzp im BSI vorhanden (N=462)	5, 6, 7	N=461 M=540 SD=136	307 (67%)	N=459 M=1.29 SD=.65	N=462 M=1.58 SD=.55	N=332 M=21.44 SD=8.69
Therapeut VT, IG, Diagnoseinfo, Mzp 1 bis 3 im BSI vorhanden (N=321)	8	N=320 M=544 SD=137	208 (65%)	N=321 M=1.32 SD=.66	N=321 M=1.61 SD=.53	N=233 M=21.76 SD=8.70
Therapeut VT, IG, Diagnoseinfo, Mzp 1 bis 3 und Endwert im BSI vorhanden (N=206)	8	N=206 M=550 SD=133	132 (64%)	N=206 M=1.28 SD=.64	N=206 M=1.58 SD=.51	N=150 M=21.16 SD=9.07

Für die Überprüfung der aufgestellten Hypothesen sind weitere Stichprobenreduktionen notwendig. Während für die Hypothesen 1 bis 4 Patienten der IG und KG untersucht werden, beschränken sich die Analysen zu den Hypothesen 5 bis 10 auf die IG, was darin begründet liegt, dass nur hier Zwischenerhebungen stattfanden, die in die Verlaufsanalysen eingehen. Eine weitere Einschränkung betrifft die eindeutige Zuordnung zu Diagnosegruppen, welche lediglich für Hypothese 1a nicht gegeben sein muss. Darüber hinaus können bei Testung der Verlaufshypothesen einige Patienten nicht eingeschlossen werden, da kei-

ne oder zu wenige Messzeitpunkte vorlagen, um die entsprechenden Analysen durchzuführen. Eine Übersicht der Schritte zur Datenreduktion mit den zugehörigen Hypothesen findet sich im Anhang (S. 171). Die Größe der untersuchten Teilstichproben, ihr Bezug zu den Hypothesen, demografische Angaben und deskriptive Informationen zur mittleren Belastung in den Instrumenten BSI, IIP und BDI sind in Tabelle 1 zusammengefasst.

Die Patienten, die in die Verlaufsfragestellungen eingehen, stellen sich im Mittel in allen relevanten Maßen als etwas belasteter dar und sind im Durchschnitt etwas älter als die Patienten der Ursprungsstichprobe.

## **4.2 Methoden zur Beantwortung der Fragestellungen (Status)**

In diesem Unterkapitel wird die Vorgehensweise zur Beantwortung der Hypothesen 1 bis 4, welche sich auf den Status der untersuchten Patienten zu Therapiebeginn beziehen, dargestellt.

### **4.2.1 Zur Beantwortung der Hypothese 1**

Der erste Hypothesenblock nimmt Bezug auf die differierenden Diagnosemethoden in Interventions- und Kontrollgruppe. Während die Therapeuten der Kontrollgruppe keine Vorgaben zur Art und Weise ihrer Diagnosestellung erhielten, sollten Therapeuten der Interventionsgruppe einem teilstrukturierten Vorgehen folgen. Zum Verzeichnen (aktuell vorhandener) komorbider Diagnosen wurde explizit aufgefordert. Pro Patient konnten maximal vier Diagnosen vergeben werden. Auf die klassifikatorische Diagnostik in der Interventionsgruppe wird im Folgenden genauer eingegangen.

#### *Klassifikatorische Diagnostik in der Interventionsgruppe*

Zunächst sollte, in der Regel beim Patienten-Erstkontakt, ein Screening für ICD-10-Diagnosen durchgeführt werden. Das Screening enthält 22 Fragen zu allen psychischen Störungen, die in das TK-Modellvorhaben eingeschlossen werden (Störungsgruppen F3 bis F6 des ICD-10). Sie entstammen dem SKID-I-Interview (Wittchen et al., 1997) und dem IDCL-Screening (Hiller et al., 2004) und werden durch Ankreuzen bejaht. Bestand der

Verdacht auf eine Störung, so wurde im Anschluss mit der entsprechenden IDCL-Checkliste (Hiller et al., 2004) näher darauf eingegangen. Die im Rahmen dieser Studie zum Einsatz kommenden Internationale Diagnose Checklisten beruhen auf den psychischen und Verhaltensstörungen nach ICD-10 (Weltgesundheitsorganisation, 2005). Jede Checkliste bezieht sich auf eine Diagnose und enthält die zur Diagnosestellung erforderlichen Kriterien und Merkmale, die weitestgehend im ICD-10-Wortlaut gehalten sind. Durch Markierung der Codierkästen („Ja“ – „Nein“ – „Verdacht“) wird der Therapeut anhand klarer Regeln durch die Diagnosestellung geleitet. Es handelt sich folglich um ein Instrument zur Fremdbeurteilung durch den Therapeuten, welcher nach hypothesengeleitetem Vorgehen eine Klassifikationsentscheidung trifft. Obwohl die IDCL auch die Erfassung von Lebenszeit-Diagnosen erlauben, sollten diese im Rahmen des TK-Modellprojektes nicht berücksichtigt werden. Hier gingen ausschließlich aktuelle Diagnosen ein; eine lebenszeitliche Betrachtung erfolgte nur dann, wenn diese für die aktuelle Diagnose relevant war (z. B. bei rezidivierenden depressiven Störungen). Für Störungen, die in den Checklisten nicht abgedeckt sind (z. B. sexuelle Funktionsstörungen oder Schlafstörungen) wurden die Therapeuten angehalten, die entsprechenden Kriterien im Kapitel V des ICD-10 (Weltgesundheitsorganisation, 2005) zu Rate zu ziehen. In einem dritten Schritt erfolgte die Diagnostik von Persönlichkeitsstörungen mit Hilfe der Internationale Diagnose Checklisten für Persönlichkeitsstörungen (IDCL-P; Bronisch et al., 1995). Darin werden zwölf Persönlichkeitsstörungen nach ICD-10 (Weltgesundheitsorganisation, 2005) aufgeführt, die analog zu den Störungen der Kategorien F3 bis F5 abgefragt werden sollten. Die Reliabilität der Checklisten wird als gut eingestuft (Hiller, Dichtl, Hecht, Hundt, & Zerssen, 1993).

#### *Zusammenfassung zu Diagnosegruppen*

Zunächst wird deskriptiv eine erste Übersicht darüber gewonnen, welche Anzahl an Diagnosen in Interventionsgruppe (unter Anwendung der Diagnosechecklisten) und Kontrollgruppe (ohne Vorgaben) gestellt wurde. Außerdem werden Mittelwertsvergleiche mit der Gruppenzugehörigkeit als Faktor angestellt.

In einem nächsten Schritt werden Diagnosegruppen gebildet. Für die nachfolgenden Analysen wird jedem Patienten in jeder dieser Kategorien entweder der Wert 1 (=mindestens eine Diagnose aus dieser Kategorie vorhanden) oder 0 (=nicht vorhanden) zugewiesen. Die

so gebildeten Kategorien sind in Tabelle 2 aufgeführt. „Andere Störungen“ werden dabei in einer eigenen Gruppe zusammengefasst.

Tabelle 2: Einteilung der Störungsdiagnosen in 16 Kategorien plus „Andere Störungen“

Diagnosegruppe	Eingehende Diagnose(n) nach ICD-10
Depressive Episode	F32.X
Rezidivierende depressive Störung	F33.X
Dysthymia	F34.1
Bipolare Störung/Manie	F30.X, F31.X, F34.0
Panikstörung/Agoraphobie	F40.00, F40.01, F41.0
Soziale Phobie	F40.1
Spezifische Phobie	F40.2
Generalisierte Angststörung	F41.1
Angst und depressive Störung, gemischt	F41.2
Zwangsstörung	F42.X
Akute Belastungsstörung	F43.0
Posttraumatische Belastungsstörung	F43.1
Anpassungsstörung	F43.2
Somatoforme Störungen	F45.X
Essstörungen	F50.X
Persönlichkeitsstörungen	F60.X, F61.X, F62.X
(Andere Störungen)	(sonstige Störungen oder nicht näher bezeichnete Störungen FXX.8 oder FXX.9, F00-F29, F38, F39, F41.3, F44, F48, F52-F59, F63-F69)

Darüber hinaus werden Komorbiditätsgruppen gebildet, welche sich an den in der Literatur besonders häufig und als relevant beschriebenen Diagnosekombinationen orientieren. So werden alle Patienten, die mindestens eine Diagnose aus der Gruppe der Persönlichkeitsstörungen (F6) sowie mindestens eine weitere Diagnose aus den Gruppen anderer im Modellprojekt eingeschlossener Störungsgruppen aufweisen, der Gruppe „komorbide mit Persönlichkeitsstörung“ zugewiesen. Alle Patienten, für die die Kriterien mindestens einer affektiven Störung erfüllt sind (F30-F34) und die darüber hinaus mindestens eine Diagnose aus dem Bereich der Angststörung erhalten haben (F40-F41) und keine Persönlichkeitsstö-

rung (F6) aufweisen, werden der Gruppe „komorbide Angst- und affektive Störungen“ zugewiesen. Alle weiteren Störungskombinationen, die nicht in die genannten Kategorien fallen, werden der Gruppe „komorbide, andere“ zugeordnet. In letztere Gruppe fallen somit auch Patienten, die zwei Diagnosen aus einem Bereich erhalten haben, etwa Patienten mit einer Kombination aus Dysthymia und Depressiver Episode (sog. „Double Depression“) oder Patienten mit mehreren Angststörungen. Eine vierte Gruppe enthält solche Patienten, für die lediglich eine Diagnose vergeben wurde („keine Komorbidität“). Die Verteilung der Patienten auf diese vier Komorbiditätsgruppen wird im Hinblick auf ihre Vergleichbarkeit mit der Gesamtstichprobe geprüft.

### *Beantwortung der Fragestellungen*

Zur Beantwortung der Frage, ob unter Verwendung der Internationale Diagnosen Checklisten (IDCL) bei mehr Patienten Komorbidität resultiert (Hypothese 1a), wird eine logistische Regression gerechnet. Darin geht die dichotome Variable „Komorbidität“ als Kriterium ein. Als Prädiktoren werden aufgenommen: Alter und Geschlecht, Eingangsbelastung erfasst über den BSI-GSI zu Therapiebeginn und die Zugehörigkeit zu Interventions- oder Kontrollgruppe. Es wird angenommen, dass – bei Kontrolle für erstere drei Variablen – die Information über Gruppenzugehörigkeit einen signifikanten Beitrag zur Zuordnung der Patienten in eine komorbide und eine nicht komorbide Gruppe leistet.

Die Hypothesen 1b und 1c werden ebenfalls mit Hilfe logistischer Regressionsanalysen überprüft. Jede der in Tabelle 2 aufgelisteten Störungsgruppen wird auf Gleichverteilung in IG und KG untersucht, und zwar ebenfalls unter Kontrolle für die möglicherweise erklärenden Variablen Alter, Geschlecht und Eingangsbelastung. Als Kriterium dient folglich jeweils eine Störungsgruppe, als Prädiktoren Alter, Geschlecht, der BSI-GSI zu Therapiebeginn sowie die Zugehörigkeit zu Interventions- oder Kontrollgruppe. Es wird angenommen, dass die Information über Gruppenzugehörigkeit einen bedeutsamen Beitrag zur Erklärung der Häufigkeit der hypothesenrelevanten Störungsgruppen Persönlichkeitsstörungen, Angststörungen und Anpassungsstörungen leistet.

#### 4.2.2 Zur Beantwortung der Hypothesen 2 und 3

Hypothesen 2 und 3 beziehen sich auf die Zusammenhänge, die zwischen Störungsdiagnosen bestehen und darauf, inwieweit bestimmte Diagnosen gehäuft parallel gestellt werden. Zunächst erfolgt auch hier eine deskriptive Zusammenschau zum gemeinsamen Auftreten von Störungsdiagnosen auf einer bivariaten Ebene. Um eine erste Übersicht zum Zusammentreffen jeweils zweier Diagnosen zu erhalten, wird für die in Tabelle 2 festgehaltenen Störungsgruppen in den insgesamt 120 möglichen Zweier-Kombinationen jeweils die Anzahl betroffener Patienten und der relative Anteil der spezifischen Kombination an der Gesamtheit aller Diagnosen einer Gruppe bestimmt. Für diese bivariaten Störungskombinationen werden im Anschluss Zusammenhangsmaße auf Grundlage von Vier-Felder-Tafeln berechnet (Hypothese 2). Der *tetrachorische Korrelationskoeffizient* bezieht sich auf den Zusammenhang zwischen zwei zweistufigen Variablen und beruht auf der Annahme, dass diese Variablen künstliche Dichotomien normalverteilter Variablen darstellen. Auf diese Annahme wird auch mit dem Begriff des „liability threshold model“ verwiesen (z. B. Neale, 2005) und sie gilt gemeinhin als für psychische Störungen angemessen (z. B. Krueger, 1999). Der tetrachorische Korrelationskoeffizient schätzt folglich die „wahre“ Korrelation zwischen den beiden Skalen (Bortz, 2005). Die Näherungsformel zur Berechnung tetrachorischer Korrelationen ist in Formel 1 dargestellt, wobei die Buchstaben a, b, c und d die Kombinationen der Merkmalsalternativen in der Vier-Felder-Tafel kennzeichnen.

$$r_{tet} = \frac{\cos(180^\circ)}{1 + \frac{\sqrt{(a \times d)}}{b \times c}} \quad (\text{Bortz, 2005, S. 231})$$

Formel 1:

Die Werte der tetrachorischen Korrelation liegen zwischen -1 und +1. Sie sollten allerdings nur dann interpretiert werden, wenn alle vier eingehenden Merkmalskombinationen mindestens eine Häufigkeit von 5 aufweisen (Wirtz, 1998). Das Programm G\*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) ermöglicht die Berechnung des Korrelationskoeffizienten  $r_{tet}$  und stellt Prozeduren zur Effektgrößenberechnung und zur Testung der Signifikanz des gefundenen Zusammenhangs zur Verfügung. Für diese Arbeit wird die Schätzung des exakten Korrelationskoeffizienten nach Brown und Benedett (1977) gewählt. Auf Grundlage der beobachteten Häufigkeiten und der gewählten Irrtumswahrscheinlichkeit wird der

Korrelationskoeffizient  $r_{tet}$  mit dem zugehörigen Standardfehler berechnet. Sofern  $n > 20$  ist, kann die Nullhypothese  $H_0: \rho_{tet} = 0$  durch den Signifikanztest nach Formel 2 überprüft werden.

Formel 2:  $z = \frac{r_{tet}}{\sigma_{tet}}$ , wobei  $\sigma_{tet} = \frac{\sqrt{p_x \times p_y \times q_x \times q_y}}{n} \times \frac{1}{\vartheta_x \times \vartheta_y}$  (Bortz, 2005 S. 230)

Dabei bezeichnen  $\vartheta_x$  und  $\vartheta_y$  die Ordinaten desjenigen z-Wertes der Standardnormalverteilung, der die Verteilung in die jeweiligen Anteile trennt. Zur Überprüfung der 120 bivariaten Störungskombinationen auf bedeutsame positive und negative Zusammenhänge hin wird jeweils der Korrelationskoeffizient  $r_{tet}$  berechnet und gegen die Nullhypothese auf Signifikanz getestet.

Nachdem auf diese Weise ein Überblick über bivariate Merkmalszusammenhänge geschaffen wurde, wird im nächsten Schritt eine multidimensionale Betrachtung angestrebt (Hypothese 3). Anstelle zahlreicher Einzelbetrachtungen von jeweils nur zwei Diagnosen sollen nun alle definierten Diagnosegruppen gemeinsam analysiert werden. Dies hat den Vorteil, dass Probleme multiplen Testens umgangen werden. Gleichzeitig kann durch den parallelen Einbezug aller verfügbaren Information ein umfassenderes Bild nicht nur zu Paaren von typischerweise gemeinsam diagnostizierten Störungen gewonnen werden, sondern differenziertere Muster von möglicherweise mehreren parallelen Diagnosen identifiziert werden. Zu diesem Zweck werden die in Tabelle 2 aufgeführten, dummy-codierten Diagnosegruppen einer *Latent Class Analyse* unterzogen. Als Latent Class Analyse wird eine statistische Methode bezeichnet, die auf der Annahme gründet, dass Zusammenhänge zwischen kategorialen beobachteten (manifesten) Variablen erklärt werden können durch eine kleinere Anzahl diskreter unbeobachteter (latenter) Variablen, die sich gegenseitig ausschließen (Clogg, 1995; Klein & Riso, 1993). Sie wird auch als das kategoriale Analogon zur Faktorenanalyse verstanden und dementsprechend als besonders geeignet für Analysen auf der Ebene von Diagnosen und Symptomen eingeschätzt (McCutcheon, 2002; Muthén, 2002; Muthén & Muthén, 2000). Das Grundprinzip, welches besagt, dass innerhalb einer gegebenen Klasse die beobachteten Variablen (in diesem Fall Diagnosegruppen) voneinander unabhängig sind und die latente Variable die beobachteten Zusammenhänge erklärt, wird lokale Unabhängigkeit („local independence“) genannt (Clogg, 1995). Latent Class

Analysen erlauben den Einbezug von Kovariaten analog zur logistischen Regressionsanalyse (Dayton & Macready, 2002; Muthén, 2002). Mit dem Ziel, die Zusammenhänge zwischen beobachteten Variablen so umfassend und sparsam wie möglich zu erklären, werden schrittweise Klassen in das Modell eingeführt, bis eine gute Modellpassung erreicht ist. Es wird folglich getestet, wie gut die erwarteten Zellbesetzungen unter Annahme des Modells die beobachteten Zellbesetzungen wiedergeben können. Die Modelltestung erfolgt über Maximum Likelihood Schätzverfahren (McCutcheon, 2002). Relevante Modellparameter sind hier die Wahrscheinlichkeit der Zugehörigkeit zu jeder Klasse („latent class probabilities“) und die Wahrscheinlichkeit, bei einer gegebenen Klassenzugehörigkeit jedes der Kriterien (in diesem Fall Diagnosegruppen) zu erfüllen („latent class scores“; Clogg, 1995; Muthén & Muthén, 2000). Zusätzlich werden für jede in die Analyse eingehende Person individuelle Schätzer für die Klassenzugehörigkeiten bereitgestellt (sog. posteriore Wahrscheinlichkeiten). Zur Auswahl des Modells mit der optimalen Anzahl an Klassen können unterschiedliche Kriterien herangezogen werden. Das *Bayes Information Criterion* (BIC; Schwartz, 1978) ist ein Index zum Ausmaß der Passung des vorhergesagten Modells zu den beobachteten Daten, der auch die Sparsamkeit eines Modells berücksichtigt. Modelle mit besserer Passung nehmen geringere Werte an (Nagin, 1999). Das *Entropie*-Kriterium (Ramaswamy, DeSabro, Reibstein, & Robinson, 1993) macht darüber hinaus Aussagen über die Qualität der Klassifikationsvorhersage, die für die Modellbewertung als ebenso wichtig wie die Passungsgüte eingeschätzt wird (Clogg, 1995). Die posterioren Wahrscheinlichkeiten sollten dazu im Idealfall pro Person für eine Klasse hoch und für die verbleibenden Klassen niedrig ausfallen. Die Entropie kann Werte zwischen 0 und 1 annehmen, wobei hohe Werte eine bessere Entropie, das heißt eine eindeutigere Zuordnung der Personen in die identifizierten Klassen bedeuten (Muthén et al., 2002). Der *Lo-Mendell-Rubin-Test* (LMR; Lo, Mendell, & Rubin, 2001) vergleicht das geschätzte Modell mit einem Alternativmodell, das eine Klasse weniger enthält. Der erhaltene *p*-Wert steht für die Wahrscheinlichkeit, dass die Daten durch ein Modell mit einer Klasse weniger besser repräsentiert werden. Bei einem niedrigen *p*-Wert sollte das geschätzte Modell beibehalten werden (Muthén & Muthén, 2010). Ähnlich ist das Prinzip des *Bootstrap-Likelihood-Ratio-Differenztests* (BLRT; Nylund, Asparouhov, & Muthén, 2007). Auch hier wird der Modellfit des Modells mit *k* Klassen gegenüber einem Modell mit *k*-1 Klassen getestet. Die Verteilung der Differenzstatistik wird dabei über Bootstrapping-Methoden geschätzt.

### 4.2.3 Zur Beantwortung der Hypothese 4

Bezüglich des Zusammenhangs zwischen Komorbidität und Belastungsschwere werden selbst- und fremdberichtete Belastung als abhängige Variablen untersucht. Sowohl bezüglich der durch den Patienten selbst erlebten als auch bezüglich der durch den Therapeuten eingeschätzten Beeinträchtigung kamen im Rahmen des TK-Modellvorhabens mehrere Instrumente zu Einsatz. Daher werden für die Untersuchung der Fragestellung zunächst zwei zusammengesetzte Beeinträchtigungsmaße, getrennt für selbst- und fremdberichtete Belastung, gebildet. Auf Seiten der Selbstauskünfte gehen der BSI-GSI (Franke, 2000), IIP-D (Horowitz et al., 2000) und SF-12 (Bullinger & Kirchberger, 1998) ein. Das Maß zu den Fremdbereichten setzt sich zusammen aus der aktuellen Einschätzung des globalen Funktionsniveaus (GAF; Endicott et al., 1976), den Einschätzungen zur allgemeinen Belastung (BSS; Schepank, 1995) und Problemrelevanten Lebensbereichen.

Die Hypothesen 4a und 4b werden mit Hilfe von Varianzanalysen geprüft. Die Information über Zugehörigkeit zu den latenten Diagnoseklassen (wie aus den Latent Class Analysen berechnet) beziehungsweise über Zugehörigkeit zu den vorab definierten Komorbiditätsgruppen dient als Gruppierungsfaktor. Bei gegebenen Gruppenunterschieden zwischen IG und KG aus Fragestellung 1 wird die Gruppenzugehörigkeit (IG/KG) als zweiter Faktor einbezogen. Sofern nicht anders angegeben, wird das Niveau des  $\alpha$ -Fehlers auf 5% festgelegt.

### 4.3 Methoden zur Beantwortung der Fragestellungen (Verlauf und Ergebnis)

Das folgende Unterkapitel befasst sich mit der Vorstellung der Methoden zur Überprüfung der Hypothesen 5 bis 10, welche sich auf den Therapieverlauf und das Therapieergebnis beziehen.

#### 4.3.1 Latent Growth Curve Models (LGM)

Zur Beantwortung des ersten Teils der Verlaufsfragestellungen wird auf *Latente Wachstumskurvenmodelle* (sog. Latent Growth Curve Models, LGM) zurückgegriffen. Dieser Ansatz kann als spezielle Untergruppe der Strukturgleichungsmodelle verstanden werden, wobei hier eine beobachtete Zielvariable in Bezug gesetzt wird zur vergangenen Zeit beziehungsweise einer mit Zeit verbundenen Variablen. Die gemessenen Variablen stellen also wiederholte Erhebungen eines Merkmals zu mehreren Messzeitpunkten dar. Diese dienen als Indikatoren für latente Variablen, die hier im Gegensatz zu konventionellen Strukturgleichungsmodellen nicht für psychologische Konstrukte stehen, sondern Merkmale der Veränderung in der Zielvariablen repräsentieren, nämlich Höhenlage (intercept) und Steigung (slope). Während der intercept-Faktor die Ausprägung der Zielvariablen zum Zeitpunkt 0 abbildet, bezeichnet der slope-Faktor die Veränderungsrate im Ergebnismaß. Es gilt die Annahme, dass die gemessenen Ergebnisvariablen direkt von den latenten Wachstumsparametern beeinflusst werden, daneben aber auch einem Messfehler unterliegen (Geiser, 2010). Im Rahmen des LGM können die mittleren Werte dieser Wachstumsparameter geschätzt werden und die individuelle Abweichung vom mittleren Trend wird durch die Varianz der Wachstumsparameter erfasst (Preacher, Wichman et al., 2008). Das Verfahren eignet sich gut zur Untersuchung der Form der Veränderung über die Zeit (Geiser, 2010). Darüber hinaus können der erklärende Einfluss von festen oder variierenden Kovariaten auf die Wachstumsparameter sowie die Kovariation zwischen intercept- und slope-Faktor untersucht werden. Damit erlaubt das Verfahren sowohl, spezifische Hypothesen zu angenommenen mittleren Verlaufsformen zu testen, als auch individuelle Unterschiede in den Verläufen über die Zeit zu erfassen und Korrelate der Veränderung zu untersuchen. Beim Einbezug von Prädiktoren in das Modell spricht man auch von konditionalen Wachstumsmodellen („conditional growth models“; u. a. Preacher, Wichman et al.,

2008). Stärken und Schwächen des LGM sind vergleichbar derer von Strukturgleichungsmodellen und über die Angemessenheit eines Modells wird auch hier mit Hilfe von Passungsgüteindizes entschieden (Duncan & Duncan, 2004). Zum Modellvergleich empfohlen werden Informationskriterien (Akaike's Information Criterion, AIC; Bayesian Information Criterion; BIC) und der Root Mean Square Error of Approximation (RMSEA; Preacher, Wichman et al., 2008). Die Parameter-Schätzung erfolgt über Maximum-Likelihood (ML)-Schätzverfahren. Dabei besteht die Annahme, dass die gemessenen Variablen multivariat normalverteilt sind (Preacher, Wichman et al., 2008). Das Verfahren erlaubt den Einschluss fehlender Werte (Bollen & Curran, 2006).

Zur Beantwortung der in den Hypothesen 5 bis 7 formulierten Fragestellungen werden getrennt für die allgemeine Belastung (BSI) und die interpersonale Belastung (IIP-D) latente Wachstumskurvenmodelle geprüft. Auf diese Weise werden Verläufe betrachtet, die möglicherweise spezifisch für bestimmte Instrumente bestehen und die in spezifischer Weise durch das Vorliegen komorbider Störungen beeinflusst sein können. Es wird sowohl im BSI als auch im IIP ein signifikanter slope-Faktor erwartet, welcher signalisiert, dass die eingehenden Patienten sich im Allgemeinen verbessern (Hypothese 5). Darüber hinaus wird erwartet, dass sowohl für den BSI als auch für den IIP ein Modell mit random slope eine bessere Passung aufweist als ein Modell mit fixed slope und dass Komorbidität als Prädiktor für die Steigung im BSI und IIP genutzt werden kann (Hypothese 6). Daneben soll explorativ untersucht werden, ob eine bedeutsame Korrelation zwischen intercept und slope in BSI und IIP besteht (Hypothese 7).

#### **4.3.1.1 Stichprobe und eingehende Messzeitpunkte**

Die Analysen beziehen sich wiederum nur auf Patienten, deren Therapeuten einen verhaltenstherapeutischen Ausbildungshintergrund haben. In die Verlaufsanalysen einbezogen werden können außerdem lediglich Patienten der IG (nur hier sah der Messplan verlaufsbegleitende Erhebungen vor). Bei Betrachtung der messwiederholten Daten wird ein typisches Problem naturalistisch erhobener Daten augenscheinlich, nämlich dass zum einen fehlende Werte vorliegen und dass zum anderen eine gewisse Variation bezüglich der tatsächlichen Erhebungszeitpunkte besteht. Laut Messplan sollten Zwischenmessungen für verhaltenstherapeutisch behandelte Patienten in der 10., 20., 40. usw. Therapiesitzung statt-

finden (s. Abbildung 1). Tabelle 3 zeigt die Kennwerte der Datenerhebungspunkte für die selektierte Patientengruppe (IG, VT). Zwar befinden sich, zumindest bis Messzeitpunkt 4, alle Lagemaße im Bereich der vorgegebenen Messzeitpunkte, jedoch besteht auch eine nicht unerhebliche Streuung. Hinzu kommt, dass für einige Patienten innerhalb kurzer Zeit mehrere Messungen im gleichen Instrument verzeichnet sind. Aus diesem Grund erfolgt zunächst eine Aufbereitung der Daten im Hinblick auf eine Vereinheitlichung der Erhebungszeitpunkte. Als der Eingangsmessung zugehörig werden dabei alle Messungen definiert, die zwischen dem Zeitpunkt vor der ersten Sitzung bis zur fünften Sitzung stattgefunden haben. Unter den Messzeitpunkt „1. Zwischenmessung“ fallen Erhebungen, die um die zehnte Sitzung streuen, und zwar in frühestens der achten, spätestens der zwölften Sitzung durchgeführt wurden. Entsprechend werden unter den Messzeitpunkt „2. Zwischenmessung“ alle Daten gefasst, die zwischen frühestens der 17. und spätestens der 23. Sitzung anfielen und unter den Messzeitpunkt „4. Zwischenmessung“ alle Daten, die zwischen der 35. und 45. Sitzung erhoben wurden. Sofern mehr als eine Messung in die so definierten Zeiträume fällt, geht der daraus gebildete Mittelwert in die Analysen ein.

Tabelle 3: Therapiesitzung, in welcher die im Messplan vorgegebenen Eingangs- und Zwischenmessungen tatsächlich stattgefunden haben (IG, VT).

Messzeitpunkt	1	2	3	4	5	6	7	8
N	685	582	493	365	156	56	12	2
MW (Sitzung)	2,63	10,37	20,78	37,43	45,22	53,79	53	77,5
SD	1,315	3,773	6,647	10,76	16,85	25,57	46,6	10,61
Modus	3	10	20	40	45	0 <sub>a</sub>	0	70 <sub>a</sub>
Median	3	10	20	40	45	49,5	60,5	77,5

<sub>a</sub>= mehrere Modi vorhanden, der kleinste wird angezeigt; N=704

Ab dem fünften Messzeitpunkt steigt der Anteil fehlender Werte deutlich an, weshalb die Verlaufsanalysen lediglich die Daten bis einschließlich Messzeitpunkt vier eingehen. Der Einbezug von vier Erhebungen ist notwendig, um auch komplexere Verlaufsformen hinsichtlich ihrer Passung gegeneinander zu testen (Curran & Hussong, 2008). Weiterhin werden pro eingehendem Patienten zwei fehlende Werte zugelassen. Patienten mit keiner oder nur einer Erhebung während der ersten vier zusammengefassten Messzeitpunkte wer-

den aus den Verlaufsanalysen ausgeschlossen (s. Anhang, S. 171). Da die spezifischen Fragestellungen der folgenden Verlaufsanalysen sich auf die mögliche Einflussnahme von Mehrfachdiagnosen beziehen, werden weiterhin nur solche Patienten aufgenommen, für die diagnostische Information gegeben ist. Es resultiert für die LGM eine Patientenstichprobe von  $N=462$ .

#### 4.3.1.2 LGM - Vorgehen

Aufgrund der fehlenden Werte werden für die Latenten Wachstumsmodelle multiple Imputationen genutzt. Bei diesem Missing-Data-Verfahren wird jeder fehlende Wert nicht durch einen einzelnen, sondern durch ein Set möglicher Werte, basierend auf den vorliegenden Werten, ersetzt (Rubin, 1987; Yuan). Es bestehen mehrere Methoden zur multiplen Imputation, eine davon beruht auf Regressionen, wobei die Annahmen multipler Normalverteilung und zufälliger fehlender Werte (missing at random, MAR) gelten (Bollen & Curran, 2006). Fehlende Werte werden wiederholt ersetzt, so dass schließlich mehrere vollständige Datensätze vorliegen, auf die die entsprechenden Analysen zunächst separat angewendet werden. Schließlich werden die aus diesen Mehrfachanalysen resultierenden Parameterschätzer gemittelt und eine  $\chi^2$ -Gesamtteststatistik zum Modellfit generiert (Muthén & Muthén, 2010). Im Vergleich zum einfachen Ersetzen fehlender Werte liegt der Vorteil multipler Imputation darin, dass durch die Generierung einer Zufallsstichprobe fehlender Werte deren Vorhersageungenauigkeit adäquater widerspiegelt wird (Yuan). Für die vorliegenden Analysen werden zunächst zehn vollständige Datensätze imputiert. Die nachfolgenden Analysen werden auf diese neu generierten Datensätze angewendet. Alle Analysen erfolgen mit Hilfe der Software Mplus 6.1 (Muthén & Muthén, 2010).

Zunächst wird für jedes Ergebnismaß (BSI und IIP) ein Grundmodell bestimmt, welches dazu in der Lage ist, die gegebenen Daten möglichst genau und sparsam zu erklären. Es wird geprüft, welche Verlaufsfunktion die Daten am besten beschreibt, wozu ein intercept-only-Modell, ein lineares, loglineares und quadratisches Modell miteinander verglichen werden. Der Modellvergleich erfolgt über RMSEA und BIC. Nach Auswahl des Modells mit der besten Passung bezüglich des Verlaufs werden die Wachstumsparameter schrittweise fixiert und gegen ein Modell mit zufällig variierenden (random) intercept- und slope-Faktoren getestet. Da zu Beginn der Therapie bedeutsame interindividuelle Unterschiede in

allgemeiner und interpersoneller Belastung erwartet werden, sollte ein Modell mit random intercept eine bessere Passung aufweisen, als ein Modell mit fixed intercept. Darüber hinaus wird angenommen, dass im Verlauf der Behandlung eine bedeutsame Besserung allgemeiner und interpersoneller Belastung erreicht wird, sprich der slope-Faktor sollte Signifikanz erreichen. Auch bezüglich der Rate der Veränderung im Verlauf der Therapie werden interindividuelle Unterschiede angenommen, weshalb ein Modell mit random slope eine bessere Passung aufweisen sollte, als ein vergleichbares Modell mit fixed slope. Schließlich wird untersucht, ob in allgemeiner und interpersoneller Belastung ein bedeutender Zusammenhang zwischen Ausgangswert und Veränderungswert besteht. Ist dies der Fall, sollte eine signifikante Korrelation zwischen den latenten Wachstumsfaktoren bestehen.

Nachdem das jeweilige Grundmodell ausgewählt wurde, welches Eingangsbelastung und Verlauf in BSI und IIP sowie die Variation der Stichprobe in den Wachstumsparametern am besten beschreibt, werden Prädiktoren in das Modell eingefügt (konditionale Wachstumsmodelle), die potenziell dazu geeignet sind, Variation in den Wachstumsparametern zu erklären. Als Kovariaten dienen hier Variablen, die über das Vorliegen von Komorbidität informieren. Bei Vorhandensein komorbider Diagnosen sollten die Ausgangswerte in BSI und IIP bedeutsam erhöht sein. Sofern das Vorliegen bestimmter Diagnosekombinationen zu langsameren Verbesserungsraten führt, sollten die Komorbiditätsvariablen auch für den slope-Faktor prädiktiv bedeutsam sein.

#### **4.3.2 Growth Mixture Modeling (GMM)**

Bezüglich der Hypothese 8, die sich unter anderem mit dem potenziellen Zusammenhang zwischen Komorbidität und latenten Veränderungsgruppen in der Therapie befasst, wird auf Growth Mixture Modeling sowie Regressionsanalysen zurückgegriffen.

Wie im Abschnitt 4.3.1 erläutert, wird die individuelle Variation im Behandlungsverlauf in konventionellen Wachstumsmodellen (LGM oder auch Hierarchische Lineare Modelle, HLM) über die Varianz der Wachstumsparameter erfasst. Daneben geht man in jenen von der Zugehörigkeit aller eingehenden Personen zu einer gemeinsamen Population aus, oder aber es wird nach a priori bekannten Subgruppen, etwa Geschlecht, differenziert. Zusammengefasst können in konventionellen Wachstumsanalysen Schätzungen erfolgen zum

durchschnittlichen Wachstum aller Personen, zur Ausprägung der Varianz zwischen den Personen bezüglich des Eingangswertes und der Steigung und zum Einfluss von Kovariaten auf diese Variation (Muthén, 2004). Während in konventionellen Wachstumsanalysen folglich davon ausgegangen wird, dass die Variabilität zwischen Personen ausreichend gut beschrieben wird durch die Varianz der Wachstumsparameter, nimmt der *Growth Mixture Modeling* (GMM)-Ansatz an, dass Personen sich in klar voneinander abgegrenzte Gruppen zusammenfassen lassen, die jeweils um eigene Wachstumskurven mit eigenen Parametern variieren (Muthén, 2004). Hier wird also die Annahme einer einzelnen Population aufgelöst zugunsten der Annahme von Parameterunterschieden in unbeobachteten Subgruppen. Die Identifikation solcher unbeobachteter Gruppen von Personen wird über die Einführung einer kategorialen latenten Variablen in das Wachstumsmodell erreicht und erfolgt über Rückschluss aus den Daten. Für jede im Rahmen der kategorialen latenten Variablen definierten Gruppe von Patienten (auch als Klasse bezeichnet) wird eine mittlere Wachstumskurve geschätzt. Individuelle Variation um diese Kurve wird erfasst durch die Varianz in den Wachstumsfaktoren. Darüber hinaus können im Rahmen von GMMs für jede Person sogenannte posteriore Wahrscheinlichkeiten bestimmt werden, das heißt individuelle Wahrscheinlichkeiten für die Mitgliedschaft in jeder der identifizierten Klassen (Colder, Campbell, Ruel, Richardson, & Flay, 2002). Für die vorliegenden Daten wird im Rahmen eines GMM-Ansatz somit angenommen, dass mehrere nicht direkt beobachtbare Patientensubpopulationen bestehen, die sich hinsichtlich ihrer Verlaufskurven in den allgemeinen Belastungsmaßen BSI und IIP unterscheiden.

#### **4.3.2.1 Stichprobe und eingehende Messzeitpunkte**

Als Grundlage dient wiederum der unter Abschnitt 4.3.1.1 beschriebene Datensatz (N=462). Da bei dieser Fragestellung jedoch explizit die frühen Veränderungen von Interesse sind, werden lediglich die zusammengesetzten Messzeitpunkte 1 bis 3 (Eingangsmessung sowie erste und zweite Zwischenerhebung) berücksichtigt. Für Patienten mit verhaltenstherapeutischer Langzeittherapie entspricht dies etwa der ersten Therapiehälfte. Falls bis zu diesem Zeitpunkt die Therapie bereits beendet ist, wird die Abschlusserhebung anstelle der zweiten Zwischenerhebung einbezogen. Die Nutzung von imputierten Daten erlaubt nicht die Zuordnung von posterioren Klassenzugehörigkeitswahrscheinlichkeiten. Da diese jedoch als Prädiktor in spätere Regressionsanalysen einfließen sollen, werden wie-

derholt Growth Mixture Analysen durchgeführt, in die in diesem Fall jedoch lediglich jene Patienten mit vollständigen Daten für die Messzeitpunkte 1 bis 3 eingehen (listenweiser Falleinschluss). Es resultiert ein Datensatz der Größe  $N=321$ . Vor einer weiteren Verwendung der posterioren Klassenzugehörigkeitswahrscheinlichkeiten werden die Ergebnisse der beiden Blöcke von Growth Mixture Analysen verglichen im Hinblick auf Anzahl, Verlauf und relative Größe der resultierenden Klassen.

#### 4.3.2.2 GMM-Vorgehen

Ausgehend von den in Abschnitt 4.3.1.2 beschriebenen Grundmodellen für BSI und IIP wird jeweils eine latente kategoriale Variable eingefügt. Beginnend mit einer latenten Klasse (was den oben genannten konventionellen Wachstumsanalyse entspricht), werden der kategorialen latenten Variablen im Modell schrittweise mehr Klassen zugeordnet, so lange bis die Anzahl an Klassen gefunden wurde, die die Daten am besten repräsentiert. Zum Modellvergleich und zur Festlegung der optimalen Anzahl an Klassen werden verschiedene Kriterien herangezogen: Das *Bayes Information Criterion* (BIC; Schwartz, 1978), der *Lo-Mendell-Rubin-Test* (LMR; Lo, Mendell, & Rubin, 2001), das *Entropie-Kriterium* (Ramaswamy, DeSabro, Reibstein, & Robinson, 1993) und der *Bootstrap-Likelihood-Ratio-Differenztests* (BLRT; u. a. Nylund, Asparouhov & Muthén, 2007). Eine genauere Beschreibung dieser Kriterien und Tests findet sich in Kapitel 4.2.2. Als weiteres Kriterium zur Bestimmung der optimalen Klassenanzahl nennen Muthén & Muthén (2000) außerdem die praktische Nützlichkeit beziehungsweise die inhaltliche Bedeutsamkeit der Hinzuziehung weiterer Klassen für die entsprechende Fragestellung. Entscheidungskriterien können hier zum Beispiel die Ähnlichkeit zweier Klassen oder die Anzahl der Individuen pro Klasse sein. Im Rahmen dieser Arbeit werden GMMs mit Hilfe der Software Mplus (Version 6.1; Muthén & Muthén, 2010) berechnet, welche für die Bestimmung der Klassen ein iteratives Maximum-Likelihood-Verfahren nutzt.

### **4.3.3 Regressionsanalysen zur Vorhersage des Therapieergebnisses und der Anzahl Sitzungen unter Einbezug von Komorbidität und Klassenzugehörigkeit**

Zur Beantwortung der Frage, inwieweit die identifizierte Klassenzugehörigkeit und Information über Komorbidität nutzbar sind zur Vorhersage des Therapieergebnisses und -aufwandes (definiert durch die Anzahl durchgeführter Therapiesitzungen) werden hierarchische Regressionsanalysen durchgeführt. In einem ersten Block von Regressionsanalysen wird die über Growth Mixture Modeling identifizierte Klassenzugehörigkeit als Prädiktor einbezogen; diese Analysen beruhen auf dem unter Abschnitt 4.3.2.1 vorgestellten Datensatz von Patienten mit vollständigen Daten für die Verlaufserhebungen 1 bis 3 (N=321). Als Kriteriumsvariablen werden hier der BSI-Differenzwert zwischen Therapiebeginn und -ende beziehungsweise Katamnese sowie die Anzahl an Sitzungen einbezogen. Als Prädiktoren werden schrittweise in die Regressionsgleichung aufgenommen: BSI-Eingangswert, Alter und Geschlecht als Kontrollvariablen, für die Vorhersage des Therapieergebnisses die Anzahl durchgeführter Sitzungen, die über GMM identifizierte wahrscheinlichste Klassenzugehörigkeit (als dummy-codierte Variablen) und die Information über Komorbidität (als dummy-codierte Variablen, wobei „keine Komorbidität“ als Referenzkategorie gilt).

In einem zweiten Block von Regressionsanalysen wird die Klassenzugehörigkeit als Prädiktor nicht berücksichtigt; hier ist grundlegend der Ausgangsdatsatz von Patienten, deren Therapeuten verhaltenstherapeutisch ausgebildet sind und die eindeutig mindestens einer Diagnosegruppe zuzuordnen sind (N=967). Als abhängige Variablen gehen die Endbeziehungsweise Katamnesewerte der zentralen Instrumente BSI und IIP sowie des BDI als das am häufigsten angewendete störungsspezifische Instrument in die Analysen ein. Im Rahmen hierarchischer Regressionsanalysen sollen blockweise folgende Prädiktoren einbezogen werden: Eingangsbelastung im jeweiligen Belastungsmaß, Alter und Geschlecht, Anzahl durchgeführter Therapiesitzungen und Komorbiditätsmuster (als Dummyvariablen werden codiert keine Komorbidität vs. Komorbidität mit Persönlichkeitsstörung vs. Komorbidität zwischen Angst- und affektiven Störungen ohne Persönlichkeitsstörung, andere Komorbidität). Darüber hinaus wird die Anzahl durchgeführter Therapiesitzungen vorhergesagt durch die Prädiktorblöcke Eingangsbelastung im BSI, Alter und Geschlecht und Komorbiditätsmuster.

## 5 ERGEBNISSE

In diesem Kapitel werden die Ergebnisse getrennt für die Statusfragestellungen (Hypothesen 1 bis 4) und die Verlaufs- und Ergebnisfragestellungen (Hypothesen 5 bis 10) berichtet.

### 5.1 Zu den Hypothesen 1 bis 4

Zunächst werden die Resultate zu den Fragestellungen dargestellt, die sich auf ein verändertes Diagnoseverhalten unter Anwendung von Diagnosechecklisten beziehen, sowie auf Art und Muster der Komorbidität und deren Konsequenzen für die Belastung zu Beginn einer Therapie.

#### 5.1.1 Ergebnisse zur Hypothese 1

Von den insgesamt eingehenden 1154 Patienten liegt für 178 Patienten (15.4%) keine diagnostische Information vor (s. Tabelle 4). Der Anteil an Patienten ohne Diagnose ist in der IG deutlich erhöht (20.7% vs. 7.1%;  $\chi^2_{(df=1)}=39.080$ ;  $p=.000$ ;  $\Phi=.184$ ). Patienten mit und ohne diagnostische Information unterscheiden sich nicht bedeutsam im Hinblick auf Alter ( $t_{(1146)}=1.274$ ,  $p=.203$ ;  $1-\beta=.99$ ), Geschlecht ( $\chi^2_{(df=1)}=.103$ ,  $p=.749$ ;  $\Phi=.009$ ) und Eingangsbelastung, gemessen am BSI-GSI ( $t_{(1071)}=-1.148$ ,  $p=.251$ ;  $1-\beta=.99$ ), jedoch im Anteil bereits beendeter Therapien (43.4% bei fehlender, 55.6% bei vorhandener diagnostischer Information;  $\chi^2_{(df=1)}=9.276$ ;  $p=.002$ ;  $\Phi=.090$ ).

Fast die Hälfte aller Patienten (48.2%) erhielten lediglich eine Diagnose (s. Tabelle 4), gut ein Viertel (25.2%) erhielt zwei, 10.1% drei und nur zwölf Patienten (1.0%) vier Diagnosen. Sofern diagnostische Information vorliegt, ist der Anteil an Mehrfachdiagnosen in der IG im Vergleich zur KG deutlich höher. Der Mittelwertunterschied für die Anzahl an Diagnosen pro Patient wird zwischen IG ( $M=1.70$ ,  $SD=.79$ ) und KG ( $M=1.41$ ,  $SD=.65$ ) signifikant ( $t_{(974)}=5.923$ ,  $p<.01$ ;  $d=.21$ ).

Tabelle 4: Anzahl vergebener Diagnosen pro Patient

		Gesamt (N=1154)	IG (n=704)	KG (n=450)
keine diagnostische Information		178 (15.4%)	146 (20.7%)	32 (7.1%)
diagnostische Information vorhanden		976 (84.6%)	558 (79.3%)	418 (92.9%)
Anzahl Diagnosen	1	556 (48.2%)	274 (38.9%)	282 (62.7%)
	2	291 (25.2%)	192 (27.3%)	99 (22.0%)
	3	117 (10.1%)	80 (11.4%)	37 (8.2%)
	4	12 (1.0%)	12 (1.7%)	0
mittlere Anzahl Diagnosen (SD) sofern mind. eine Diagnose vorhanden		1.57 (.75) (N=976)	1.70 (.79) (N=558)	1.41 (.65) (N=418)

*Anmerkung:* In diese Tabelle gehen auch Diagnosen mit unvollständigen oder fehlenden ICD-10-Nachkommastellen ein

Tabelle 5: Häufigkeiten und relative Anteile der Haupt- und Nebendiagnosen zu Therapiebeginn

	Gesamt	IG	KG
Hauptdiagnose	976 (100%)	558 (100%)	418 (100%)
davon mit Zweitdiagnose	420 (43.0%)	284 (50.9%)	136 (32.5%)
davon mit Drittdiagnose	129 (30.7%)	92 (32.4%)	37 (27.2%)
davon mit Viertdiagnose	12 (9.3%)	12 (13.0%)	0

*Anmerkung:* In diese Tabelle gehen auch Diagnosen mit unvollständigen oder fehlenden ICD-10-Nachkommastellen ein

Die Anzahl und der relative Anteil weiterer Diagnosen bei Vorliegen von bereits einer, zwei oder drei Diagnosen ist in Tabelle 5 zusammengefasst. Hier wird der Unterschied zwischen IG und KG noch einmal deutlich: während etwa bei Vorliegen einer Hauptdiagnose in der IG rund 50% eine Zweitdiagnose erhalten, sind dies in der KG nur 32.5%. Von den 420 Patienten mit mindestens zwei Diagnosen sind 83 (8.5%) der Gruppe „komorbid mit Persönlichkeitsstörung“ zuzuordnen, 114 (11.7%) der Gruppe „komorbide Angst- und affektive Störungen“ und 223 (22.8%) der Gruppe „komorbide, andere“. Diese Anteile

unterscheiden sich nicht bedeutsam von denjenigen Patienten der TK-Studie, die eine Diagnose erhielten, jedoch nicht von einem verhaltenstherapeutisch ausgebildeten Therapeuten behandelt wurden ( $\chi^2_{(df=3)}=2.550$ ;  $p=.47$ ;  $CI = .042$ ).

Zur Frage, ob unter Verwendung der Internationale Diagnosen Checklisten (IDCL) und damit in der IG mehr Komorbidität resultiert als in der KG, wurde außerdem eine logistische Regressionsanalyse gerechnet, wobei Alter, Geschlecht und BSI-GSI als Kontrollvariablen dienten. Ergebnisse sind in Tabelle 6 zusammengefasst. Auch unter Einbezug der Eingangsbelastung, welche als bedeutsamer Prädiktor für Komorbidität resultiert, erreicht die Information zur Gruppenzugehörigkeit Signifikanz. Die Zugehörigkeit zur Kontrollgruppe im Vergleich zur Interventionsgruppe resultiert bedeutsam seltener in Komorbidität.

Tabelle 6: Ergebnisse der logistischen Regressionsanalyse zur Vorhersage von Komorbidität

	$\beta$	OR	95% CI
Alter	-.001	.999	.998-1.000
Geschlecht <sup>1</sup>	-.115	.891	.668-1.189
BSI-GSI	.318**	1.375	1.124-1.681
Gruppe <sup>2</sup>	-.758**	.469	.354-.620

<sup>1</sup>als Referenzkategorie diente männliches Geschlecht,

<sup>2</sup>als Referenzkategorie diente Interventionsgruppe

Ergebnisse zu den störungsbezogenen Fragestellungen sind in Tabelle 7 zusammengefasst. Nach Kontrolle für Alter, Geschlecht und globale psychische Belastung zu Beginn der Therapie resultieren für einige Störungsgruppen Unterschiede bezüglich ihrer Diagnosehäufigkeit unter strukturierter Diagnostik (IG) und nach nicht strukturierter Diagnoseerfassung (KG). So wird etwa die soziale Phobie in der IG zu einem größeren Anteil diagnostiziert als in der KG (für die Diagnose in der KG im Vergleich zur IG ergibt sich ein Odds Ratio (OR) =.486,  $p<.01$ ). Gleiches gilt auch für spezifische Phobien (OR=.353,  $p<.01$ ), nicht jedoch für den Bereich Panikstörung/Agoraphobie. Während die Diagnose einer gemischten Angst- und depressiven Störung in der IG bedeutsam seltener als in der KG vergeben wird (OR=2.744,  $p<.01$ ), bestätigt sich diese Annahme nicht für die Gruppe der Anpassungsstörungen. Das häufigere Vorkommen von Persönlichkeitsstörungen in der IG gegenüber der KG erreicht hingegen Signifikanz (OR=.510,  $p<.01$ ). Darüber hinaus erge-

ben sich bedeutsame Unterschiede hinsichtlich der Diagnose affektiver Störungen: sowohl rezidivierende depressive Störungen (OR=.538,  $p<.01$ ) als auch Dysthymia (OR=.512,  $p<.01$ ) werden mit einem bedeutsam höheren Anteil nach Verwendung der IDCL-Checklisten diagnostiziert. Die Diagnose Depressive Episode findet sich hingegen etwas häufiger in der KG, ohne dass dieser Unterschied statistische Bedeutsamkeit erreicht.

Tabelle 7: Anzahl Patienten mit mindestens einer Nennung der Diagnose

Diagnose nach ICD-10	N (967 <sup>1</sup> )	IG (n=554)	KG (n=413)	OR	95% CI	99% CI
Depressive Episode	258	141 (25.5%)	117 (28.3%)	1.142	.834-1.546	
Rezidivierende depressive Störung	183	123 (22.2%)	60 (14.5%)	.538	.372-.779*	**
Dysthymia	97	69 (12.5%)	28 (6.8%)	.512	.315-.834*	**
Bipolare Störung/Manie	10	6 (1.1%)	4 (1.0%)	.981	.273-3.522	
Panikstörung/ Agoraphobie	88	51 (9.2%)	37 (9.0%)	.972	.610-1.548	
soziale Phobie	89	64 (11.6%)	25 (6.1%)	.486	.291-.813*	**
spezifische Phobie	47	38 (6.9%)	9 (2.2%)	.353	.168-.741*	**
GAS	59	41 (7.4%)	18 (4.4%)	.508	.271-.949*	
Angst/Depression gemischt	33	11 (2.0%)	22 (5.3%)	2.744	1.288-5.843*	**
Zwangsstörung	38	24 (4.3%)	14 (3.4%)	.890	.444-1.784	
Akute Belastungsstörung	13	7 (1.3%)	6 (1.5%)	1.371	.455-4.126	
PTBS	40	31 (5.6%)	9 (2.2%)	.431	.202-.920*	
Anpassungsstörungen	205	107 (19.3%)	98 (23.7%)	1.332	.962-1.844	
Somatoforme Störungen	89	53 (9.6%)	36 (8.7%)	.972	.613-1.543	
Essstörungen	44	28 (5.1%)	16 (3.9%)	.710	.354-1.425	
Persönlichkeitsstörungen	85	60 (10.8%)	25 (6.1%)	.510	.297-.876*	
Sonstige Störungen	110	59 (10.6%)	51 (12.3%)	1.269	.839-1.918	

<sup>1</sup>Die veränderte Stichprobengröße kommt dadurch zustande, dass bei einigen Patienten die vorliegende diagnostische Information unzureichend war, um sie eindeutig einer der Störungsgruppen zuordnen zu können.

## 5.1.2 Ergebnisse zu den Hypothesen 2 und 3

In diesem Abschnitt sind die Ergebnisse zu den bivariaten (Hypothese 2) und multivariaten Zusammenhängen (Hypothese 3) zwischen Kategorien psychischer Störung zusammengestellt.

### 5.1.2.1 Deskriptive Darstellung der Diagnosekombinationen und bivariate Zusammenhänge

Eine Übersicht zum gemeinsamen Auftreten von Störungsdiagnosen findet sich in Tabelle 8. Es wird deutlich, dass, wie in einer ambulanten Stichprobe zu erwarten, affektive Störungen und Angststörungen, aber auch Anpassungsstörungen, besonders häufig diagnostiziert werden. Affektive Störungen werden zu allen weiteren Störungen besonders häufig komorbide diagnostiziert. So erhalten zum Beispiel fast ein Viertel der Patienten mit sozialer Phobie (23.6%) zusätzlich die Diagnose einer depressiven Episode. Hohe Anteile komorbider affektiver Störungen ergeben sich auch bei generalisierter Angststörung, bei Posttraumatischer Belastungsstörung, somatoformen Störungen und Persönlichkeitsstörungen. Von den insgesamt 120 theoretisch möglichen bivariaten Diagnosekombinationen finden sich 90 tatsächliche Kombinationsmöglichkeiten in den Daten wieder. Von diesen werden jedoch nur 46 Kombinationen fünf Mal oder häufiger vergeben.

Die tetrachorischen Korrelationen dieser Kombinationen sind in Tabelle 9 wiedergegeben. Der tetrachorische Korrelationskoeffizient sollte jedoch nur dann interpretiert werden, wenn alle Felder der Vierfeldertafel mindestens eine Häufigkeit von 5 aufweisen (Wirtz & Nachtigall, 1998).

Tabelle 8: Häufigkeiten und relative Anteile der bivariaten Störungskombinationen

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Depressive Episode		0	5	0	12	21	6	14	2	8	5	7	13	16	8	16
			5.2%		13.6%	23.6%	12.8%	23.7%	6.1%	21.1%	38.5%	17.5%	5.3%	18.0%	13.2%	18.8%
2 Rezidivierende depressive Störung	0		6	0	12	11	8	16	1	7	1	10	4	17	9	30
			6.2%		13.6%	12.4%	17.0%	27.1%	3.0%	18.4%	7.7%	25.0%	2.0%	19.1%	20.5%	35.3%
3 Dysthymia	5	6		0	9	10	3	4	0	1	0	3	5	9	6	7
	1.9%	3.3%			10.2%	11.2%	6.4%	6.8%	0	2.6%	0	7.5%	2.4%	10.1%	12.6%	8.2%
4 Bipolare Störung/Manie	0	0	0		0	2	0	0	0	1	0	1	1	0	0	2
						2.2%				2.6%	0	2.5%	0.5%			2.4%
5 Panikstörung/Agoraphobie	12	12	9	0		6	4	5	0	2	1	2	9	5	0	6
	4.7%	6.6%	9.3%			6.7%	8.5%	8.5%		5.3%	7.7%	5.0%	4.4%	5.6%		7.1%
6 soziale Phobie	21	11	10	2	6		4	7	0	7	1	4	7	3	5	9
	8.1%	6.0%	10.3%	20.0%	6.8%		8.5%	11.9%		18.4%	7.7%	10.0%	3.4%	3.4%	11.6%	10.6%
7 spezifische Phobie	6	8	3	0	4	4		1	0	2	0	4	8	2	0	3
	2.3%	4.4%	3.1%		4.5%	4.5%		1.7%		5.3%	0	10.0%	3.9%	2.2%		3.5%
8 Generalisierte Angststörung	14	16	4	0	5	7	1		0	3	0	1	7	7	3	6
	5.4%	8.7%	4.1%		5.7%	7.9%	2.1%			7.9%	0	2.5%	3.4%	7.9%	6.8%	7.1%
9 Angst/Depression gemischt	2	1	0	0	0	0	0	0		0	0	0	1	4	0	2
	0.8%	0.5%											0.5%	4.5%		2.4%
10 Zwangsstörung	8	7	1	1	2	7	2	3	0		0	4	2	2	0	1
	3.1%	3.8%	1.0%	10.0%	2.3%	7.9%	4.3%	5.1%				10.0%	1.0%	2.2%		1.2%
11 Akute Belastungsstörung	5	1	0	0	1	1	0	0	0	0		0	0	0	0	1
	1.3%	0.5%			1.1%	1.1%										1.2%
12 PTBS	7	10	3	1	2	4	4	1	0	4	0		2	2	2	4
	2.7%	5.5%	3.1%	10.0%	2.3%	4.5%	8.5%	1.7%		10.5%			1.0%	2.2%	4.5%	4.7%
13 Anpassungsstörungen	13	4	5	1	9	7	8	7	1	2	0	2		11	1	10
	5.0%	2.2%	5.2%	10.0%	10.2%	7.9%	17.0%	11.9%	3.0%	5.3%		5.0%		12.4%	2.3%	11.8%
14 Somatoforme Störungen	16	17	9	0	5	3	2	7	4	2	0	2	11		4	7
	6.2%	9.3%	9.3%		5.7%	3.4%	4.3%	11.9%	12.1%	5.3%		5.0%	5.4%		9.1%	8.2%
15 Essstörungen	8	9	6	0	0	5	0	3	0	0	0	2	1	4		5
	3.1%	4.9%	6.2%			5.6%		5.1%				5.0%	0.5%	4.5%		5.9%
16 Persönlichkeitsstörungen	16	30	7	2	6	9	3	6	2	1	1	4	10	7	5	5
	6.2%	16.4%	7.2%	20.0%	6.7%	10.1%	6.4%	10.2%	6.1%	2.6%	7.7%	10.0%	4.9%	7.9%	11.4%	11.4%
GESAMT	258	183	97	10	88	89	47	59	33	38	13	40	205	89	44	85

Tabelle 9: Tetrachorische Korrelationskoeffizienten der bivariaten Diagnosekombinationen (Standardfehler)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Depressive Episode																
2. Rezidivierende depr. Störung	-.82*															
3. Dysthymia	-.47*	-.32*														
4. Bipolare Störung/Manie	-.37*	-.29	-.15													
5. Panikstörung/Agoraphobie	-.24*	-.11	.02	-.13												
6. soziale Phobie	-.05	-.14	.02	.23	-.07											
7. spezifische Phobie	-.22*	-.02	-.08	-.02	.01	.00										
8. GAS	-.04	.15	.08	-.06	.00	.09	-.15									
9. Angst/Depression gemischt	-.34*	-.34*	-.35*	.05*	-.33*	-.34*	-.22	-.26								
10. Zwangsstörung	-.07*	.01	-.21	.26	-.08	.22	.06	.09	-.18							
11. Akute Belastungsstörung	.14	-.14	-.20	.20	.03	.03	-.06	-.10	.00	-.02						
12. PTBS	-.13	.11	-.04	.25	-.10	.05	.21*	-.11	-.19	.26	-.03					
13. Anpassungsstörungen	-.52*	-.60*	-.39*	.21	-.23*	-.30*	-.06	-.17	-.36*	-.31*	-.35*	-.32*				
14. Somatoforme Störungen	-.15	.01	.01	-.13	-.12	-.22*	-.14	.09	.10	-.09	-.18	-.10	-.18*			
15. Essstörungen	-.12	.04	.10	.00	-.38*	.08	-.27	.06	-.21	-.23	-.05	.06	-.42*	.02		
16. Persönlichkeitsstörungen	-.13	.29*	-.05	.24	-.06	.06	-.05	.06	-.05	-.18	-.04	.06*	-.19*	-.02	.09	
	-.05)	-.07)	-.04)	-.01)	-.04)	-.01)	-.05)	-.01)	-.05)	-.08)	-.05)	-.02)	-.07)	-.03)	-.02)	-.03)

Anmerkung: \* $p < .05$ ; die Angaben sind grau dargestellt, sofern die Schnittmenge der Störungsgruppen  $< 5$  beträgt

Zwölf Diagnosekombinationen erreichen statistische Signifikanz, wovon jedoch nur eine auf einen bedeutsamen positiven Zusammenhang hinweist, nämlich die Kombination aus rezidivierender depressiver Störung und Persönlichkeitsstörungen ( $r_{\text{tel}}=.29, p<.05$ ). Inwieweit sich Personen mit Komorbidität dieser beiden Störungsbilder unterscheiden von solchen Personen, die nur eine dieser Diagnosen aufweisen, wird auch bei den folgenden Fragestellungen gesondert berücksichtigt. Bedeutsame negative Zusammenhänge ergeben sich zwischen Dysthymia und anderen unipolaren affektiven Störungen, aber auch zwischen depressiver Episode und Panikstörung/Agrophobie beziehungsweise spezifischer Phobie und in besonders großem Ausmaß zwischen Anpassungsstörungen und anderen Diagnosen (s. Tabelle 9).

### 5.1.2.2 Ergebnisse der Latent Class Analysen

In die Analysen gingen die 16 oben aufgeführten Diagnosegruppen plus Restkategorie ein. Es wurden Modelle für ein bis sechs latente Klassen getestet (s. Tabelle 10). Gewählt wurde das Modell mit fünf Klassen.

Tabelle 10: Ergebnisse der Latent Class Analysen

Klassen	H0 loglike- lihood	Anzahl freier Pa- rameter	BIC	Entropie	LMR	BLR
1	-5.239.813	19	10.610.235	-	-	-
2	-4.410.245	36	9.067.961	1.000	0	0
3	-4.282.005	55	8.942.091	1.000	0	0
4	-4.224.355	74	8.957.400	0.987	0	0
5	-4.185.999	93	9.011.299	0.987	0	0
6	-4.152.314	112	9.074.538	0.973	0,9867	0

*Anmerkung:* BIC = Bayes Information Criterion; LMR = Lo-Mendell-Rubin-Test; BLR = Bootstrap-Likelihood-Ratio-Differenztest

Trotz eines leichten Anstiegs des BIC bei Einfügen einer vierten und fünften Klasse sprechen die beiden Indizes zur Beurteilung des relativen Modellfits, LMR und BLR, für die Aufnahme weiterer Klassen in das Modell. Erst im 6-Klassen-Modell zeigt der LMR-Test an, dass ein Modell mit einer Klasse weniger die beobachteten Daten besser abbildet. An-

teile der jeweiligen Klassen an der Gesamtstichprobe und die bedingte Wahrscheinlichkeit des Vorliegens der Diagnosen bei gegebener Klassenzugehörigkeit sind in Tabelle 11 wiedergegeben. Eine grafische Darstellung der bedingten Wahrscheinlichkeiten, interpretierbar als die Prozentzahl der Klassenmitglieder, die die Kriterien für eine spezifische Diagnose erfüllen, für die 5-Klassen-Lösung findet sich in Abbildung 2.

Tabelle 11: Anteil der Latenten Klassen und bedingte Wahrscheinlichkeiten

	Latente Klasse	C1	C2	C3	C4	C5
	Anteil der latenten Klasse	18,5	32,6	26,5	3,3	19,1
Latent Class Scores: bedingte Wahrscheinlichkeit, bei Zugehörigkeit zur entsprechenden Klasse die Diagnose zu erhalten	Depressive Episode	0	0	1	0,06	0
	Rezidivierende depr. Störung	1	0	0	0	0,02
	Dysthymia	0,03	0,26	0,02	0	0,02
	Bipolare Störung/Manie	0	0,03	0	0	0,01
	Panikstörung/Agoraphobie	0,07	0,19	0,05	0	0,03
	Soziale Phobie	0,06	0,17	0,08	0	0,03
	Spezifische Phobie	0,05	0,08	0,02	0	0,04
	Generalisierte Angststörung	0,09	0,07	0,06	0	0,03
	Angst/Depression gemischt	0,01	0	0	1	0
	Zwangsstörung	0,04	0,07	0,03	0	0,01
	Akute Belastungsstörung	0,01	0,02	0,02	0	0
	PTBS	0,06	0,07	0,03	0	0,01
	Anpassungsstörungen	0	0,02	0,05	0,03	1
	Somatoforme Störungen	0,09	0,14	0,06	0,13	0,05
	Essstörungen	0,05	0,08	0,03	0	0,01
Persönlichkeitsstörungen	0,17	0,09	0,06	0,06	0,05	
Sonstige Störungen	0,13	0,15	0,09	0,06	0,07	

In Klasse 1 fallen 179 Personen (18.5%). Alle diese Personen erhielten die Diagnose einer rezidivierenden depressiven Störung. Als relevante weitere Störungen sind allen voran die Persönlichkeitsstörungen (17%), aber auch die generalisierte Angststörung (9%) und die Gruppe der sonstigen Störungen (13%) zu nennen. Die höchste Zugehörigkeitswahrscheinlichkeit für Klasse 2 ergibt sich für 311 Personen (32.6%). Diese Klasse ist gekennzeichnet durch vergleichsweise geringe bedingte Wahrscheinlichkeiten für alle Diagnosegruppen. In

Klasse 2 besteht jedoch von allen Klassen die größte Wahrscheinlichkeit für die Diagnosen Dysthymia (26%), Panikstörung/Agoraphobie (19%), soziale Phobie (17%) und somatoforme Störungen (14%). In Klasse 3 fallen 256 Personen und damit gut ein Viertel (26.5%) der Stichprobe. Mitglieder dieser Klasse erhalten zu 100% die Diagnose einer depressiven Episode, die Wahrscheinlichkeit für andere, weitere Störungen ist sehr gering. Klasse 4 stellt mit 32 Personen (3.3%) die kleinste Klasse dar. Sie enthält fast alle Personen mit der Diagnose Angst und depressive Störung gemischt. Die Wahrscheinlichkeit für somatoforme Störungen (13%) ist hier im Vergleich zu den anderen Klassen erhöht. In Klasse 5 schließlich fallen 189 Personen (19.1%). Diese Klasse ist vor allen Dingen charakterisiert durch die Diagnose Anpassungsstörung, während die Wahrscheinlichkeiten für andere Diagnosen minimal sind. Insgesamt finden sich in den Daten 183 unterschiedliche Muster von Diagnosekombinationen.

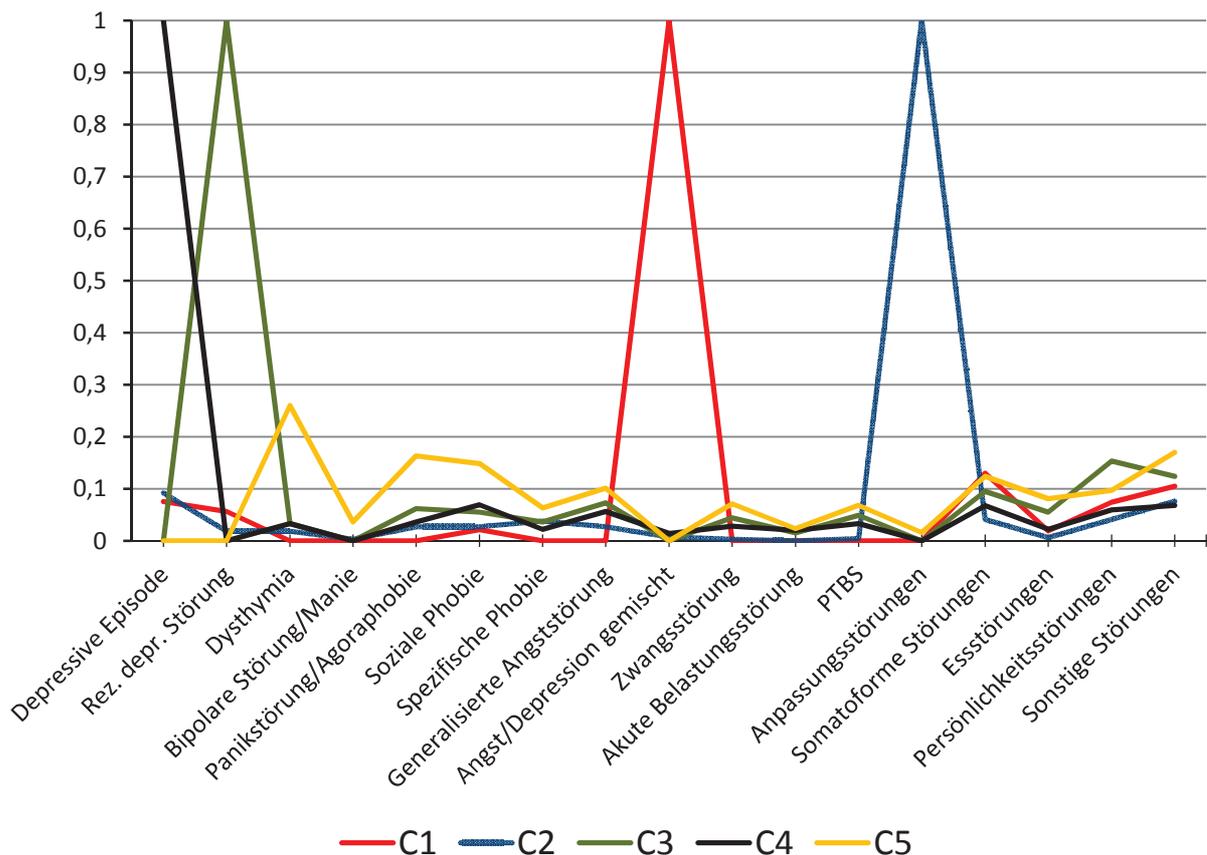


Abbildung 2: 5-Klassen-Lösung der Latent Class Analysen

*Abhängigkeit der identifizierten Klassen von der Art der Diagnosestellung*

In die Latent Class Analysen ging jeweils die Gruppenzugehörigkeit (IG/KG) als Prädiktor ein. Dabei ergeben sich bedeutsame Zusammenhänge zwischen Gruppen- und Klassenzugehörigkeit. So führt die Zugehörigkeit zur KG im Vergleich zur IG zu einem Abfall der Zugehörigkeitswahrscheinlichkeit für Klasse 1 im Vergleich zu Klasse 3 ( $\beta=-.506$ ;  $p<.05$ ), Klasse 4 ( $\beta=-1.498$ ;  $p<.01$ ) und Klasse 5 ( $\beta=-.654$ ;  $p<.01$ ). Dagegen führt die Zugehörigkeit zu KG im Vergleich zur IG zu einem Anstieg der Zugehörigkeitswahrscheinlichkeit zu Klasse 4 im Vergleich zu Klasse 2 ( $\beta=1.179$ ;  $p<.01$ ), Klasse 3 ( $\beta=.992$ ;  $p<.05$ ) und Klasse 5 ( $\beta=.845$ ;  $p<.05$ ).

Die unterschiedliche Verteilung der Gruppen auf die Klassen wird signifikant ( $\chi^2_{(df=4)}=19.230$ ,  $p<.01$ ;  $CI=.141$ , s. a. Tabelle 12), ebenso wie die Verteilung der Klassen auf Komorbiditätsmuster ( $\chi^2_{(df=12)}=102.435$ ,  $p<.001$ ;  $CI=.188$ ). Dabei ist Klasse 1 gekennzeichnet durch die insgesamt höchste Komorbiditätsrate (61.5%) und durch den im Vergleich zu den anderen Klassen mit Abstand größten Anteil „komorbide mit Persönlichkeitsstörung“ (16.8%). Klasse 2 ist gekennzeichnet durch einen relativ geringen Anteil an Komorbidität (36.3%); wenn sich hier Mehrfachdiagnosen finden, dann vor allem die Gruppe „komorbide, andere“. Klasse 3 weist einen mittleren Anteil an Komorbidität auf (46.9%), wobei die Gruppen „komorbide Angst- und affektive Störungen“ und „komorbide, andere“ gleich häufig auftreten (jeweils 20.3%). Eine besonders geringe Komorbiditätsrate (31.2%) findet sich wie zu erwarten in Klasse 4. Klasse 5 schließlich weist ebenfalls eine geringe Rate an Mehrfachdiagnosen auf (35.4%), wobei „komorbide Angst- und affektive Störungen“ nicht vertreten sind und 29.6% der Gruppe „komorbide, andere“ zuzuordnen ist.

Tabelle 12: Verteilung der Patienten in IG und KG auf die Latenten Diagnoseklassen

	Latente Diagnose-Klasse					Gesamt
	C1	C2	C3	C4	C5	
IG	120 (21.7%)	185 (33.4%)	141 (25.5%)	10 (1.8%)	98 (17.7%)	554
KG	59 (14.3%)	126 (30.5%)	115 (27.8%)	22 (5.2%)	91 (22.0%)	413
Gesamt	179 (18.5%)	311 (32.2%)	256 (26.5%)	32 (3.3%)	181 (18.7%)	967

### 5.1.3 Ergebnisse zur Hypothese 4

In den folgenden Abschnitten wird getrennt für Selbst- und Fremdberturmsmaße dargestellt, inwieweit Personen mit Mehrfachdiagnosen sich von solchen mit nur einer Diagnose unterscheiden.

#### 5.1.3.1 Komorbidität und selbstberichtete Beeinträchtigung

Um ein zusammengesetztes Beeinträchtigungsmaß zu bilden, wurden die drei eingehenden Selbstauskunft-Instrumente (BSI-GSI, Franke, 2000; IIP-D, Horowitz et al., 1988 und SF-12, Bullinger, & Kirchberger, 1998) zunächst per Lineartransformation auf eine einheitliche Metrik von 0 (starke Beeinträchtigung) bis 100 (keine Beeinträchtigung) gebracht und im Anschluss der Mittelwert berechnet. Voraussetzung für die Zusammenfassung zu einem Gesamtwert war jedoch, dass mindestens BSI-GSI und IIP-D vorhanden waren. Das zusammengesetzte Ergebnismaß für den Selbstbericht konnte für  $N=917$  Personen berechnet werden. Der Mittelwert beträgt 54.07 ( $SD=10.37$ ), bei einem Range von 17.39 bis 95.67. Die interne Konsistenz dieses neu gebildeten Maßes liegt bei Cronbachs Alpha=.698. Die Verteilungskennwerte für Interventions- und Kontrollgruppe sind im Anhang (S. 174) wiedergegeben.

Das zusammengesetzte Ergebnismaß geht als abhängige Variable in eine Varianzanalyse ein. Die deskriptiven Statistiken der zweifaktoriellen Varianzanalysen sind in Tabelle 13 und Tabelle 14 abgebildet. Werden als Gruppierungsfaktoren die Zugehörigkeit zu den vier Komorbiditätsmustern (Faktor Komorbidität) und die Zugehörigkeit zu Interventions- und Kontrollgruppe (Faktor Gruppe) gewählt, so resultiert ein signifikanter Haupteffekt „Komorbidität“ ( $F_{(3,909)}=5.860, p<.01, \eta^2=.019$ ). Im Post-hoc-Vergleich ergeben sich bedeutsame Unterschiede zwischen den Gruppen „keine Komorbidität“ und „komorbide Persönlichkeitsstörung“, sowie „keine Komorbidität“ und „komorbide Angst- und affektive Störung“ und zwischen „komorbide Persönlichkeitsstörung“ und „komorbide, andere“. Die größte selbsteingeschätzte Belastung weisen Personen mit komorbider Persönlichkeitsstörung auf. Sowohl der Haupteffekt für den Faktor Gruppe, als auch die Wechselwirkung zwischen den Faktoren Komorbidität und Gruppe erreichen keine statistische Bedeutsamkeit ( $1-\beta=.99$ ). Wird statt des Faktors Komorbidität die Zugehörigkeit zu latenten Diagnoseklassen (Faktor Klasse) einbezogen, so ergibt sich ein ähnliches Ergebnis: der Hauptef-

fekt „Klasse“ wird hoch signifikant ( $F_{(4,907)}=8.779$ ,  $p < .001$ ,  $\eta^2=.037$ ), mit bedeutsamen Unterschieden zwischen den Klassen 1 und 2, 1 und 5, sowie 3 und 5. Der Haupteffekt Gruppe und auch die Wechselwirkung Klasse x Gruppe sind ohne Einfluss ( $1-\beta=.99$ ).

Tabelle 13: Deskriptive Statistik zur zweifaktoriellen Varianzanalyse mit den Faktoren Gruppenzugehörigkeit und Latente Diagnoseklasse.

		Latente Diagnose-Klasse				
		C1	C2	C3	C4	C5
zusammengesetzte Beeinträchtigung Selbstbericht		N=120	N=184	N=141	N=10	N=98
	IG	M=50.61	M=54.02	M=51.44	M=53.69	M=54.45
		SD=9.05	SD=10.97	SD=9.52	SD=10.15	SD=9.82
		N=47	N=115	N=98	N=20	N=84
	KG	M=47.62	M=55.28	M=53.51	M=50.18	M=55.71
		SD=9.52	SD=11.43	SD=11.11	SD=9.97	SD=8.90
zusammengesetzte Beeinträchtigung Fremdbericht		N=119	N=185	N=140	N=10	N=98
	IG	M=49.37	M=56.39	M=50.13	M=58.60	M=58.52
		SD=12.34	SD=13.70	SD=11.76	SD=7.31	SD=13.86
		N=59	N=126	N=115	N=22	N=91
	KG	M=44.21	M=53.78	M=49.17	M=51.15	M=52.86
		SD=10.83	SD=12.15	SD=12.45	SD=12.02	SD=10.29

### 5.1.3.2 Komorbidität und fremdberichtete Beeinträchtigung

Auch hier wurde ein zusammengesetztes Beeinträchtigungsmaß gebildet, wozu die eingehenden Skalen (GAF, Endicott et al., 1976; BSS, Schepank, 1995 und Problemrelevante Lebensbereiche) zunächst per Linearkombination in die einheitliche Metrik von 0 (starke Beeinträchtigung) bis 100 (keine Beeinträchtigung) gebracht wurden. Sofern mindestens sechs der insgesamt elf Skalen vorlagen, wurde der Mittelwert berechnet. Das zusammengesetzte Ergebnismaß in der Fremdbeurteilung konnte für N=965 Personen berechnet werden. Der Mittelwert liegt bei 52.46 (SD=12.90), bei einem Range von 14.09 bis 88.57. Die interne Konsistenz des zusammengesetzten Ergebnismaßes Fremd liegt bei Cronbachs Alpha=.672.

Tabelle 14: Deskriptive Statistik zur zweifaktoriellen Varianzanalyse mit den Faktoren Gruppenzugehörigkeit und Komorbiditätsmuster.

		Komorbiditätsmuster			
		keine Komorbidität	Komorbidität mit PS	Komorbidität Angst/affektive Störungen	Komorbidität andere Muster
zusammengesetzte Beeinträchtigung Selbstbericht		N=269	N=59	N=76	N=149
	IG	M=54.54	M=48.04	M=50.95	M=52.09
		SD=10.01	SD=9.03	SD=10.38	SD=9.72
		N=244	N=19	N=32	N=69
	KG	M=53.89	M=52.04	M=49.72	M=54.96
		SD=10.40	SD=9.46	SD=11.14	SD=11.93
zusammengesetzte Beeinträchtigung Fremdbbericht		N=269	N=59	N=76	N=148
	IG	M=56.42	M=44.39	M=50.64	M=54.59
		SD=12.99	SD=12.23	SD=12.82	SD=12.94
		N=277	N=24	N=38	N=74
	KG	M=52.40	M=41.24	M=49.83	M=48.34
		SD=11.55	SD=12.06	SD=11.44	SD=12.54

Die Verteilungskennwerte für Interventions- und Kontrollgruppe sind ebenfalls im Anhang (S. 174) wiedergegeben, die Mittelwerte in den Tabellen 13 und 14. Gehen in die Varianzanalysen als Gruppierungsfaktoren die Zugehörigkeit zu Komorbiditätsmustern (Faktor Komorbidität) und die Zugehörigkeit zu Interventions- und Kontrollgruppe (Faktor Gruppe) ein, so resultiert auch für die fremdbeurteilte Beeinträchtigung ein signifikanter Haupteffekt „Komorbidität“ ( $F_{(3,957)}=20.006$ ,  $p<.001$ ,  $\eta^2=.059$ ), wobei sich im Post-hoc-Vergleich bedeutsame Unterschiede zwischen den Gruppen „keine Komorbidität“ und „komorbide Persönlichkeitsstörung“, „keine Komorbidität“ und „komorbide Angst- und affektive Störung“, „komorbide Persönlichkeitsstörung“ und „komorbide Angst- und affektive Störung“ sowie „komorbide Persönlichkeitsstörung“ und „komorbide, andere“ ergeben. Die größte fremdeingeschätzte Belastung weisen Personen mit komorbider Persönlichkeitsstörung auf. Hier ergibt sich auch ein signifikanter Haupteffekt „Gruppe“ ( $F_{(1,957)}=9.718$ ,  $p<.01$ ,  $\eta^2=.010$ ), wobei die Patienten der IG als weniger belastet eingeschätzt werden als die der KG. Die Wechselwirkung Klasse x Gruppe erreicht keine statistische Bedeutsamkeit. Wird statt des Faktors Komorbidität die Information zur Klassenzugehörigkeit in die Analysen einbezogen, so ergibt sich auch hier ein ähnliches Bild. Der

Haupteffekt „Klasse“ wird signifikant ( $F_{(4,955)}=18.337, p<.001, \eta^2=.071$ ), mit bedeutsamen Unterschieden zwischen den Klassen 1 und 2, 1 und 5, 2 und 3, sowie 3 und 5. Auch hier zeigen sich Unterschiede zwischen IG und KG (Haupteffekt Gruppe:  $F_{(1,955)}=14.069, p<.001, \eta^2=.015$ ), mit geringerer Belastung in der IG gegenüber der KG. Die Wechselwirkung Klasse x Gruppe wird auch hier nicht statistisch bedeutsam. Eine Übersicht zur selbst- und fremdberichteten Belastung, differenziert nach Komorbiditätsgruppen und latenten Diagnoseklassen findet sich auch im Anhang (S. 175f., S. 177f.).

### 5.1.3.3 Belastung bei Komorbidität aus rezidivierender depressiver Störung und Persönlichkeitsstörung

Unter Betrachtung allein derjenigen Personen, die eine Kombination aus rezidivierender depressiver Störung und Persönlichkeitsstörung aufweisen gegenüber solchen, die nur eine dieser beiden Diagnosen (entweder in Reinform oder mit anderer Komorbidität) erhalten, ergibt sich folgendes Bild (s. Tabelle 15).

Tabelle 15: Selbst- und fremdberichtete Beeinträchtigung für die Kombination aus rezidivierender depressiver Störung und Persönlichkeitsstörung

	Komorbiditätsmuster		
	rezidivierende depressive Störung ohne Persönlichkeitsstörung	Persönlichkeitsstörung ohne rezidivierende depressive Störung	rezidivierende depressive Störung und Persönlichkeitsstörung
zusammengesetzte Beeinträchtigung Selbstbericht	N=144 M=49.34 SD=9.36	N=53 M=48.25 SD=9.64	N=27 M=51.71 SD=8.83
zusammengesetzte Beeinträchtigung Fremdbbericht	N=152 M=48.94 SD=11.94	N=55 M=45.29 SD=12.21	N=30 M=40.74 SD=10.97

Bezogen auf die selbstberichtete Belastung zeigen sich keine bedeutsamen Unterschiede zwischen den drei untersuchten Gruppen ( $F_{(2,221)}=1.229, p=.295, 1-\beta=.92$ ). In der fremdberichteten Belastung ergibt sich ein bedeutsamer Unterschied zwischen den Gruppen rezidivierende depressive Störung ohne komorbide Persönlichkeitsstörung und der Gruppe mit

einer Kombination aus diesen beiden Diagnosen ( $F_{(2,234)}=6.743$ ,  $p<.001$ ,  $\eta^2=.054$ ), Insbesondere im GAF-Wert spiegeln sich Unterschiede zwischen diesen beiden Gruppen wider (rezidivierende depressive Störung ohne Persönlichkeitsstörung  $N=145$ ,  $M=52.24$ ,  $SD=11.09$ ; mit komorbider Persönlichkeitsstörung  $N=29$ ,  $M=48.31$ ,  $SD=10.52$ ).

## 5.2 Zu den Hypothesen 5 bis 10

Die im Folgenden dargestellten Resultate befassen sich mit den Konsequenzen von Komorbidität auf Therapieverläufe und –ergebnisse in den abhängigen Variablen BSI, IIP und BDI.

### 5.2.1 Verlaufsanalysen - BSI als abhängige Variable

Für die Latent Growth Curve Analysen gilt die Annahme multivariater Normalverteilung, welche partiell durch die Überprüfung der Normalität individueller Variablen untersucht werden kann. Zwar sind die Verteilungen des BSI zu allen vier einbezogenen Messzeitpunkten abweichend von Normalität (s. Anhang, S. 179). Jedoch können im Rahmen der Analysen mit Mplus Schätzer gewählt werden, die für Verletzungen der Normalverteilungsannahme robust sind. In den folgenden Analysen wird der MLR-Schätzer verwendet (Maximum Likelihood Schätzung mit robusten Standardfehlern; Muthén & Muthén, 2010).

#### 5.2.1.1 LGM - Abhängige Variable BSI

Wie in Abschnitt 4.3.1.2 erläutert, wurde zunächst ein Grundmodell bestimmt, welches in der Lage ist, den Verlauf im Ergebnismaß BSI adäquat zu beschreiben. Als konkurrierende Verlaufsmodelle wurden getestet: ein Modell, welches rein durch den Ausgangswert bestimmt wird (intercept-only), ein lineares, loglineares und quadratisches Modell. Wie aus Tabelle 16 ersichtlich ist, weist der RMSEA auf den besten Gesamt-Modell-Fit für das Modell mit loglinearem Verlauf hin, gestützt durch das BIC, welches für das loglineare Modell den geringsten Wert annimmt. Die folgenden Analysen beziehen sich folglich auf die Annahme eines loglinearen Verlaufs im BSI. Alternativ zu dem loglinearen Modell mit freiem intercept und slope wurden Gegenmodelle getestet, in denen einer oder beide

Wachstumsfaktoren fixiert waren. Dies führte nicht zu einer Besserung der Modellpassung (s. Tabelle 16), weshalb das Modell mit freien Parametern beibehalten wurde.

Tabelle 16: Ergebnisse der LGM-Analysen für die abhängige Variable BSI

		loglikelihood (H0)	RMSEA	CFI	BIC	
1	Testung der Verlaufsform	intercept-only	-1.445.163	0.185	0.617	2.927.139
		linear	-1.307.660	0.103	0.925	2.670.539
		loglinear	-1.291.814	0.050	0.983	2.638.847
		quadratisch	-1.282.693	0.064	0.994	2.645.148
2	loglinearer Verlauf - Testung von Gegenmodellen mit Fixierungen	fixed i, random s	-1.530.990	0.276	0.249	3.104.930
		random i, fixed s	-1.297.708	0.053	0.973	2.638.364
		fixed i, fixed s	-1.826.660	0.323	0.000	3.690.134
3	Einbezug von Komorbiditätsmustern als Prädiktor		-1.279.390	0.051	0.976	2.650.813

In diesem Modell sind beide Wachstumsfaktoren signifikant von Null verschieden ( $M_{\text{intercept}}=1.284$ ,  $z=41.591$ ,  $p<.001$ ;  $M_{\text{slope}}=-.791$ ,  $z=-13.050$ ,  $p<.01$ ) und es besteht bedeutsame Variabilität in Ausgangslage und Steigung ( $\text{Var}_{\text{intercept}}=.304$ ,  $z=10.119$ ,  $p<.01$ ;  $\text{Var}_{\text{slope}}=.243$ ;  $z=2.052$ ,  $p=.04$ ). Somit besteht eine signifikante Unterschiedlichkeit der betrachteten Patienten hinsichtlich ihrer latenten Ausgangswerte und Veränderung über die Zeit. Die Korrelation zwischen intercept und slope ist nicht signifikant ( $r=-.221$ ,  $z=-1.518$ ,  $p=.13$ ), es besteht kein Zusammenhang zwischen Ausgangslage und Verbesserungsrate. Die aufgeklärte Varianz ( $R^2$ ) der BSI-Werte durch die beiden latenten Wachstumsfaktoren liegt zwischen .697 und .712. In einem weiteren Schritt wurden Komorbiditätsmuster (keine Komorbidität, Komorbidität mit Persönlichkeitsstörung, Komorbidität zwischen Angst und affektiven Störungen ohne Persönlichkeitsstörung und andere Komorbiditäten) als Prädiktoren für intercept und slope in das Modell aufgenommen. Im Vergleich zu Patienten mit reiner Diagnose weisen Patienten mit komorbider Persönlichkeitsstörung erhöhte Eingangswerte im BSI auf (.254,  $z=2.480$ ,  $p<.05$ ); gleiches gilt für Patienten mit einer Kombination aus Angst und affektiven Störungen (.215,  $z=2.382$ ,  $p<.05$ ). Bezüglich der BSI-Ausgangswerte ergeben sich darüber hinaus keine bedeutsamen Unterschiede zwischen den Komorbiditätsgruppen. Betrachtet man den Verbesserungsverlauf über die Zeit,

so weisen Patienten aus der Gruppe „keine Komorbidität“ eine größere Steigung auf als solche aus der Gruppe „andere Komorbidität“ ( $-0.255$ ,  $z=2.129$ ,  $p<.05$ ). Weitere Unterschiede in den Verbesserungsraten in Abhängigkeit von der Komorbidität ergeben sich nicht (s. Tabelle 17).

Tabelle 17: Ergebnisse der Regression der Wachstumsfaktoren auf Komorbidität (BSI)

Referenz	keine Komorbidität	komorbid mit PS	komorbid Angst/Affekt.St.	komorbid, andere Kombinationen
		.254	.215	.106
keine Komorbidität	/	$z=2.480$ $p=.013$	$z=2.382$ $p=.017$	$z=1.466$ $p=.143$
komorbid mit PS	-.058 $z=-0.356$ $p=.722$	/	-.039 $z=-.320$ $p=.749$	-.148, $z=-1.334$ , $p=.182$
komorbid Angst/Affekt.St.	.015, $z=.094$ , $p=.925$	0.073 $z=.373$ $p=.709$	/	-.109 $z=-1.092$ $p=.275$
komorbid, andere Kombinationen	-.255, $z=-2.129$ , $p=.033$	-0.197 $z=-1.147$ $p=.251$	-.270 $z=-1.576$ $p=.115$	/

*Anmerkung:* über der Diagonalen: Prädiktion für intercept, unter der Diagonalen: Prädiktion für slope

### 5.2.1.2 GMM – Abhängige Variable BSI

In das einfache Wachstumsmodell mit drei Messzeitpunkten wurde eine weitere latente Variable eingefügt, die die Zugehörigkeit zu distinkten Klassen, sich ergebend aus Ähnlichkeiten in Ausgangslage und Steigung, erfassen sollte. Damit erfuhr das einfache Wachstumsmodell eine Erweiterung zum Growth Mixture Modell. Schrittweise wurden mehr latente Klassen eingefügt. Die Passungsgütekriterien zur Bestimmung der optimalen Anzahl an Klassen legen ein Modell mit vier Klassen als beste Lösung nahe. Das BIC nimmt bis zu vier Klassen schrittweise niedrigere Werte an (2091.548, 2051.060, 2046.478, 2038.423) und nimmt bei Einführung einer fünften Klasse wieder zu (2045.889). Nach dem Lo-Mendell-Rubin (LMR)-Test wird ebenfalls ein Modell mit vier Klassen präferiert; der Einbezug einer fünften Klasse führt nicht zu einer signifikanten Verbesserung des Modellfits (1 Klasse vs. 2 Klassen:  $p<.01$ ; 2 Klassen vs. 3 Klassen:  $p<.01$ ; 3 Klassen

vs. 4 Klassen:  $p < .01$ ; 4 Klassen vs. 5 Klassen:  $p = .34$ ). Die Entropie als Maß für die Güte der Klassifikation der Individuen in die identifizierten Klassen fällt für das vier-Klassen-Modell akzeptabel aus (.789).

Die Growth Mixture Analysen wurden wiederholt für einen Datensatz mit listenweisem Fallausschluss, so dass für alle eingehenden Personen alle drei Messzeitpunkte gegeben waren ( $N=321$ ). Auch hier resultiert ein loglineares Modell mit vier Klassen als beste Pasung. Hierfür sprechen sowohl das BIC (1 Klasse: 1467.294, 2 Klassen: 1431.428, 3 Klassen: 1412.313, 4 Klassen: 1399.480, 5 Klassen: 1405.136) als auch der LMR-Test (1 Klasse vs. 2 Klassen:  $p < .05$ ; 2 Klassen vs. 3 Klassen:  $p < .05$ ; 3 Klassen vs. 4 Klassen:  $p < .05$ ; 4 Klassen vs. 5 Klassen:  $p = .60$ ) und die Entropie (.884 bei 4 Klassen).

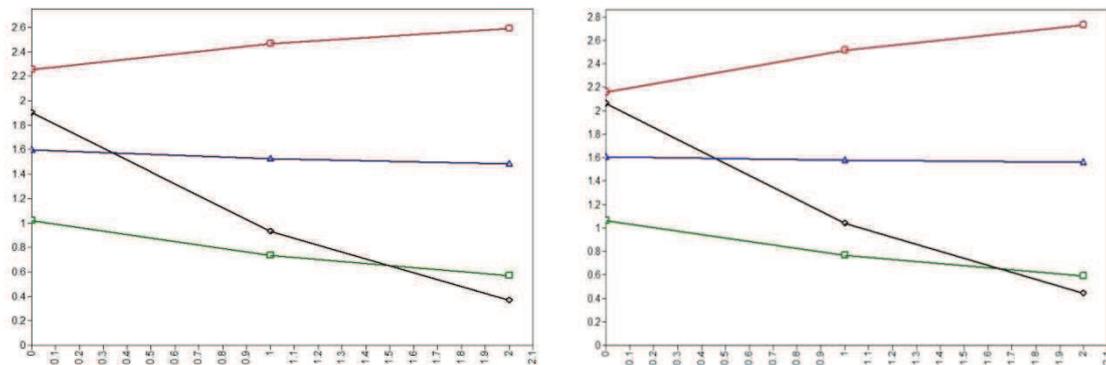


Abbildung 3: Latente Verlaufsklassen, abhängige Variable BSI, links resultierend aus imputierten Daten ( $N=462$ ); rechts resultierend aus listenweisem Fallausschluss ( $N=321$ ).

Die geschätzten mittleren Verlaufskurven, die sich aus den beiden GMM-Analysen ergeben, sind in Abbildung 3 dargestellt. Trotz der differierenden Stichproben (Imputation fehlender Werte vs. listenweiser Fallausschluss) resultieren ähnliche Verläufe. Es ergibt sich jeweils eine sehr große Klasse von Patienten, die die Therapie mit im Mittel relativ gering ausgeprägter Belastung im BSI beginnen und sich im Laufe der Behandlung leicht verbessern (Klasse 1, „leicht verbessert“; imputierte Stichprobe: 61.4%, listenweiser Fallabschluss 63.9%). Eine weitere Klasse umfasst jeweils rund ein Viertel der Patienten, die zu Beginn eine mittlere Belastung im BSI aufweisen und im Verlauf im Mittel keine Veränderung zeigen (Klasse 2 „unverändert“; imputierte Stichprobe: 27.9%, listenweiser Fallabschluss 25.5%). Jeweils zwei Klassen beginnen die Therapie mit sehr hoher subjektiver

Belastung. Ein Teil dieser Patienten verschlechtert sich im Mittel im Verlauf der Behandlung leicht (Klasse 3 „verschlechtert“; imputierte Stichprobe: 4.7%, listenweiser Fallausschluss 4.4%). Die vierte Klasse von Patienten startet ebenfalls mit im Mittel sehr hohen BSI-Werten, weist aber zu den beiden folgenden Messzeitpunkten sehr deutliche Verbesserungen auf (Klasse 4 „schnell verbessert“; imputierte Stichprobe: 6.0%, listenweiser Fallausschluss 6.2%).

Tabelle 18: Anzahl (relativer Anteil) Patienten in den identifizierten Klassen, in Abhängigkeit des Komorbiditätsmusters (listenweiser Fallausschluss, N=321).

Komorbiditätsmuster	wahrscheinlichste Klasse				Gesamt
	Klasse 1 leicht verbessert	Klasse 2 unverändert	Klasse 3 verschlechtert	Klasse 4 schnell verbessert	
keine Komorbidität	123 (75.9%)	25 (15.4%)	5 (3.1%)	9 (5.6%)	162 (100%)
Komorbidität mit Persönlichkeitsstörung	20 (55.6%)	13 (36.1%)	2 (5.6%)	1 (2.8%)	36 (100%)
Komorbidität Angst/Affekt. St.	22 (53.4%)	13 (30.9%)	1 (2.4%)	6 (14.3%)	42 (100%)
Komorbidität andere Muster	40 (49.4%)	31 (38.3%)	6 (7.4%)	4 (4.9%)	81 (100%)
Gesamt	205 (63.9%)	82 (25.5%)	14 (4.4%)	20 (6.2%)	321

*Anmerkung:* die Prozentangaben sind zeilenweise zu lesen.

Tabelle 18 zeigt die Anzahl und den relativen Anteil der Patienten in den identifizierten Klassen, in Abhängigkeit der Komorbiditätsmuster. Die Berechnung eines  $\chi^2$ -Tests wäre aufgrund der ungleichen Zellbesetzungen nicht aussagekräftig. Dennoch ist aus Tabelle 18 ablesbar, dass tendenziell ein größerer Anteil von Patienten mit nur einer Diagnose in die Klasse 1 „leicht verbessert“ fällt. In Klasse 2 „unverändert“ fallen verhältnismäßig häufig Personen mit einer komorbiden Persönlichkeitsstörung oder einem anderen Komorbiditätsmuster. Ähnliches gilt für Klasse 3 „verschlechtert“. In Klasse 4 „schnell verbessert“ sind Patienten aus allen Komorbiditätsgruppen vertreten, wobei ein vergleichsweise hoher Anteil auf die Kombination „Angst/affektive Störungen ohne Persönlichkeitsstörung“ fällt. Die Klassen unterscheiden sich nicht bedeutsam hinsichtlich des Anteils abgeschlossener

Therapien im Erhebungszeitraum ( $\chi^2_{(df3)}=1.975$ ,  $p=.578$ ;  $CI=.078$ ), jedoch in der Anzahl Therapiesitzungen, sofern die Therapie beendet wurde ( $F_{(3,238)}=4.477$ ,  $p<.05$ ,  $\eta^2=.053$ ), mit den kürzesten Therapien für Klasse 4 (schnell verbesserte Patienten;  $M=37.53$ ,  $SD=14.30$ ) und den längsten Therapien für Klasse 2 (unverändert;  $M=50.02$ ,  $SD=15.27$ ). Während in der Gruppe der schnell verbesserten Patienten alle Personen eine reliable Verbesserung im BSI sowohl zum Therapieende als auch in der Katamnese aufweisen, schwanken die Anteile für die übrigen Gruppen zwischen 70% und 87% (s. Anhang, S. 189).

### **5.2.2 Regressionsanalysen erster Block: zur Vorhersage von Therapieergebnis und –dauer unter Berücksichtigung von Komorbidität und Klassenzugehörigkeit**

Vor Berechnung der hierarchischen Regressionsanalysen wurden zunächst die Voraussetzungen geprüft. Nach Tabachnick und Fidell (2007) kann Multikollinearität an sehr hoch korrelierenden Variablen ( $>.70$ ) erkannt werden. Dies war für keine Kombination der Prädiktorvariablen der Fall (s. Anhang, S. 181). Darüber hinaus kann der sog. Toleranzwert für die Kollinearitätsdiagnose genutzt werden. Im Falle der Unabhängigkeit ist er 1 und steigt mit wachsender linearer Abhängigkeit. Urban & Mayerl (2006, S. 232) empfehlen, dass der Toleranzwert nicht unter .25 und der VIF (variance inflation factor, beruhend auf der Toleranz) nicht über 5.0 liegen sollte. Dies ist für die getesteten Variablen der Fall. Für die Anzahl an Therapiesitzungen als abhängige Variable war eine Verletzung der Normalverteilungsannahme gegeben. Verschiedene Datentransformationen wurden versucht, wobei durch die Anwendung der Quadratwurzel die Ursprungsdaten am ähnlichsten einer Normalverteilung wurden. Somit ging die Quadratwurzel der Sitzungsanzahl als abhängige Variable in die schrittweisen Regressionen ein.

#### **5.2.2.1 Differenz Eingangs- minus Endwert im BSI als Kriterium**

Für  $N=206$  Personen liegen BSI-Endwerte vor, so dass der Verbesserungswert als Differenz zwischen Ein- und Ausgangswert berechnet werden konnte. Schrittweise in die Regressionsgleichung aufgenommen wurden BSI-Eingangswert, Alter und Geschlecht, Anzahl Sitzungen, wahrscheinlichste Klassenzugehörigkeit (als dummy-codierte Variablen, wobei Klasse 4 „schnell verbessert“ als Referenzkategorie galt) und Komorbiditätsmuster

(als dummy-codierte Variablen, wobei „keine Komorbidität“ als Referenzkategorie galt). Die standardisierten Regressionskoeffizienten ( $\beta$ ), erklärter Varianzanteil ( $R^2$ ) und Änderung in  $R^2$  durch den Einschluss von Prädiktoren sind Tabelle 19 zu entnehmen. Nach jedem Schritt ergibt sich ein  $R$  signifikant unterschiedlich zu 0 (nach dem letzten Schritt  $R=.748$ ,  $F_{(10,195)}=24.798$ ,  $p<.01$ ). Wie zu erwarten hat der Eingangswert im BSI hohen Erklärungswert für die Veränderung zum Therapieende ( $\beta=.701$ ,  $p<.01$ ;  $R^2=.322$ ,  $p<.01$ ). Alter und Geschlecht, Anzahl Sitzungen sowie die Information über Komorbidität haben geringen Erklärungswert (keine bedeutsam von Null verschiedene Zunahme in  $R^2$  nach Aufnahme dieser Variablen ins Regressionsmodell). Dagegen leistet die über GMM identifizierte Klassenzugehörigkeit bedeutsame Varianzaufklärung über die BSI-Eingangsbelastung hinaus (Zunahme in  $R^2=.233$ ,  $p<.01$ ).

### 5.2.2.2 Differenz Eingangs- minus Katamnesewert im BSI als Kriterium

Für  $N=138$  Personen liegen BSI-Endwerte vor. Es wurden die gleichen Prädiktoren wie zur Vorhersage der BSI-Differenz zum Therapieende genutzt.

Die Ergebnisse für die Vorhersage der Verbesserung im BSI zwischen Therapiebeginn und Katamnese sind ebenfalls in Tabelle 19 dargestellt. Nach jedem Schritt ergibt sich ein  $R$  signifikant unterschiedlich zu 0 (nach dem letzten Schritt  $R=.727$ ,  $F_{(10,127)}=14.252$ ,  $p<.01$ ). Auch hier hat der BSI-Eingangswert hohen Erklärungswert ( $\beta=.550$ ,  $p<.01$ ;  $R^2=.305$ ,  $p<.01$ ). Alter und Geschlecht, Anzahl Sitzungen sowie die Information über Komorbidität erreichen als Prädiktoren keine Signifikanz. Die über GMM identifizierte Klassenzugehörigkeit hingegen leistet wiederum bedeutsame Signifikanzaufklärung (Zunahme in  $R^2=.191$ ,  $p<.01$ ).

### 5.2.2.3 Regression zur Prädiktion der Therapiedauer

Für  $N=242$  beendete Therapie liegen Angaben zur Anzahl Sitzungen vor. Schrittweise wurden folgende Prädiktoren aufgenommen: BSI-Eingangswert, Alter und Geschlecht, wahrscheinlichste Klassenzugehörigkeit und Komorbiditätsmuster. Nach dem letzten Schritt ergibt sich  $R=.408$ ,  $F_{(9,232)}=5.144$ ,  $p<.01$ . Sowohl die anfängliche Symptombelastung im BSI ( $\beta=.115$ ,  $p>.05$ ;  $R^2=.013$ ,  $p>.05$ ) als auch die soziodemografischen Variablen

(Alter:  $\beta=-.003$ ,  $p>.05$ , Geschlecht:  $\beta=-.039$ ,  $p>.05$ ;  $R^2=.002$ ,  $p>.05$ ) haben nur geringe und nicht signifikante Erklärungskraft für die Anzahl an Therapiesitzungen. Demgegenüber führt die Aufnahme der Blöcke Verlaufsklasse und Komorbidität jeweils zu einem bedeutsamen Anstieg erklärter Varianz (Verlaufsklasse:  $R^2=.057$ ,  $p<.01$ ; C2 vs. C4  $\beta=.371$ ,  $p<.01$ , C3 vs. C4  $\beta=.309$ ,  $p<.01$ ; Komorbidität:  $R^2=.094$ ,  $p<.01$ ; Komorbidität mit Persönlichkeitsstörung vs. keine Komorbidität  $\beta=.320$ ,  $p<.01$ ).

Tabelle 19: Regressionsanalysen zur Vorhersage von Therapieergebnis und –dauer unter Berücksichtigung von Komorbidität und Klassenzugehörigkeit

Prädiktor		Outcome					
		BSI-Differenz Ende (n=206)		BSI-Differenz Katamnese (n=138)		Anzahl Sitzungen (n=242)	
		$\beta$	Zunahme in $R^2$	$\beta$	Zunahme in $R^2$	$\beta$	Zunahme in $R^2$
Anfängl. Symptombelastung	BSI-GSI	.701 **	.322**	.550 **	.305**	.115	.013
soziodem. Variablen	Alter	.027		.072		-.003	
	Geschlecht	-.014	.000	-.074	.005	-.039	.002
Dauer	Anzahl Sitzungen	.024	.002	-.061	.012	-	-
Verlaufsklasse	C1	-.407 **	.233**	-.484 **	.191**	.080	.057**
	C2	-.617 **		-.553 **		.371 **	
	C3	-.261 *		-.466 **		.309 *	
Komorbidität	mit PS	.003	.002	.111	.015	.320 **	.094**
	Angst/Depr	-.049		-.041		.117	
	andere	-.005		-.017		.097	
Aufgeklärter Varianzanteil	$R^2$		.560		.529		.166

Anmerkung: \* $p<.05$ ; \*\* $p<.01$

### 5.2.3 Verlaufsanalysen – IIP als abhängige Variable

In den folgenden Abschnitten sind die Ergebnisse für das Ergebnismaß IIP dargestellt, wobei das Vorgehen analog zum Unterkapitel 5.2.1 (BSI) erfolgte. Die Verteilungskennwerte des IIP zu den vier eingehenden Messzeitpunkten finden sich im Anhang (S. 180).

#### 5.2.3.1 LGM - Abhängige Variable IIP

Wiederum spiegelt ein loglinearer Verbesserungsverlauf die beste Passung an die Daten wider, gestützt durch RMSEA und BIC (s. Tabelle 20). Das Modell mit freien Parametern schneidet im Vergleich zu den Gegenmodellen mit Fixierungen der Wachstumsfaktoren besser ab. In diesem Modell sind beide Wachstumsfaktoren signifikant von Null verschieden ( $M_{\text{intercept}}=1.593$ ,  $z=61.295$ ,  $p<.001$ ;  $M_{\text{slope}}=-.514$ ,  $z=-12.999$ ,  $p<.001$ ) und es besteht bedeutsame Variabilität in Ausgangslage und Steigung ( $\text{Var}_{\text{intercept}}=.229$ ,  $z=10.627$ ,  $p<.001$ ;  $\text{Var}_{\text{slope}}=.152$ ,  $z=2.029$ ,  $p=.042$ ). Es besteht eine signifikante Unterschiedlichkeit der betrachteten Patienten hinsichtlich ihrer Ausgangswerte und Veränderung über die Zeit. Die Korrelation zwischen intercept und slope ist auch im IIP nicht signifikant ( $r=.172$ ,  $z=.727$ ,  $p=.467$ ), es besteht somit kein Zusammenhang zwischen Ausgangslage und Verbesserungsrate. Die aufgeklärte Varianz ( $R^2$ ) der BSI-Werte durch die beiden latenten Wachstumsfaktoren liegt zwischen .713 und .818.

Tabelle 20: Ergebnisse der LGM-Analysen für die abhängige Variable IIP

		loglikelihood (H0)	RMSEA	CFI	BIC	
1	Testung der Verlaufsform					
		intercept-only	-1.126.440	0.195	0.774	2.289.706
		linear	-1.007.770	0.091	0.969	2.070.779
		loglinear	-1.007.694	0.086	0.973	2.070.627
2	loglinearer Verlauf - Testung von Gegenmodellen mit Fixierungen	quadratisch	-988.111	0.091	0.994	2.056.012
		fixed i, random s	-1.305.911	0.333	0.424	2.654.786
		random i, fixed s	-1.021.330	0.096	0.952	2.085.624
		fixed i, fixed s	-1.666.211	0.436	0.000	3.369.248
3	Einbezug von Komorbiditätsmustern als Prädiktor	-988.996	0.079	0.965	2.070.058	

Schließlich wurden Komorbiditätsmuster (keine Komorbidität, Komorbidität mit Persönlichkeitsstörung, Komorbidität zwischen Angst und depressiven Störungen ohne Persön-

lichkeitsstörung und andere Komorbiditäten) als Prädiktoren für intercept und slope in das Modell aufgenommen (s. Tabelle 20). Im Vergleich zu Patienten mit reiner Diagnose weisen Patienten mit komorbider Persönlichkeitsstörung erhöhte Eingangswerte im IIP auf (.424,  $z=5.043$ ,  $p<.01$ ). Gleiches gilt für Patienten mit einer Kombination aus Angst und affektiven Störungen (.196,  $z=2.624$ ,  $p<.01$ ), welche wiederum verglichen mit Patienten mit komorbiden Persönlichkeitsstörungen signifikant geringere IIP-Eingangsbelastung zeigen (-.228,  $z=-2.256$ ,  $p<.05$ ). Auch komorbide Patienten mit anderen Kombinationen weisen bedeutsam geringere Eingangswerte auf als solche mit komorbider Persönlichkeitsstörung (-.364,  $z=-4.014$ ,  $p<.01$ ). Betrachtet man den Verbesserungsverlauf über die Zeit, so weisen Patienten aus der Gruppe „keine Komorbidität“ eine größere Steigung auf als solche aus der Gruppe „andere Komorbidität“ (-.255,  $z=2.129$ ,  $p=.033$ ). Weitere Unterschiede in den Verbesserungsraten in Abhängigkeit von der Komorbidität ergeben sich nicht (s. Tabelle 21).

Tabelle 21: Ergebnisse der Regression der Wachstumsfaktoren auf Komorbidität (IIP)

Referenz	keine Komorbidität	komorbid mit PS	komorbid Angst/Affekt. St.	komorbid, andere Kombinationen
keine Komorbidität	/	.424 $z=5.043$ $p<.01$	.196 $z=2.624$ $p=.009$	.060 $z=.993$ $p=.321$
komorbid mit PS	.036 $z=.278$ $p=.781$	/	-.228 $z=-2.256$ $p=.024$	-.364 $z=-4.014$ $p<.01$
komorbid Angst/Affekt. St.	.044 $z=-.361$ $p=.718$	-.080 $z=-.505$ $p=.614$	/	-.136 $z=-1.658$ $p=.097$
komorbid, andere Kombinationen	-.216 $z=-2.100$ $p=.036$	-.252 $z=-1.781$ $p=.075$	-.172 $z=-1.355$ $p=.175$	/

*Anmerkung:* über der Diagonalen: Prädiktion für intercept, unter der Diagonalen: Prädiktion für slope

### 5.2.3.2 GMM – Abhängige Variable IIP

In das einfache Wachstumsmodell mit drei Messzeitpunkten wurde wiederum eine latente Variable eingefügt, die die Zugehörigkeit zu distinkten Klassen erfassen sollte (GMM).

Analog zum oben beschriebenen Vorgehen wurden schrittweise mehr latente Klassen eingefügt. Bei Einführung einer zweiten Klasse nimmt das BIC leicht zu (1577.363 bei einer Klasse vs. 1578.793 bei zwei Klassen) und der Lo-Mendell-Rubin (LMR)-Test wird gerade nicht mehr signifikant (1 Klasse vs. 2 Klassen:  $p=.051$ ). Gegen eine zwei-Klassen-Lösung spricht auch das Größenverhältnis der identifizierten Klassen: während eine Klasse fast 99% aller Personen enthält, erreicht die zweite lediglich eine absolute Größe von 5. Dennoch wurde ein weiteres Modell mit drei Klassen getestet. Dabei kam es zu einem weiteren Anstieg im BIC (1580.400), der LMR-Test ist nicht signifikant (2 Klasse vs. 3 Klassen:  $p=.23$ ). Auch hier ergäbe sich eine sehr kleine Klasse (3% der eingehenden Personen), deren klinische Bedeutsamkeit fraglich ist. Die Ergebnisse, insbesondere der ansteigende BIC, welcher die Modellsparsamkeit berücksichtigt, weisen darauf hin, dass mit dem Ein-Klassen-Modell die beobachtete Verteilung ausreichend gut wiedergegeben ist. Regressionsanalysen zur Vorhersage von Therapieergebnis und –dauer unter Einbezug der Klassenzugehörigkeit im IIP erfolgten daher nicht.

#### **5.2.4 Regressionsanalysen zweiter Block: zur Vorhersage von Therapieergebnis und –Dauer unter Berücksichtigung von Komorbidität**

In Vorab-Datenanalysen zeigte sich, dass Angaben über die Anzahl therapeutischer Sitzungen bei abgeschlossenen Therapien in der Kontrollgruppe sehr lückenhaft waren (IG 350 von 360 abgeschlossenen Therapien, entspricht 97%; KG nur 40 von 183 abgeschlossenen Therapien, entspricht 22%). Die nachfolgend dargestellten Ergebnisse der Regressionsanalysen beziehen sich daher lediglich auf Patienten der IG (Ergebnisse s. Tabelle 22).

##### **5.2.4.1 Vorhersage des Therapieergebnisses im BSI zum Abschluss- und Katamnesezeitpunkt**

Zum Zeitpunkt des Abschlusses des Therapie gehen  $N=302$  Personen in die Analysen ein, zur Katamnese  $N=197$ . Bei Testung der Voraussetzungen für lineare Regressionsanalysen zeigte sich, dass die Daten zu beiden Zeitpunkten deutlich rechtssteil ausfielen, weshalb sie einer  $\ln$ -Transformation unterzogen wurden (Histogramme, Streudiagramme sowie Boxplots finden sich im Anhang, S. 183ff.). Nach dem letzten Schritt ergibt sich für die Vorhersage des BSI-Endwertes  $R=.550$ ,  $F_{(7,204)}=18.247$ ,  $p<.01$  und für die Vorhersage des

BSI-Katamnesewertes  $R=.464$ ,  $F_{(7,189)}=7.406$ ,  $p<.01$ . Sowohl für die Vorhersage des BSI-Wertes zum Ende der Therapie als auch zur Katamnese hat die anfängliche Belastung im BSI großen Einfluss (Abschluss:  $R^2=.277$ ,  $p<.01$ ,  $\beta=.489$ ,  $p<.01$ ; Katamnese:  $R^2=.184$ ,  $p<.01$ ,  $\beta=.410$ ,  $p<.01$ ). Darüber hinaus ergibt sich in beiden Fällen ein bedeutsamer Anstieg erklärter Varianz durch die Hinzunahme der Anzahl Sitzungen (Abschluss: Zunahme in  $R^2=.012$ ,  $p<.05$ ,  $\beta=.084$ ,  $p>.05$ ; Katamnese: Zunahme in  $R^2=.020$ ,  $p<.05$ ,  $\beta=.132$ ,  $p>.05$ ). Der Prädiktor Komorbidität bleibt hingegen ohne bedeutsamen Einfluss.

#### **5.2.4.2 Vorhersage des Therapieergebnisses im IIP zum Abschluss- und Katamnesezeitpunkt**

Hier gehen zum Therapieende  $N=304$  Personen in die Analysen ein, zur Katamnese  $N=197$ . Bezüglich der Verteilung der IIP-Werte zu den untersuchten Zeitpunkten war keine Verletzung der Normalverteilungsannahme zu erkennen, weshalb in diesem Fall die Rohwerte als abhängige Variablen genutzt wurden (Histogramme, Streudiagramme sowie Boxplots siehe Anhang, S. 183ff.). Nach dem letzten Schritt ergibt sich für die Vorhersage des IIP-Endwertes  $R=.581$ ,  $F_{(7,296)}=21.584$ ,  $p<.01$  und für die Vorhersage des IIP-Katamnesewertes  $R=.568$ ,  $F_{(7,189)}=12.842$ ,  $p<.01$ . Die IIP-Eingangsbelastung resultiert als signifikanter Prädiktor sowohl zur Vorhersage des End- als auch des Katamnesewertes (Abschluss:  $R^2=.322$ ,  $p<.01$ ,  $\beta=.544$ ,  $p<.01$ ; Katamnese:  $R^2=.284$ ,  $p<.01$ ,  $\beta=.487$ ,  $p<.01$ ). Die Hinzunahme der Anzahl Sitzungen erhöht den Anteil aufgeklärter Varianz nicht bedeutsam. Jedoch führt das Vorhandensein von Komorbidität zwischen Angst und affektiven Störungen im Vergleich zu keiner Komorbidität zu signifikant erhöhten IIP-Endwerten.

#### **5.2.4.3 Vorhersage des Therapieergebnisses im BDI zum Abschluss- und Katamnesezeitpunkt**

Für das störungsspezifische Instrument BDI konnten zum Therapieende  $N=209$  Personen in die Analysen einbezogen werden, zur Katamnese  $N=144$ . Auch hier lag eine Verletzung der Normalverteilungsannahme vor, so dass statt des Rohwertes die ln-transformierten BDI-Werte als abhängige Variablen genutzt werden (Histogramme, Streudiagramme sowie Boxplots siehe Anhang, S. 183ff.). Nach dem letzten Regressionsschritt ergibt sich für die

Vorhersage des BDI-Endwertes  $R=.468$ ,  $F_{(7,201)}=8.059$ ,  $p<.01$  und für die Vorhersage des BDI-Katamnesewertes  $R=.457$ ,  $F_{(7,136)}=5.120$ ,  $p<.01$ . Wiederum zeigt sich der Eingangswert als bedeutsamer Prädiktor sowohl für den End- als auch den Katamnesezeitpunkt (Abschluss:  $R^2=.179$ ,  $p<.01$ ,  $\beta=.384$ ,  $p<.01$ ; Katamnese:  $R^2=.129$ ,  $p<.01$ ,  $\beta=.311$ ,  $p<.01$ ). Bezüglich der Vorhersage des BDI-Endwertes erreicht die Anzahl durchgeführter Therapiesitzungen keinen bedeutsamen zusätzlichen Einfluss (Zunahme in  $R^2=.012$ ,  $p>.05$ ), dagegen führt jedoch Komorbidität zwischen Angst- und affektiven Störungen im Vergleich zu Einfachdiagnosen zu erhöhten BDI-Endwerten ( $\beta=.136$ ,  $p<.05$ ; für den gesamten Block Komorbidität Zunahme in  $R^2=.025$ ,  $p>.05$ ). Bezüglich der Prädiktion des BDI-Katamnesewertes leistet die Anzahl Sitzungen einen bedeutsamen Beitrag zur Varianzaufklärung (Zunahme in  $R^2=.060$ ,  $p<.01$ ;  $\beta=.245$ ,  $p<.01$ ). Die Hinzunahme der Information über Komorbidität bleibt indes ohne bedeutsamen Einfluss.

#### 5.2.4.4 Vorhersage der Anzahl Therapiesitzungen

Zur Vorhersage der Anzahl durchgeführter Therapiesitzungen bei beendeten Therapien konnten Daten von  $N=347$  Personen genutzt werden. Auch diese Daten entsprachen nicht der Annahme einer Normalverteilung, weshalb in diesem Fall die Bildung der Quadratwurzel gewählt wurde, um die Daten normalverteilungsähnlicher zu machen (Histogramme, Streudiagramme sowie Boxplots siehe Anhang, S. 183ff.). Nach dem letzten Schritt ergibt sich für die Vorhersage der Anzahl Sitzungen  $R=.304$ ,  $F_{(6,340)}=5.778$ ,  $p<.01$ . Auch hier zeigt sich die Eingangsbelastung (in diesem Fall wurde der BSI als allgemeines Instrument gewählt) als bedeutsam für die Vorhersage ( $R^2=.031$ ,  $p<.01$ ,  $\beta=.128$ ,  $p<.05$ ). Darüber hinaus können durch Hinzunahme des Blocks Komorbidität weitere Anteile an Varianz aufgeklärt werden (Zunahme in  $R^2=.055$ ,  $p<.01$ ; für Komorbidität mit Persönlichkeitsstörung  $\beta=.221$ ,  $p<.01$ ; für Komorbidität aus Angst- und affektiven Störungen  $\beta=.148$ ,  $p<.05$ ). Bis zum Therapieende benötigen Patienten mit nur einer Diagnose im Mittel 35 ( $M=35.87$ ,  $SD=16.52$ ), Patienten mit mehr als einer Störung 44 ( $M=44.90$ ,  $SD=21.07$ ) Sitzungen ( $t_{(348)}=-4.480$ ;  $p<.01$ ;  $1-\beta=.99$ ). Die längsten Therapien haben Patienten mit komorbider Persönlichkeitsstörung ( $M=51.21$ ,  $SD=20.79$ ), gefolgt von solchen mit einer Kombination aus Angst- und affektiven Störungen ( $M=46.67$ ,  $SD=20.14$ ) und anderen Diagnosekombinationen ( $M=40.96$ ,  $SD=21.09$ ) (s. Anhang, S. 191).

### 5.2.4.5 Weiterführende Analysen

Zusätzlich in den in Hypothese 9 und Hypothese 10 formulierten Fragestellungen wurden weitere Analysen durchgeführt, um den Einfluss komorbider Diagnosen auf verschiedene Ergebnismaße zu untersuchen. Als abhängige Variablen wurden wiederum BSI, IIP und BDI herangezogen, allerdings die Werte zu den Verlaufsmessungen 2 (um die 10. Sitzung) und 3 (um die 20. Sitzung). Um die Variablen normalverteilungsähnlicher zu machen, wurde für BSI und BDI eine Transformation (Bildung der Quadratwurzel) durchgeführt (Verteilungskennwerte s. Anhang, S. 187). Als unabhängige Variablen wurden schrittweise aufgenommen: Anfängliche Symptombelastung im jeweiligen Ergebnismaß, Alter und Geschlecht und Komorbidität. Die Ergebnisse sind im Anhang (S. 190) aufgeführt. Während die Eingangsbelastung durchgängig einen hoch bedeutsamen Prädiktor für den Status zu den beiden Zwischenmessungen darstellt, leisten die übrigen Variablen keine zusätzliche Varianzaufklärung. Lediglich zur Vorhersage des BSI-Wertes zur dritten Zwischenmessung erweist sich Komorbidität als einflussreiche Variable zusätzlich zur BSI-Eingangsbelastung.

Darüber hinaus wurde der Status der Patienten zum Therapieende und zum Katamnesezeitpunkt in Abhängigkeit der Komorbiditätsgruppe untersucht. Unter Betrachtung der relativen Verbesserung (Differenz zwischen Eingangs- und Endwert) ergeben sich keine signifikanten Unterschiede zwischen den Komorbiditätsgruppen in den drei zentralen Ergebnismaßen BSI ( $F_{(3,460)}=1.398, p=.243, 1-\beta=.99$ ), IIP ( $F_{(3,461)}=.291, p=.832, 1-\beta=.99$ ) und BDI ( $F_{(3,326)}=.503, p=.681, 1-\beta=.98$ ). Betrachtet man jedoch das absolute Therapieergebnis, so zeigen sich in allen Maßen bedeutsame Unterschiede zwischen den Komorbiditätsgruppen (BSI:  $F_{(3,464)}=6.596, p<.01, \eta^2=.041$ ; IIP:  $F_{(3,463)}=5.716, p<.01, \eta^2=.036$ ; BDI:  $F_{(3,329)}=6.276, p<.01, \eta^2=.054$ ), mit signifikant besseren Endwerten für Patienten mit nur einer Diagnose im Vergleich zu solchen mit komorbider Persönlichkeitsstörung oder einer Kombination aus Angst- und affektiven Störungen (s. a. Anhang, S. 191).

Daneben wurde für jeden Patienten bestimmt, ob er in den zentralen Ergebnismaßen BSI, IIP und BDI reliable Veränderung aufweist oder beide Kriterien einer klinisch signifikanten Verbesserung (nach Jacobson & Truax, 1991) erfüllte. Die Ergebnisse finden sich im Anhang (S. 192f.). Aufgrund der ungleichen Zellbesetzungen wurden keine Tests auf Unabhängigkeit durchgeführt, jedoch lassen sich aus den prozentualen Angaben einige Tendenzen ablesen. Zunächst einmal ist festzuhalten, dass die relativen Anteile in den ver-

schiedenen Veränderungsgruppen für alle Ergebnismaße zwischen Therapieende und Katamnese relativ stabil bleiben. Die Patientengruppe mit nur einer Diagnose weist bezüglich der Ergebnismaße BSI und IIP den höchsten Anteil klinisch signifikanter Verbesserung auf (zwischen 51% und 57%), während diese Anteile im BDI deutlich geringer ausfallen (um die 30%). Der Anteil nicht reliabel verbesserter Patienten ist unter Heranziehung des Ergebnismaßes IIP deutlich höher (ca. 44%) als unter Betrachtung des BSI (14%-18%) und des BDI (um die 30%). Dementsprechend ist auch im BSI der Anteil reliabel verbesserter Patienten am höchsten (78%), gefolgt von BDI (68%) und IIP (50%).

Die Gruppen reliabel verbesserter Patienten und solcher ohne reliable Verbesserung in den drei Ergebnismaßen zum Therapieende und in der Katamnese wurden außerdem bezüglich ihrer mittleren Therapiedauer untersucht. Patienten mit reliabler Verbesserung im BSI zum Therapieende oder zum Katamnesezeitpunkt unterscheiden sich hinsichtlich ihrer Therapielänge nicht bedeutsam von denjenigen ohne reliable Verbesserung (Therapieende  $t_{(327)}=-.351$ ,  $p=.725$ ;  $1-\beta=.96$ ; Katamnese  $t_{(221)}=-.337$ ,  $p=.737$ ;  $1-\beta=.87$ ). Jedoch verbleiben diejenigen Patienten, die zum Ende der Therapie im IIP im Vergleich zum Aufnahmewert reliabel verbessert sind bedeutsam länger in der Therapie als die Vergleichsgruppe (reliable Verbesserung  $N=181$ ,  $M=40.96$ ,  $SD=20.37$ , keine reliable Verbesserung  $N=150$ ,  $M=36.29$ ,  $SD=19.18$ ;  $t_{(329)}=-2.128$ ,  $p<.05$ ;  $d=.24$ ); dieser Unterschied findet sich jedoch nicht zum Katamnesezeitpunkt ( $t_{(221)} = -1.523$ ,  $p=.129$ ;  $1-\beta=.96$ ). Unter Betrachtung des BDI als Ergebnismaß lassen sich ebenfalls Mittelwertsunterschiede finden; hier weisen reliabel verbesserte Patienten eine größere Anzahl Sitzungen auf, und zwar unabhängig ob zur Bestimmung des Status der Abschluss- oder Katamnesezeitpunkt herangezogen wird (Ende: reliable Verbesserung  $N=223$ ,  $M=42.26$ ,  $SD=20.52$ , keine reliable Verbesserung  $N=109$ ,  $M=31.81$ ,  $SD=16.62$ ,  $t_{(330)}=-4.624$ ,  $p<.01$ ;  $d=.56$ ; Katamnese: reliable Verbesserung  $N=152$ ,  $M=39.72$ ,  $SD=20.72$ , keine reliable Verbesserung  $N=73$ ,  $M=32.93$ ,  $SD=17.87$ ,  $t_{(223)}=-2.403$ ,  $p<.05$ ;  $d=.35$ ).



## 6 DISKUSSION

Das Ziel dieser Studie ist die Untersuchung des gemeinsamen Auftretens psychischer Störungen sowie des Einflusses komorbider Diagnosen auf Eingangsstatus, Therapieverlauf und –ergebnis ambulanter Psychotherapiepatienten. Im Folgenden werden die Ergebnisse im Hinblick auf die zu Beginn formulierten Fragestellungen diskutiert und in den Forschungskontext gestellt. Nach einer allgemeinen Reflexion der Ergebnisse werden kritische Aspekte dieser Arbeit beleuchtet und schließlich ein allgemeines Fazit gezogen sowie ein Ausblick auf mögliche zukünftige Forschungsfragen gegeben.

### 6.1 Reflexion der Ergebnisse

In den nachstehenden Abschnitten werden die Resultate mit Blick auf die verschiedenen Gruppen von Fragestellungen dieser Arbeit reflektiert und in Bezug zu früherer Forschung gesetzt.

#### 6.1.1 Zur Veränderung des Diagnoseverhaltens unter Anwendung von IDCL

Während den Therapeuten der KG keine Vorgaben hinsichtlich der Art und Weise der Diagnosestellung gemacht wurden, wurden die Therapeuten der IG in einer einführenden Schulung und den projektbezogenen Informationsmaterialien dazu aufgefordert, ein strukturiertes Vorgehen zu wählen, in welchem Diagnosechecklisten (Hiller et al., 2004) als Kernstück dienten.

Sofern mindestens eine Diagnose pro Patient vorliegt, werden in der IG signifikant mehr Diagnosen pro Patient gestellt als in der KG. Auch unter Berücksichtigung der Eingangsbelastung differenziert die Zugehörigkeit zu IG oder KG zwischen komorbiden und nicht komorbiden Patienten. Der Anteil komorbider Patienten in der IG liegt mit um die 50% in vergleichbarer Größenordnung wie nach Durchführung eines strukturierten Interviews (Wittchen, 1996; Zimmerman & Mattia, 1999b) und auch der Anteil komorbider Störungen in der KG (um die 33%) entspricht Angaben zur klinischen Routinediagnostik bei unstandardisiertem Vorgehen (Wittchen, 1996). In der IG werden hypothesenkonform häufi-

ger als in der KG Diagnosen aus dem Bereich der Angststörungen (mit Ausnahme der Panikstörung) und Persönlichkeitsstörungen gestellt, daneben werden hier im Vergleich zur KG häufiger rezidivierende depressive Störungen und Dysthymia diagnostiziert. Das gehäufte Vorkommen von Angststörungen und Persönlichkeitsstörungen unter strukturiertem Vorgehen ist konform mit in der Literatur berichteten Ergebnissen (Shear et al., 2000; Zimmerman & Mattia, 1999b). Hingegen findet sich in der KG signifikant häufiger die Diagnose Angst/Depression gemischt. Die Annahme eines gehäuften Vorkommens der Anpassungsstörungen in der KG wird nicht bestätigt, wenngleich ein entsprechender Trend, ebenso wie für die depressive Episode, aus den Daten ablesbar ist. Auf Unterschiede in der allgemeinen psychischen Belastung zu Therapiebeginn sind diese Ergebnisse nicht zurückzuführen.

Möglicherweise haben die Vorgaben zur Diagnostik bei den Therapeuten der IG zu einem präziseren diagnostischen Vorgehen geführt. Aus der Anwendung strukturierter klinischer Interviews ist bekannt, dass deren Daten als vollständiger und valider einzustufen sind als solche aus unstrukturierten Befragungen (z. B. Zimmerman & Mattia, 1999a). In ähnlicher Weise können die Checklisten dazu anregen, neben ad-hoc-Verdachtsdiagnosen Alternativen in das Blickfeld des Diagnostikers zu rücken und Information zu erfragen, die ansonsten möglicherweise übersehen würde. Gerade die unterschiedlichen Verteilungen affektiver Störungen können ein Hinweis darauf sein, dass die Therapeuten der IG gezielter nach Ausprägung und zeitlicher Ausdehnung der Symptomatik gefragt haben.

Die Ergebnisse dieser ersten Fragestellung sind allerdings mit Vorsicht zu interpretieren, da im Rahmen des TK-Modellvorhabens keine Kontrolle darüber bestand, inwieweit die Therapeuten sich im Einzelfall tatsächlich an die diagnostischen Vorgaben gehalten haben und wie der zeitliche Mehraufwand durch die Anwendung der IDCL eingeschätzt wird. Somit können auch Therapeuten der KG ein strukturiertes und Therapeuten der IG ein eher unstrukturiertes Vorgehen gewählt haben. Aus den demografischen und Belastungsmaßen ergeben sich keine Anhaltspunkte dafür, dass die Patienten der KG ein unterschiedliches Klientel als Patienten der IG darstellen. Jedoch bestehen einige konfundierende Faktoren, die die erhöhte Komorbiditätsrate in der IG zum Teil erklären könnten. So wurden die Therapeuten der IG im Gegensatz zu denjenigen der KG geschult und für die Studienteilnahme bezahlt, was zu einem erhöhten Commitment beigetragen haben mag. Eine offene Frage ist darüber hinaus, warum der Anteil an Patienten ohne diagnostische Information

insgesamt sehr hoch ist (ca. 15%) und in der IG noch einmal deutlich größer ausfällt als in der KG. Das Vorhandensein von mindestens einer ICD-10-Diagnose galt im Modellvorhaben als Voraussetzung für die Einleitung des modifizierten Gutachterverfahrens. Die Tatsache, dass für rund 93% der Patienten der IG zur Therapiebewilligung kein externer Gutachter eingeschaltet wurde (Wittmann et al., 2011, S. 133) spricht dafür, dass hier zwar Diagnosen gestellt, nicht aber im Datensatz verzeichnet wurden. Einschränkend ist weiterhin anzumerken, dass sich im Rahmen des TK-Modellvorhabens keine Unterschiede hinsichtlich des Therapieerfolgs zwischen IG und KG ergeben (Wittmann et al., 2011), dass es damit also keine Hinweise auf einen Zusammenhang zwischen der Art der Diagnostik und dem Ergebnis gibt. Inwieweit die Anwendung von Diagnosechecklisten Aspekte der Indikation und Therapieplanung beeinflusst, könnte in nachfolgenden Studien durch eine Befragung von Therapeuten geklärt werden. Daneben sollte, etwa in quasiexperimenteller Form, eine vergleichende Untersuchung des Diagnoseverhaltens unter den drei Bedingungen Diagnosechecklisten, strukturiertes klinisches Interview und unstrukturiertes Vorgehen durchgeführt werden, um in Zukunft Vor- und Nachteile sowie klinischen Nutzen dieser Methoden unter verschiedenen Gegebenheiten besser abwägen zu können.

### **6.1.2 Zum gemeinsamen Auftreten psychischer Störungen**

Es wurde untersucht, ob und wenn ja welche der vorab bestimmten 16 Störungsgruppen häufiger gemeinsam diagnostiziert werden als unter Annahme der Gleichverteilung zu erwarten wäre.

Ein bedeutsamer positiver Zusammenhang, der so auch in der Literatur zu finden ist (u. a. Diguier, Barber, & Luborsky, 1993; Hardy et al., 1995; Shea et al., 1990), ergibt sich lediglich für die Kombination von Persönlichkeitsstörungen mit rezidivierender depressiver Störung. Dies entspricht der Beobachtung aus zahlreichen Forschungsarbeiten, dass Persönlichkeitsstörungen als Risikofaktoren für die Entwicklung spezifischer psychischer Störungen, insbesondere affektiver Störungen, gelten können (Fiedler, 1998). Die Tatsache, dass sich hier ein Zusammenhang der Persönlichkeitsstörungen mit rezidivierender depressiver Störung, nicht aber mit der einzelnen depressiven Episode ergibt, ist angesichts der den Achse-II-Störungen inhärenten tief verwurzelten und weitgehend stabilen Mustern, die sich wiederholt in akuten Schwierigkeiten manifestieren können, nur schlüssig. Für die Kombination aus Depression und Generalisierter Angststörung findet sich nahezu ein

Nullzusammenhang. Dieses Resultat steht im Widerspruch zu zitierten Studien (z. B. Borkovec, Abel et al., 1995; Newman et al., 2010); jedoch wird an anderer Stelle auch bemerkt, dass die Zweitdiagnose Depression bei GAS gegenüber den Phobien vergleichsweise selten vergeben wird (Becker & Margraf, 2007). Die oft berichtete Komorbidität zwischen verschiedenen Angststörungen spiegelt sich in den vorliegenden Daten nicht wider. Dies kann zum einen daran liegen, dass für diese Studie, im Unterschied zu vielen zitierten Untersuchungen, das ICD-10 als Diagnosesystem zugrunde lag. Hier bestehen für einige Angststörungen Ausschlusshinweise und insbesondere für die GAS auch Ausschlussdiagnosen aus dem Bereich der Angststörungen. Somit ist etwa die gemeinsame Diagnose von GAS und Panikstörung nicht zulässig und sollte demgemäß in den Daten nicht zu finden sein. Ein weiterer Grund für die geringe Komorbidität zwischen verschiedenen Angststörungen liegt möglicherweise in der für diese Arbeit getroffenen Einteilung der Störungsgruppen. Bei der Größe des gegebenen Datensatzes und im Durchschnitt weniger als zwei Diagnosen pro Patient sind einige Zellen, die das gemeinsame Auftreten zweier Diagnosen anzeigen, notwendigerweise nur schwach oder gar nicht besetzt. So kann die Wahl von 16 Störungsgruppen zu spezifisch sein, um noch bedeutsame Unterschiede zu identifizieren. Allerdings wurde bei der Diagnosegruppenbildung bewusst darauf geachtet, Zusammenfassungen zu in sich sehr heterogenen und damit für Komorbiditätsaussagen wenig relevanten Gruppen zu vermeiden. Für diejenigen Diagnosen, welche nur selten oder gar nicht in der untersuchten Stichprobe vorhanden sind, kann keine Aussage über spezifische Komorbiditätsmuster getroffen werden, wobei auch nicht geschlossen werden kann, dass keine Verbindung zwischen zwei Störungen besteht. Es wäre sinnvoll, die hier gewählte Aufspaltung an Diagnosen in Zukunft an einem größeren Datensatz anzuwenden.

Neben den prognostizierten positiven Beziehungen finden sich auch einige signifikante negative Zusammenhänge zwischen den Störungsgruppen. In besonders großem Ausmaß gilt dies für die Anpassungsstörungen (negative Relation zu depressiver Episode, Dysthymia, sozialer Phobie und spezifischer Phobie). Diese Beobachtung deckt sich mit der Vorgabe, dass diese Diagnose in der Regel nur dann vergeben werden soll, wenn die gegebene Problematik nicht durch eine andere Störung besser erklärt werden kann (Bengel & Hubert, 2010). Die Tatsache, dass Dysthymia einen signifikanten negativen Zusammenhang zu anderen affektiven Störungen aufweist zeigt, dass die sogenannte „Double Depression“ in der ambulanten Praxis nur selten diagnostiziert wird. Ob dies daran liegt, dass nach Be-

stätigung einer affektiven Störung nicht nach weiteren geschaut wurde oder dass diese Störungen tatsächlich nur gelegentlich gemeinsam vorliegen (die lebenszeitliche Komorbidität liegt bei ca. 1.5%, Wittchen & Hoyer, 2011), kann an dieser Stelle nicht beantwortet werden. Ins Auge fallen außerdem die hohen negativen Korrelationen zwischen depressiver Störung und Panikstörung/Agoraphobie beziehungsweise spezifischer Phobie. In der Literatur wird zu beiden Angststörungen Komorbidität mit Depression berichtet (u. a. Kessler et al., 2005; Noyes et al., 1990), was auch als Folge eingeschränkten Funktionierens im Alltag klinisch einleuchtend wäre. Eine Erklärungsmöglichkeit für die hier gefundene Negativkorrelation ist, dass Patienten mit diesen Angststörungen zu einem frühen Zeitpunkt in die Therapie kommen, eben dann, wenn noch keine Komplikation im Sinne einer komorbiden depressiven Störung aufgetreten ist. Insgesamt ergeben sich für die depressive Episode deutlich mehr negative Zusammenhangstendenzen und weniger Komorbidität als für die rezidivierende depressive Störung. Personen mit letzterer Diagnose weisen naturgemäß eine höhere Chronizität auf, was Komplikationen wie weitere psychische Störungen mit sich bringen kann. In der Literatur zur Komorbidität affektiver Störungen wird nur selten explizit eine Unterscheidung getroffen zwischen einfacher und rezidivierender depressiver Episode. Die hier dargestellten Ergebnisse zeigen, dass diese Differenzierung im Hinblick auf Komorbidität und damit auch auf die Therapieplanung jedoch durchaus von Relevanz ist.

Aus dem Vergleich der Überschneidungen zwischen den Störungsgruppen und ihrer relativen Anteile auf der einen Seite und den tetrachorischen Korrelationen auf der anderen Seite wird der Unterschied zwischen den Konzepten „co-occurrence“ und Komorbidität (Krueger, 2002) deutlich. Beispielsweise tritt eine depressive Episode in rund 24% der Patienten mit sozialer Phobie auf; dieser Zusammenhang ist jedoch nicht statistisch bedeutsam, da beide Störungen in der Gesamtstichprobe relativ häufig vorkommen. Ähnliches gilt auch für die rezidivierende depressive Störung bei Generalisierter Angststörung und noch weitere Störungskombinationen. Das alleinige Betrachten der Überschneidungsraten würde zu missverständlich hohen Annahmen gemeinsamen Auftretens führen, die sich unter Einbezug der Basisraten wieder relativieren.

Als Nachteil der gewählten Störungseinteilung (eine Person fällt in die jeweilige Störungsgruppe, sofern mindestens eine Diagnose aus dem definierten Bereich gestellt wurde) ist zu nennen, dass damit lediglich eine Untersuchung von Komorbidität zwischen, jedoch nicht

innerhalb von Störungsgruppen möglich ist. Die häufig berichtete Komorbidität etwa von Persönlichkeitsstörungen untereinander (Fiedler, 1998; Stuart et al., 1998; Zimmerman, 1994) konnte somit nicht berücksichtigt werden. Als methodische Einschränkung der tetrachorischen Korrelationsanalysen ist weiterhin festzuhalten, dass hier sowohl Patienten der IG als auch der KG eingingen. Auf eine getrennte Betrachtung wurde aufgrund der ohnehin oft geringen Zellbesetzungen verzichtet. Jedoch wäre es aufschlussreich, ähnliche Analysen für Datensätze durchzuführen, die auf einer jeweils einheitlichen Diagnostik beruhen.

### **6.1.3 Zur Identifikation latenter Klassen von Diagnosekombinationen**

Im Rahmen einer multivariaten Betrachtungsweise wurden Latent Class Analysen (LCA) durchgeführt, um aus den 16 definierten Diagnosegruppen plus Restkategorie nicht direkt beobachtbare Klassen von Diagnosekombinationen zu identifizieren.

Im gewählten 5-Klassen-Modell sind vier Klassen durch jeweils eine Störung (depressive Episode, rezidivierende depressive Störung, Anpassungsstörung, Angst/Depression gemischt) charakterisiert, die alle Patienten in dieser Klasse erfüllen; die Wahrscheinlichkeit für weitere Diagnosen fällt jeweils gering aus. Die Ergebnisse der tetrachorischen Korrelationen werden bestätigt, indem der Anteil der Patienten mit Persönlichkeitsstörung in Klasse 1 am größten ist, jener Klasse die vor allem durch das Vorhandensein rezidivierender depressiver Störungen gekennzeichnet ist. Anpassungsstörungen haben sehr geringe Überschneidungsraten mit anderen Diagnosen, was sich auch in der identifizierten Klasse 5 widerspiegelt. Klasse 2 kann als „Restklasse“ verstanden werden, die vorrangig jene Patienten zusammenfasst, die keine der vier genannten populären Diagnosen aufweisen. Gleichzeitig stellt Klasse 2 mit 315 Personen die größte Klasse dar, weshalb es Gewinn bringend sein könnte, diese Klasse genauer im Hinblick auf spezifische Untergruppierungen zu untersuchen. Möglicherweise wäre hier auch eine weitere Zusammenfassung der Diagnosegruppen hilfreich, da einige Gruppen, wie etwa bipolare Störungen oder akute Belastungsstörungen, nur sehr schwach besetzt sind. Insgesamt werden 183 verschiedene Diagnosekombinationen gefunden, davon tritt jede einzelne im Schnitt 5.3 Mal auf. Stellt man dies der Fülle (61489) potenzieller Diagnosekombinationen in diesem Datensatz gegenüber, so legt dies nahe, dass die Anzahl vergebener Diagnosen und deren Kombination nicht rein zufällig zustande kommen, sondern dass bestimmte Diagnosekombinationen überzufällig gehäuft bestehen. Im Gegensatz zu den tetrachorischen Korrelationen konnte

hier die Gruppenzugehörigkeit als Kovariate einbezogen werden. Dabei ergeben sich bedeutsame Unterschiede in der Verteilung der Klassen auf IG und KG, was die Ergebnisse zum veränderten Diagnoseverhalten unter Verwendung von Diagnosechecklisten in der IG noch untermauert.

Zusammenfassend zu den Ergebnissen der Latent Class Analysen ist zu sagen, dass sich zwar eindeutig fünf Klassen identifizieren lassen, die sich jedoch als insgesamt wenig prägnant herausstellen. Ableitungen für die klinische Praxis, etwa besondere Aufmerksamkeit auf mögliche weitere Diagnosen bei Vorliegen einer bestimmten Störung zu lenken, lassen sich daraus nur schwer ziehen. Die sich hier ergebenden Klassen können als Resultat der im Datensatz insgesamt relativ geringen Komorbiditätsrate (im Schnitt 1.57 Diagnosen pro Patient) gesehen werden, wodurch die drei häufigsten Diagnosen (depressive Episode, Anpassungsstörung, rezidivierende depressive Störung) zwangsläufig die größte differenzierende Funktion einnehmen.

#### **6.1.4 Zum Status komorbider Patienten zu Therapiebeginn**

Im Rahmen zweifaktorieller Varianzanalysen wurden Personen mit und ohne Komorbidität im Hinblick auf die selbst- und fremdeingeschätzte Belastung zu Therapiebeginn untersucht. Der Faktor Komorbidität wurde dabei zum einen als latente Diagnoseklassen, wie über die LCA identifiziert, zum anderen als vier vorab definierte Muster operationalisiert. Aufgrund der bereits dargestellten Unterschiede zwischen IG und KG wurde die Gruppenzugehörigkeit als weiterer Faktor aufgenommen. Um die Probleme multiplen Testens zu umgehen, wurde vorab ein zusammengesetztes Ergebnismaß sowohl für die Belastung im Selbst- als auch im Fremdbereich berechnet.

In Übereinstimmung mit den Hypothesen ergibt sich für beide abhängigen Variablen ein signifikanter Haupteffekt „Komorbidität“. Patienten mit nur einer Diagnose geben eine geringere Belastung an als Patienten aller drei Komorbiditätsgruppen. Dieses Ergebnis wird weitestgehend durch die Sichtweise der Therapeuten gestützt und untermauert bisherige Befunde zur Belastungsschwere bei Vorliegen komorbider Diagnosen (u. a. Kessler et al., 2005). Patienten mit komorbider Persönlichkeitsstörung erweisen sich nach eigenem Bekunden und nach Auskunft ihrer Therapeuten als die am stärksten belastete Gruppe; ein bedeutsamer Unterschied zu Patienten mit Komorbidität aus Angst und affektiven Störun-

gen ergibt sich allerdings nur für den Fremdbbericht. Ein Grund hierfür liegt möglicherweise darin, dass in das zusammengesetzte Maß zum Selbstbericht explizit auch interpersonale Probleme (IIP) aufgenommen wurden, die nicht nur bei Persönlichkeitsstörungen (Hardy et al., 1995), sondern auch bei affektiven Erkrankungen massiv sein können (u. a. Davila, Hammen, Burge, Paley, & Daley, 1995). In die Therapeutenangaben ging zwar der Beeinträchtigungsschwere-Score für den sozialkommunikativen Bereich ein, der interpersonale Aspekt ist hier insgesamt jedoch geringer gewichtet. Möglicherweise unterliegen Therapeuten auch einem Bias, indem bei Vorliegen einer Persönlichkeitsstörung, welche oft als schwerer behandelbar als reine Achse-I-Störungen angesehen werden, unmittelbar eine große Belastung unterstellt wird. Diese kann zwar bei Personen, die eine Behandlung aufsuchen gegeben sein, muss aber gerade bei Persönlichkeitsstörungen im subjektiven Erleben nicht zwangsläufig vorliegen. Im Vergleich dazu mag die subjektiv empfundene Belastungsschwere bei Patienten mit Diagnosen aus den Bereichen Angst- und affektive Störungen unterschätzt werden. Die besonders hohe Belastung von Patienten mit komorbider Persönlichkeitsstörung zeigt sich auch unter Einbezug der latenten Diagnoseklassen als Faktor. Hier ergeben sich für Patienten der Klasse 1 (charakterisiert durch rezidivierende depressive Störung mit häufiger paralleler Persönlichkeitsstörung) die höchsten, für Patienten der Klasse 5 (charakterisiert durch Anpassungsstörungen) die geringsten Belastungswerte. Bemerkenswert ist weiterhin, dass sich im Fremd-, nicht jedoch im Selbstbericht, ein signifikanter Haupteffekt „Gruppe“ ergibt, indem von Therapeutenseite die Belastung in der IG im Vergleich zur KG geringer eingestuft wird. Bezüglich der Ursachen für diesen Unterschied kann nur spekuliert werden. Denkbar ist, dass IG-Therapeuten, die ihre Patienten als sehr schwer belastet ansahen, diesen die Teilnahme am Modellvorhaben nicht angeboten haben, da sie den Mehraufwand für nicht zumutbar hielten. In der KG hingegen war eine Teilnahme sowohl für Patienten als auch für Therapeuten mit vergleichsweise geringem Aufwand verbunden, so dass hier möglicherweise mehr Patienten eingingen, die aus Therapeutensicht schwer belastet sind. Die Befragungen zur Nichtteilnahme (Wittmann et al., 2011, S. 52) stützen diese Vermutung nicht, wobei hierzu lediglich Angaben für knapp ein Zehntel aller Nichtteilnehmer vorliegen. Es muss auch betont werden, dass sich diese Unterschiede nicht zeigen, wenn Patienten selbst Auskunft über ihre Symptom- und allgemeine Belastung geben. Befunde zur Vorhersage des Therapieergebnisses legen nahe, dass insbesondere die Therapeuteneinschätzung fehleranfällig ist (u. a. Hannan et al.,

2005), weshalb in diesen Studien, so auch zur Generierung von Rückmeldungen im TK-Projekt, auf Patientenangaben zurückgegriffen wird.

### **6.1.5 Zu den Verlaufsfragestellungen**

Zur Untersuchung des Einflusses komorbider Diagnosen auf den Therapieverlauf wurden latente Wachstumsanalysen durchgeführt, wobei Effekte der Zeit und Komorbidität erwartet wurden.

Zunächst einmal ist festzuhalten, dass die Psychotherapie für alle Patientengruppen, unabhängig von der Art und Anzahl der Diagnosen, wirksam ist. Über die Zwischenmessungen hinweg erreichen sie im Durchschnitt eine bedeutsame Verbesserung in beiden untersuchten Ergebnismaßen. Jedoch wird kein bedeutsamer Zusammenhang zwischen der Ausgangslage und der Veränderung in BSI und IIP gefunden. Dies widerspricht der Annahme, dass zu Beginn sehr stark belastete Patienten, worunter auch komorbide Patienten fallen (s. Ergebnisse zur Hypothese 4), mehr „Raum zur Verbesserung“ haben und daher besonders stark von einer Psychotherapie profitieren könnten. Demgegenüber ist nach diesen Ergebnissen eine hohe Ausgangsbelastung jedoch genauso wenig mit einem komplizierteren oder weniger erfolgreichen Verlauf gleichzusetzen; die Ausprägung, mit der Patienten starten, erlaubt alleine keine Vorhersage darüber, ob sie sich besonders schnell oder langsam verbessern.

Im Rahmen konditionaler Wachstumsmodelle wird der Befund einer erhöhten selbstberichteten Belastung unter Komorbidität noch einmal bestätigt. Patienten mit komorbider Persönlichkeitsstörung oder mit einer Kombination aus Angst und affektiven Störungen beginnen die Therapie in beiden Ergebnismaßen BSI und IIP mit signifikant größeren Belastungswerten als Patienten mit nur einer Diagnose. Daneben zeigt sich der besondere Stellenwert von interpersonaler Belastung bei Persönlichkeitsstörungen: bezüglich interpersonaler Belastung, nicht aber bezüglich allgemeiner psychischer Belastung ergeben sich bedeutsame Unterschiede in den Ausgangswerten zwischen der Gruppe Komorbidität mit Persönlichkeitsstörung und den anderen beiden Komorbiditätsgruppen. Jedoch schlagen sich diese Eingangsdifferenzen nicht im Verlauf nieder: sowohl bei Vorliegen einer komorbiden Persönlichkeitsstörung als auch bei einer Kombination aus Angst und affektiven Störungen werden vergleichbare Verbesserungsraten erzielt wie bei Einfachdiagnosen.

Die Beobachtung paralleler Verbesserung wird auch an anderer Stelle berichtet (u. a. Diguier, Barber, & Luborsky, 1993; Dreessen et al., 1994; Joormann et al., 2005; Rief et al., 2000). Bezogen auf die Gruppe anderer Komorbiditäten ist festzuhalten, dass diese sich in ihrer Eingangsbelastung in BSI und IIP nicht bedeutsam von der Patientengruppe mit nur einer Diagnose unterscheiden; allerdings ergibt sich hier in beiden Ergebnismaßen ein bedeutsamer Einfluss auf die Veränderungsrate. Trotz vergleichbarer psychometrischer Ausgangslage zeigen Patienten mit nur einer Diagnose eine schnellere Verbesserung als solche aus der Gruppe „andere Komorbidität“. Diese Ergebnisse zusammenfassend hat Komorbidität durchaus Relevanz für Therapieverläufe, und zwar in Abhängigkeit der Art der Diagnosekombination. Liegt eine parallele Persönlichkeitsstörung vor oder ein Muster aus Angst und affektiven Störungen, so ist nicht zwangsläufig mit einer langsameren Verbesserung als unter Einfachdiagnosen zu rechnen. Aufgrund der in der Regel erhöhten Belastungswerte zu Beginn der Therapie muss jedoch bei gleichbleibender Therapielänge davon ausgegangen werden, dass zwar die gleiche relative Verbesserung, jedoch ein schlechteres absolutes Therapieergebnis erreicht werden. Auch bei im Vergleich mit monomorbiden Patienten ähnlichen Ausgangsvoraussetzungen scheint sich „Komorbidität mit anderen Diagnosekombinationen“ negativ auszuwirken, indem diese Patienten mehr Zeit benötigen um ein ähnliches Therapieergebnis wie solche mit Einfachdiagnosen zu erreichen. Diese Resultate widersprechen jenen von Newman et al. (2010) und in Teilen der Arbeit von Joormann et al. (2005), wo unter Komorbidität steilere Verbesserungskurven als bei Einfachdiagnosen gefunden wurden. Unter Betrachtung angstspezifischer Ergebnismaße fanden Joormann et al. (2005) jedoch keinen Zusammenhang zwischen Verbesserungsverläufen und dem parallelen Vorliegen von Angststörungen und Depression – ein Befund, welcher sich mit den Ergebnissen dieser Arbeit deckt. Die Tatsache, dass sich Patienten aus der Gruppe „andere Komorbidität“ in beiden Maßen langsamer verbessern als nicht komorbide Patienten weist darauf hin, dass die Einschränkungen beim Vorliegen von Mehrfachdiagnosen sich nicht zwangsläufig in einer erhöhten Ausgangslage widerspiegeln müssen, sondern dass offensichtlich komplexere Faktoren einwirken. Alter, Geschlecht und allgemeines Funktionsniveau können Unterschiede im Verlauf dieser beiden Gruppen nicht erklären (s. a. Anhang, S. 173). Jedoch ist ein Einblick dahingehend, welche Störungen sich in den Gruppen tatsächlich wiederfinden, hilfreich. Sowohl in der Gruppe „keine Komorbidität“ als auch in der Gruppe „andere Komorbidität“ finden sich besonders häufig die Diagnosen Akute Belastungsstörung und Depressive Episode. Jedoch ist die Gruppe

„andere Komorbidität“ darüber hinaus charakterisiert durch das häufige Vorkommen insbesondere Somatoformer Störungen, welche hier nur in sehr geringem Ausmaß als alleinige Diagnose vergeben werden (s. a. Anhang, S. 172f.). Somatoforme Störungen sind geprägt durch die Wahrnehmung vornehmlich körperlicher Beschwerden, welche durch BSI und IIP nur in Teilen erfasst werden – dies könnte die vergleichsweise geringe Belastung in diesen Ergebnismaßen zu Therapiebeginn erklären. Gleichzeitig weisen Patienten mit somatoformen Störungen oft einen bereits chronifizierten Verlauf auf, wenn sie sich in psychotherapeutische Behandlung begeben (Rief & Hiller, 2011), so dass hier die Wahrscheinlichkeit für eine schnelle Besserung gering ist.

In einem weiteren Schritt wurden kategoriale latente Variablen in die Wachstumsmodelle eingefügt um zu prüfen, ob sich latente Klassen der Veränderung in BSI und IIP identifizieren lassen. Für den BSI ergeben sich vier Klassen, geprägt durch spezifische Eingangs- und Verlaufscharakteristika: Klasse 1 stellt die größte Patientengruppe dar, und umfasst solche Patienten, die die Behandlung mit vergleichsweise geringer Ausgangsbelastung beginnen und sich im Verlauf etwas weiter verbessern. Klasse 2 beginnt mit mittleren Belastungswerten und zeigt im Mittel keine Veränderung über die Zeit. Klasse 3 umfasst die kleinste Gruppe und beinhaltet Patienten, welche mit sehr hoher Belastung starten und sich verschlechtern. Patienten der Klasse 4 beginnen mit im Mittel ähnlich hohen BSI-Werten, verbessern sich aber deutlich im Verlauf. Für den IIP als Ergebnismaß war es nicht möglich, latente Verlaufsklassen zu identifizieren.

Vergleichbar schnell verbesserte Patientengruppen wie Klasse 4 im BSI lassen sich immer wieder in psychotherapeutischen Verlaufsdaten finden, wenn auch meist in einer höheren Größenordnung (Lutz, Stulz, et al., 2009, 2007; Stulz & Lutz, 2007; Stulz, Lutz, Leach, Lucock, & Barkham, 2007). Zu Korrelaten der Mitgliedschaft insbesondere in der vergleichsweise früh und schnell verbesserten Gruppe von Patienten liegen erst wenige Untersuchungen vor (z. B. Stulz, Lutz, Leach, et al., 2007). Aus den Ergebnissen der vorliegenden Arbeit lassen sich lediglich Tendenzen des Zusammenhangs zwischen Komorbiditätsgruppen und der Mitgliedschaft in latenten Veränderungsklassen ableiten. So finden sich in den Klassen 2 („unverändert“) und 3 („verschlechtert“) gehäuft Patienten mit komorbider Persönlichkeitsstörung und in Klasse 4 („schnell verbessert“) tendenziell mehr Patienten mit einer Kombination aus Angst und affektiven Störungen. Es muss jedoch betont werden, dass sich alle Komorbiditätsmuster in allen Klassen zeigen, dass also Komorbidität,

wenn überhaupt, nur ein schwacher Prädiktor für die Mitgliedschaft in latenten Veränderungsklassen sein kann. Weitere Forschung ist notwendig um besser zu verstehen, was genau insbesondere die schnell verbesserten Patienten charakterisiert, auch um die im Rahmen von Rückmeldesystemen umgesetzten differenziellen Verlaufsvorhersagen zu optimieren. Potenziell erklärende Variablen könnten etwa die Therapiemotivation oder Erwartungen des Patienten, aber auch die Einnahme von Medikamenten oder Spontanremissionseffekte sein. Eine weitere Erklärungsmöglichkeit könnte am Konzept der „sudden gains“ ansetzen: womöglich haben in den schnell verbesserten Gruppen vergleichsweise mehr Patienten eine große Symptomreduktion zwischen zwei Behandlungssitzungen erlebt, welche in positivem Zusammenhang zum Therapieergebnis steht (Stiles et al., 2003; Tang & DeRubeis, 1999). Zur Überprüfung dieser Annahme wäre es notwendig, die hier durchgeführten Analysen zu kombinieren mit der Untersuchung sitzungsweise erhobener Daten.

Bezüglich der Verlaufsanalysen sind einige Einschränkungen zu nennen. Zum einen konnten diese Analysen zwangsläufig nur auf einen stark reduzierten Datensatz angewendet werden (Therapeut mit verhaltenstherapeutischer Ausrichtung, IG, diagnostische Information und ein Minimum eingehender Messzeitpunkte mussten gegeben sein). Im Vergleich zum TK-Gesamtdatensatz sind die hier untersuchten Patienten etwas älter und etwas stärker belastet (s. Tabelle 1); daneben bestehen möglicherweise Unterschiede in weiteren Variablen, die hier nicht untersucht wurden. Ob die gefundenen Ergebnisse typisch für ambulante Psychotherapiepatienten in Deutschland sind, lässt sich also an dieser Stelle nicht sagen. Hierzu wären weitere Studien notwendig, vorrangig mit einem Studiendesign, welches spezifischer auf die Untersuchung von Verlaufsfragestellungen ausgerichtet ist, als dies im TK-Modellvorhaben gegeben war (s. Abschnitt 6.2).

Ein weiteres Befund ist, dass vermutlich nicht alle Ergebnismaße gleichermaßen zur Identifikation latenter Veränderungsgruppen geeignet sind. Während sich unter dem BSI als Ergebnismaß vier distinkte Verlaufsklassen mit guten Klassifikationsvorhersagen herausbilden, ist es mit dem Ergebnismaß IIP nicht möglich, klar voneinander abgegrenzte Patientensubgruppen zu bestimmen. Wenn den Daten mehrere kategoriale latente Variablen aufgezwungen werden, ergibt sich nach allen relevanten Indizes eine ungünstige Modellpassung. Dies bedeutet nicht, dass die einbezogenen Patienten sich hinsichtlich ihrer Ausgangs- und Veränderungswerte im IIP nicht unterscheiden, dass also keine Heterogenität

besteht. Bei Betrachtung der deskriptiven Statistiken wird deutlich, dass zu allen eingehenden Messzeitpunkten, genauso wie für den BSI, eine beträchtliche Streuung in den Daten vorliegt. So kann jeder Patient einen eigenen, spezifischen Verlauf aufweisen, wobei diese Verläufe sich jedoch nicht in typische Subgruppen zusammenfassen lassen. Die Unterschiedlichkeit in BSI und IIP hinsichtlich der Möglichkeit, latente Veränderungsgruppen zu identifizieren, mag auch in den Verteilungseigenschaften der Ausgangsvariablen begründet liegen (dazu Bauer & Curran, 2003).

Ein weiterer Punkt betrifft die Aufnahme der latenten Veränderungsklassen als Prädiktor in die Regressionsanalysen. Aufgrund der gegebenen Datenstruktur (Erhebungen zu Beginn der Therapie, zur 10. und 20. Sitzung usw.) und der gewählten Methodik musste der Zeitraum zur Modellierung der frühen Veränderungsklassen relativ lang, nämlich als Veränderung bis zur 20. Sitzung, definiert werden. Der Mittelwert durchgeführter Sitzungen lag zwischen 38 (Klasse 1) und 50 Sitzungen (Klasse 4). Somit liegt die Veränderung, welche in die Generierung der Klassen eingeht, vergleichsweise nahe an der zum Therapieende erreichten Verbesserung, was die hohe Prädiktionskraft der Verlaufsklassen erklärt. Von größerer klinischer Relevanz und auch für die Verbesserung von Vorhersage- und Rückmeldesystemen besser nutzbar wäre Information über Verbesserung innerhalb sehr früher Phasen der Therapie. Dass sich Veränderungsmuster schon nach den ersten Therapiesitzungen modellieren lassen und prädiktiven Wert besitzen, konnte in früheren Studien gezeigt werden (Lutz, Stulz, et al., 2007; Stulz, Lutz, et al., 2007). Jedoch stammen die dazu verwendeten Daten aus dem US-amerikanischen und britischen Versorgungskontext und umfassen im Durchschnitt deutlich kürzere Therapiedauern, weshalb die daraus gewonnenen Ergebnisse nur eingeschränkt auf die deutsche Psychotherapielandschaft übertragbar sind. Demgegenüber stellt die vorliegende Untersuchung die erste dieser Art an einem in der deutschen Versorgungsroutine erhobenen Datensatz dar und lässt Folgeuntersuchungen vielversprechend erscheinen, in welchen eine Identifikation früher Veränderungen auf sitzungsweise erhobenen Daten beruhen sollte.

Es bedarf darüber hinaus auch einer grundsätzlichen Diskussion des Konzeptes früher Veränderungen („early changes“) in der Therapie. Wie bereits in den einleitenden Kapiteln angedeutet, besteht hierzu keine einheitliche Definition. So werden sowohl das Mindestmaß an Veränderung, als auch der Begriff „früh“, sprich der maximal betrachtete Zeitraum bis zur Erreichung dieser Veränderung, unterschiedlich konzipiert (etwa Haas et al., 2002;

Hayes et al., 2007; Lutz, Stulz, et al., 2007; Renaud et al., 1998), natürlich auch in Abhängigkeit von der jeweiligen Datenlage. Interpretationsspielraum lässt vor allem die Frage, ab wann von einer frühen Veränderung gesprochen werden kann. Eine Möglichkeit besteht darin, eine bestimmte Sitzungsanzahl festzulegen, unabhängig von der letztendlichen Dauer der Behandlung. Führt man diesen Gedanken weiter, so stellt sich die inhaltliche Frage, wann genau eine Behandlung beginnt: bereits mit Anmeldung zur Psychotherapie, mit Beginn der probatorischen Sitzungen oder mit der ersten bewilligten Therapiestunde? Vieles spricht dafür, frühe Veränderung ab dem ersten therapeutischen Kontakt zu modellieren, um auch unspezifische Wirkfaktoren zu erfassen. Daneben ist zu überlegen, frühe Veränderung in Relation zu den im jeweiligen Versorgungskontext üblichen Therapiedauern zu setzen. Bei in den USA üblichen kurzen Therapiedauern von oft nur sechs bis sieben Sitzungen (u. a. Haas et al., 2002) entspricht die Definition früher Veränderungen über die ersten drei Sitzungen im Durchschnitt bereits der Hälfte der insgesamt dargebotenen Behandlung – das Verhältnis von früher Veränderung zu durchschnittlicher Behandlungsdauer wäre also dasselbe wie in der im Rahmen dieser Arbeit durchgeführten Untersuchungen. Aus inhaltlichen Gesichtspunkten ist es fraglos wichtig, sich abzeichnende Veränderungen so früh und so genau wie möglich zu erfassen, um sie in die differenzielle Therapieplanung mit einzubeziehen. Aufgrund sich stetig verbessernder technischer Möglichkeiten und der im Rahmen von Wissenschaftler-Praktiker-Netzwerken erhobenen breiten Referenzdatenbanken ist auch eine unmittelbare Analyse von in frühen Therapiephasen gewonnenen Daten und deren Rückmeldung in den Therapieprozess möglich. Die Methode des Growth Mixture Modeling scheint dabei insbesondere geeignet, um Patienten mit besonders gutem Verlauf frühzeitig zu identifizieren und von denjenigen Patienten zu differenzieren, welche zu Beginn der Therapie ähnlich hohe Belastungswerte angeben, sich im Verlauf jedoch kaum oder gar nicht verbessern. Da die latenten Veränderungsgruppen bisher ausschließlich über Selbstberichtmaße der Patienten identifiziert wurden, muss auch die Möglichkeit bedacht werden, dass sich die schnell verbesserte Gruppe als Resultat des Antwortverhaltens ergibt. So könnten sich einige Patienten zu Beginn der Therapie als besonders belastet darstellen, um ihre Hilfsbedürftigkeit und die Notwendigkeit einer sofortigen Behandlung zu demonstrieren. Jedoch würde dies nicht dem Konzept der frühen Verbesserungen widersprechen; vielmehr wäre es notwendig, auch diese Patientengruppe zuverlässig zu identifizieren.

### 6.1.6 Zur Vorhersage des Therapieergebnisses und der Anzahl Sitzungen

Im Rahmen der Prädiktion des Therapieergebnisses unter Einbezug der identifizierten Verlaufsklassen wurden die Ergebnisse früherer Studien bestätigt, die finden, dass die Mitgliedschaft in der Gruppe früh verbesserter Patienten ein positives Ergebnis vorhersagt beziehungsweise dass das Nicht-Vorliegen früher Verbesserung mit einem eher negativen Ergebnis assoziiert ist (Haas et al., 2002; Hannan et al., 2005; Kuyken, 2004; Lambert et al., 2002; Lutz, Stulz, et al., 2009; Stulz & Lutz, 2007). Die Mitgliedschaft in Klasse 4 („schnell verbessert“) im Vergleich zu den anderen Klassen resultiert in einer signifikant größeren Prä-Post-Differenz im Ergebnismaß BSI sowohl zum Ende der Behandlung als auch zum Katamnesezeitpunkt. Dass es sinnvoll ist, zur Vorhersage des Therapieergebnisses zusätzlich zur Eingangsbelastung auch die Mitgliedschaft in latenten Veränderungsklassen zu nutzen, konnte bereits in früheren Studien gezeigt werden (Lutz, Stulz, & Köck, 2009). Die zusätzliche Aufnahme der Komorbiditätsvariablen in die Regressionsanalysen führt nicht zu einer signifikant verbesserten Vorhersage des Therapieergebnisses. Neben Eingangsbelastung und Mitgliedschaft in latenten Verlaufsklassen besitzt diese Information keine zusätzliche Erklärungskraft und auch wenn die Variable Komorbidität im vierten Schritt vor den Verlaufsklassen eingefügt wird, leistet sie keine bedeutsame Varianzaufklärung. Anders verhält es sich bei der Vorhersage der Anzahl in Anspruch genommener Therapiesitzungen. Die Vorhersagekraft der aufgenommenen Variablen ist hier insgesamt geringer, wobei jedoch die Anfangsbelastung keine entscheidende Rolle mehr einnimmt und stattdessen die Variablen Verlaufsklasse und Komorbidität große Varianzaufklärung leisten. Die Mitgliedschaft in Klasse 2 („unverändert“) oder 3 („verschlechtert“) ist mit signifikant mehr Therapiesitzungen assoziiert als die Mitgliedschaft in Klasse 4 („schnell verbessert“). Ein ähnliches Ergebnis findet sich auch bei Stulz et al. (2007). Daneben führt das Vorliegen komorbider Persönlichkeitsstörungen im Vergleich zu reinen Diagnosen zu bedeutsam längeren Therapien.

Zusammengefasst eignet sich Komorbidität bei Vorliegen von Information über frühe Veränderungen nicht zur Vorhersage des Therapieergebnisses (definiert als Differenzmaß). Dieses Resultat war zu erwarten angesichts der nicht signifikant unterschiedlichen Verbesserungsverläufe in den latenten Wachstumsanalysen. Positiv formuliert, können Patienten mit mehreren Diagnosen im gleichen Ausmaß, im Sinne der erreichten Prä-Post-Differenz, von einer Psychotherapie (hier Verhaltenstherapie) profitieren, wie Patienten mit nur einer

Diagnose. Allerdings sollte zur Therapieplanung insbesondere die Komorbidität von Persönlichkeitsstörungen bedacht werden, da diese mit einem längeren Verlauf assoziiert sind. Erwähnenswert ist weiterhin, dass zur Vorhersage der Prä-Post-Differenz im BSI die Anzahl in Anspruch genommener Therapiesitzungen keinen bedeutsamen Einfluss hat, wenn bereits die Eingangsbelastung berücksichtigt wurde. Dieses Ergebnis widerspricht den Annahmen des dose-effect-Modells (Howard, 1986), welches davon ausgeht, dass die im Laufe einer Therapie erreichte Besserung mit der Sitzungsanzahl assoziiert ist.

Der besonders große Einfluss der Eingangsbelastung auf das Therapieergebnis zeigt sich auch im zweiten Block von Regressionsanalysen. Die hier eingehenden Daten unterscheiden sich insofern von den ersten Regressionen, als dass hier keine Einschränkungen bezüglich der Anzahl vorliegender Zwischenmessungen gemacht wurden. Voraussetzung war die Zuordnung zur IG, das Vorliegen diagnostischer Information und eine Beendigung der Therapie im Studienzeitraum. Als Kriterium dienten die End- und Katamnesewerte in BSI, IIP und BDI. Für alle drei Ergebnismaße erweist sich die anfängliche Symptombelastung als herausragender Prädiktor – ein Ergebnis, welches so auch in der Literatur zu finden ist (Clarkin & Levy, 2004; Kuyken, 2004; Lutz et al., 1999). Zusätzlichen Erklärungswert besitzt die Anzahl durchgeführter Therapiesitzungen für die Vorhersage des Therapieerfolgs im BSI zum Therapieende und zum Katamnesezeitpunkt, hier allerdings nur marginal, sowie in größerem Ausmaß für den BDI in der Katamnese. Jedoch fällt letzterer Zusammenhang nicht, wie laut dose-effect-Modell (Howard et al., 1986) anzunehmen, positiv aus (je höher die Dosis, desto besser das absolute Therapieergebnis). Stattdessen sind mehr Therapiesitzungen mit einem schlechterem Ergebnis in der Katamnese, sprich höheren BDI-Werten, assoziiert; für die BDI-Werte zum Therapieende ergibt sich kein signifikanter Einfluss der Dauer. Eine mögliche Erklärung für dieses zunächst kontraintuitiv erscheinende Ergebnis kann im phasenhaften Verlauf der dahinterstehenden Diagnosen zu sehen sein. Möglicherweise werden Therapien bei rezidivierenden depressiven Störungen oder Zykllothymia dann beendet, wenn die Person sich besser fühlt, und dies relativ unabhängig von der vergangenen Therapiezeit. Jedoch erscheint es gerade hier besonders wichtig, auf die Stabilität der erreichten Besserung zu achten und es besteht die Gefahr einer voreiligen Beendigung, sobald ein erstes Abklingen der Symptombelastung bemerkbar ist. Bei langandauernden affektiven Störungen mag es folglich zwar länger dauern, bis eine erste Verbesserung erreicht ist, wobei bei einer Therapiebeendigung zu diesem Zeitpunkt das Er-

gebnis nicht schlechter ausfällt als bei kürzeren Therapien (kein signifikanter Einfluss der Dauer auf das unmittelbare Ergebnis). Jedoch ist möglicherweise gerade hier die Gefahr eines Rezidivs erhöht, so dass sich in der Katamnese wiederum ein positiver Zusammenhang zwischen Sitzungsanzahl und Symptomatik ergibt. Diese Ergebnisse machen deutlich, dass es wichtig ist, zwischen unmittelbaren und längerfristigen Therapieerfolgen zu differenzieren. So ist etwa bei (Eldredge et al., 1997) ein positiver Zusammenhang zwischen der Therapiedauer und dem Ansprechen auf Behandlung zu finden (hier profitieren ursprüngliche Nonresponder im Hinblick auf eine Symptomreduktion von längeren Therapien), allerdings schließt diese Untersuchung keine follow-up-Daten mit ein.

Für das Therapieergebnis im IIP scheint die Sitzungsanzahl zu beiden Zeitpunkten ohne Einfluss, sofern bereits die Eingangsbelastung berücksichtigt wurde. Dies kann in gewissem Widerspruch zu den Annahmen des Phasenmodells (Howard et al., 1993) verstanden werden, wonach die Verbesserung interpersonaler Probleme am ehesten der Rehabilitationsphase zuzuordnen ist, welche erst spät im therapeutischen Prozess durchlaufen wird.

Zusammenfassend ist festzuhalten, dass nach den hier vorgestellten Resultaten der Einfluss vergangener Therapiesitzungen auf das Ergebnis in Abhängigkeit des gewählten Ergebnismaßes variiert, dass dieser Einfluss – wenn bereits die Eingangsbelastung berücksichtigt wurde – insgesamt aber gering ausfällt und eher in eine negative Richtung weist. Insgesamt können diese Ergebnisse also die Annahmen des dose-effect-Modells (Howard et al., 1986) nicht stützen. In eine ähnliche Richtung weisen auch die Befunde von Baldwin et al. (2009). Sie untersuchten, ebenfalls in einem naturalistischen Datensatz, den Zusammenhang zwischen der Anzahl an Therapiesitzungen und der Veränderungsrate in einem globalen Ergebnismaß (OQ-45; Lambert et al., 2004). Es resultierte eine größere empirische Unterstützung für das good-enough-level-Modell (Barkham, Connell et al., 2006) als für das dose-effect-Modell (Howard et al., 1986) und die Autoren schließen damit, dass die Dosis keinen Prädiktor für das Ansprechen auf eine Behandlung darstellt, sondern vielmehr einen Marker für die Geschwindigkeit der Verbesserung. Für das hier gefundene tendenziell schlechtere Abschneiden bei längeren Therapien könnten auch nicht berücksichtigte Drittvariablen eine Rolle spielen. So mögen kompliziertere individuelle Problemstellungen sowohl mit einer längeren Therapiedauer als auch mit einem insgesamt schlechteren Abschneiden assoziiert sein. Allerdings können solche möglichen Komplikationen nicht alleine in der Komorbidität begründet liegen, denn auch wenn im Rahmen schrittwei-

ser Regressionen die Komorbiditätsvariable vorrangig eingefügt wird, bleibt der Einfluss der Sitzungsanzahl bestehen.

Die Aufnahme von Komorbidität als Prädiktor des Therapieergebnisses führt nicht zu einem signifikanten Anstieg erklärter Varianz; jedoch ergeben sich bedeutsame Teilergebnisse für zwei Untergruppen von Komorbidität. So weisen Personen mit einer Kombination aus Angst- und affektiven Störungen zum Ende der Therapie höhere Werte im Bereich interpersonaler Probleme auf als nicht komorbide Patienten und Personen mit komorbider Persönlichkeitsstörung erreichen im störungsspezifischen Maß BDI zum Therapieende größere Belastungswerte als nicht komorbide Personen. Für die anderen Ergebnismaße ergibt sich ein nicht signifikanter Einfluss von Komorbidität, jedoch mit Tendenz zu einem schlechteren Resultat bei Vorliegen komorbider Diagnosen. Obwohl also komorbide Patienten absolut gesehen schlechter abschneiden als nicht komorbide Patienten, eignet sich diese Variable in der Regel nicht als Prädiktor des Therapieergebnisses, wenn die Eingangsbelastung bereits berücksichtigt wurde. Somit kann das schlechtere Ergebnis interpretiert werden als zu großen Teilen zurückführbar auf erhöhte Belastungswerte zu Beginn der Therapie und nicht als Folge von mit Komorbidität verbundenen weiteren Komplikationen. Auch zur Vorhersage der Ergebnisse zur zweiten und dritten Zwischenerhebung erweist sich die Eingangsbelastung im jeweiligen Instrument als stärkster und Komorbidität als nur schwacher Prädiktor.

Der Anteil reliabel und klinisch signifikant verbesserter Patienten in Abhängigkeit der Komorbiditätsgruppe schwankt mit dem gewählten Ergebnismaß, ist aber alles in allem zum Therapieende für Patienten mit nur einer Diagnose leicht erhöht und entspricht somit der in der Komorbiditätsliteratur oft berichteten Beobachtung eines schlechteren Abschneidens unter Mehrfachdiagnosen (u. a. Dreessen et al., 1994; Noyes et al., 1990; Reich & Green, 1991; Viinamäki et al., 2002). Jedoch erreichen auch Patienten mit Komorbidität bemerkenswert hohe Verbesserungsraten zum End- und Katamnesezeitpunkt. Dass hier Komorbidität ohne erheblichen negativen Einfluss auf das Therapieergebnis zum Katamnesezeitpunkt ist, widerspricht den Ergebnissen von Ilardi, Craighead und Evans (1997), welche sehr viel höhere Rückfallraten bei Patienten mit komorbider Persönlichkeitsstörung im Vergleich zu nicht komorbiden Patienten fanden. In der zitierten Studie fanden die Katamneseerhebungen im Schnitt 50 Monate nach Entlassung statt, also nach einem sehr viel längeren Zeitraum als im TK-Modellvorhaben. Möglicherweise hätten sich in den unter-

suchten Daten zu einem späteren Zeitpunkt ebenfalls negative Effekte komorbider Störungen ergeben.

Dass nach den hier vorgestellten Daten die Therapieeffekte auch für komorbide Patienten scheinbar anhaltend sind, ist ein ermutigendes Ergebnis, welches indes mit Vorsicht zu betrachten ist, da in die Katamneseerhebung jeweils deutlich geringere Stichprobengrößen (nur ca. 65% der zum Therapieende vorliegenden Daten) eingehen. Die Gruppe von Patienten, welche eine reliable Verschlechterung aufweisen, liegt mit 4% bis 8% in der Größenordnung dessen, was in der Psychotherapieforschung häufig gefunden wird (Clarkin & Levy, 2004) und scheinen nicht in Zusammenhang mit den Komorbiditätsgruppen zu stehen.

Bei Vorhersage der Anzahl durchgeführter Sitzungen resultiert wiederum ein deutlich geringerer aufgeklärter Varianzanteil als bei der Vorhersage des Ergebnisses. Auch hier besitzt die anfängliche Symptombelastung im BSI vergleichsweise große Erklärungskraft (s. a. Anderson & Lambert, 2001), jedoch resultiert zusätzlich Komorbidität als signifikanter Prädiktor. Je mehr Diagnosen vorliegen, desto mehr Therapiesitzungen werden in Anspruch genommen, mit der im Durchschnitt höchsten Anzahl bei komorbiden Persönlichkeitsstörungen, gefolgt von Komorbidität aus Angst- und affektiven Störungen. Hinweise auf eine Abhängigkeit der Therapiedauer von unterschiedlichen Symptommustern finden sich auch in der Literatur (etwa bei Howard et al., 1986; Kopta et al., 1994; Korobkin, Heron, & Ramirez, 1998; Maling, Gurtman, et al., 1995) – insbesondere auch die in dieser Arbeit identifizierten längeren Therapien bei Persönlichkeitsstörungen. Der hier gefundene Unterschied lässt sich nicht erklären durch die bei Komorbidität erhöhte Eingangsbelastung oder durch divergierende Anteile abgebrochener Therapien in den Komorbiditätsgruppen. Bemerkenswert ist weiterhin, dass die in diese Analysen eingehenden Patienten alle unter Bedingungen der IG des TK-Modellvorhabens behandelt wurden, sprich, dass hier bei gegebener Belastung ohne Einleitung eines Gutachterverfahrens von vornherein das Sitzungskontingent einer Langzeittherapie bewilligt wurde (entspricht in der Verhaltenstherapie 45 Sitzungen). Es stand somit allen Patienten die gleiche Sitzungsanzahl zur Verfügung, die in Abhängigkeit von der Anzahl an Diagnosen und der Art der Komorbidität jedoch mehr oder weniger ausgeschöpft wurde. Dies spricht für eine differenzielle Anpassung der hier durchgeführten ambulanten Psychotherapien in Abhängigkeit nicht nur von der durch den Patienten selbst berichteten und mit psychometrischen Instrumenten

erfassten Eingangsbelastung, sondern auch von der kategorialen Eingangsdiagnostik (Fremdbeurteilung durch Therapeuten). Patienten mit nur einer Diagnose scheinen deutlich schneller den gewünschten Therapieerfolg zu erreichen und, wie den Katamnesedaten zu entnehmen ist, auch zu halten, so dass diese im Mittel deutlich weniger Sitzungen in Anspruch nehmen (die durchschnittliche Sitzungsanzahl liegt zwischen der einer verhaltenstherapeutischen Kurz- und Langzeittherapie). Unter den Bedingungen des Projektes scheint es so gewesen zu sein, dass die Therapeuten ihr Handeln und insbesondere ihre therapeutischen Ressourcen auf die Bedürfnisse der Patienten einstellen konnten, so dass diejenigen, die mehr Therapie benötigten diese auch bekamen, wenngleich dieser Mehraufwand nicht dazu führte, dass die zu Anfang bestehenden Belastungsunterschiede zwischen komorbiden und monomorbiden Patienten zum Therapieende ausgeglichen waren. In diesem Zusammenhang muss mit Bedacht werden, dass Therapeuten der IG regelmäßige Rückmeldungen über den Symptomverlauf ihrer Patienten erhalten hatten und sich in ihrer Entscheidung über eine Therapiebeendigung auch daran orientiert haben. Studien zur Wirksamkeit von Rückmeldungen zeigen nicht nur einen im allgemeinen positiven Effekt, sondern im spezifischen, dass Rückmeldungen Einfluss auf das Therapeutenverhalten ausüben, indem diese als kritisch benannte Fälle länger in der Therapie halten und bei sich gut entwickelnden Patienten die Therapie früher beenden (Lambert et al., 2003, 2001).

Zusammen genommen profitieren komorbide Patienten im TK-Modell trotz größerer Ausgangsbelastung in ähnlichem Ausmaß wie Patienten mit nur einer Diagnose, weisen aber bezüglich des absoluten Wertes ein schlechteres Therapieergebnis auf. Sie verbleiben länger in der Therapie, was jedoch nicht zur Folge hat, dass die anfangs bestehenden Belastungsunterschiede zwischen den Gruppen ausgeglichen werden können. Die Tatsache, dass die Variable Komorbidität in den Regressionsanalysen zur Vorhersage des Therapieergebnisses nur marginal signifikant wird spricht nicht zwangsläufig dafür, dass sie keine therapeutische Relevanz hat (vgl. Stiles, Honos-Webb, & Surko, 1998). Sofern der Therapeut diese mögliche Komplikation wahrnimmt und darauf in angemessener Form (etwa durch eine Anpassung der Sitzungszahl) reagiert, reduziert sich ihr Einfluss auf das Ergebnis.

Es resultieren nicht für alle Ergebnisvariablen und im Kontext der anfänglichen Symptombelastung auch nur geringe Zusammenhänge zwischen der Therapiedauer und dem Ergebnis, und zwar in der Form, dass längere Therapien mit einem eher schlechteren Ergebnis assoziiert sind. Insofern lässt sich aus diesen Ergebnissen keine Unterstützung für das Do-

sis-Wirkungs-Modell der Psychotherapie ableiten. Jedoch ist einschränkend anzumerken, dass die Operationalisierung der „Dosis“ hier lediglich über die Anzahl an Sitzungen erfolgte. Es gibt Hinweise darauf, dass der zusätzliche Einbezug der Gesamtbehandlungsdauer beziehungsweise der Behandlungsdichte zu divergierenden Ergebnissen führt (Reardon et al. 2002). Daneben wurden mögliche Interaktionseffekte zwischen der Eingangsbelastung und der Dauer hier nicht beachtet; denkbar ist, dass nur bestimmte Untergruppen von Patienten von einer größeren Sitzungsanzahl profitieren (vgl. Shapiro et al., 1994). Darüber hinaus wurde hier zur Prädiktion der Sitzungsanzahl die Eingangsbelastung lediglich über den globalen BSI-Score definiert. Ob der IIP als Prädiktor einen ähnlichen Stellenwert eingenommen hätte oder, wie in einer Untersuchung von Puschner, Kraft, & Bauer (2004), ohne Einfluss auf die Sitzungszahl geblieben wäre, lässt sich an dieser Stelle nicht sagen.

## **6.2 Kritische Betrachtung der Untersuchung**

In Ergänzung zu den bei den spezifischen Fragestellungen bereits aufgeführten Nachteilen geht es in den kommenden Abschnitten um die Auseinandersetzung mit allgemeinen Stärken und Schwächen der durchgeführten Analysen. Dabei wird ein Schwerpunkt gelegt auf problematische Aspekte des Designs der TK-Studie im Hinblick auf die hier untersuchten Fragestellungen sowie auf Schwierigkeiten, die im Zusammenhang mit der Zusammenfassung zu Diagnosegruppen zu diskutieren sind.

### **6.2.1 Probleme des Studiendesigns**

Die im Rahmen dieser Arbeit untersuchten Fragestellungen ergeben sich als Nebenfragestellungen einer groß angelegten Studie, in deren Interesse vorrangig andere Zielbereiche lagen als die Untersuchung diagnostischer Methoden und der Auswirkung von Mehrfachdiagnosen. Dementsprechend ist das Studiendesign nicht spezifisch auf die Hypothesen dieser Arbeit abgestimmt. Hinzu kommen Einschränkungen, die im Kontext naturalistisch erhobener Datensätze zwangsläufig gemacht werden müssen. Im Vergleich zu kontrollierten Studien (RCTs) fällt die externe Validität größer aus auf Kosten einer verringerten in-

ternen Validität. Für die vorliegende Untersuchung ist besonders nachteilig, dass keine Kontrolle darüber besteht, welche Diagnosemethode beim einzelnen Patienten tatsächlich zum Einsatz kam, so dass die hier getroffenen Aussagen lediglich auf der Gruppeneinteilung und dem laut Studiendesign gewünschten Verhalten des Therapeuten beruhen können. Daneben wurde im Rahmen der TK-Studie keine Auskunft darüber eingeholt, welche Therapiemethoden faktisch durchgeführt wurden, beziehungsweise in welchem Ausmaß die teilnehmenden Therapeuten sich adhärent zu dem Therapieverfahren verhielten welches sie angegeben hatten. Somit ist nicht auszuschließen, dass etwa frühe Veränderungsverläufe oder die Therapiedauer durch verfahrens- und therapiemethodenspezifische Aspekte beeinflusst sind. Da in diese Studie lediglich Patienten eingingen, deren Therapeuten einen verhaltenstherapeutischen Ausbildungshintergrund angegeben hatten, konnten keine Wechselwirkungen zwischen der Komorbidität und den Therapieverfahren auf das Ergebnis untersucht werden. Daneben wurden aufgrund des naturalistischen Designs der TK-Studie weitere möglicherweise erklärende Variablen nicht oder nur unzureichend erfasst; so ist es denkbar, dass etwa die Einnahme von Medikamenten, externe Lebensereignisse oder soziale Unterstützung einen Einfluss insbesondere auf den Verlauf und das Ergebnis ausüben, der hier nicht weiter untersucht werden konnte. Auch bezüglich des wenig ausgeprägten Zusammenhangs zwischen Komorbidität und Therapieergebnis könnten moderierende Effekte etwa der Therapiemotivation, Hoffnung auf Besserung, Erfahrung des Therapeuten oder früherer Behandlungen bestehen. Ein weiteres typisches Problem naturalistischer Datensätze stellen fehlende Werte dar, wobei für diese Studie vor allem das Fehlen von Diagnosedaten zu beklagen ist. Darüber hinaus konnten im Studienzeitraum nicht alle Therapien beendet werden, weshalb die Analysen zur Vorhersage des Therapieergebnisses und der Anzahl Sitzungen zwangsläufig auf einer selektierten Stichprobe beruhen. Für nur etwas über 50% der Patienten liegen Angaben zum Therapieabschluss (Ende und/oder Kattamnese) vor. Obwohl sich im Vergleich zwischen den zum Studienende abgeschlossenen und noch laufenden Behandlungen keine Unterschiede bezüglich der Eingangsbelastung und Komorbidität ergeben ist es möglich, dass gerade besonders komplexe und langandauernde Therapien systematisch aus den Analysen ausgeschlossen wurden. Bezüglich der Abbruchraten scheinen laut früheren Untersuchungen vor allem ethnische Herkunft, Bildungsstand und Einkommen relevant (Wierzbicki & Pekarik, 1993). Ob und inwieweit auch Komorbidität einen Einfluss auf die Art einer Therapiebeendigung (reguläre Beendigung versus Abbruch) ausübt, bleibt eine offene Frage. Darüber hinaus bleibt anzumerken,

dass die in der TK-Studie bestehende Clusterung der Patienten in Therapeuten in den Analysen nicht berücksichtigt wurde und damit keine Angabe zu möglicherweise bestehenden Therapeuteneffekten gemacht werden können.

### **6.2.2 Diagnosestellung und Zusammenfassung zu Diagnosegruppen**

Alle eingehenden Diagnosen wurden auf zweierlei Art zusammengefasst: zum einen in insgesamt 17 Diagnosegruppen, zum anderen in vier Komorbiditätsgruppen. Es lässt sich diskutieren, ob mit dieser Einteilung ein Gleichgewicht zwischen Spezifität und Allgemeingültigkeit getroffener Aussagen eingehalten werden kann. Für die Untersuchung der bivariaten und multivariaten Diagnosezusammenhänge erschien es wichtig, eine eher feine Einteilung der Diagnosegruppen zu wählen und dabei möglichst viele Diagnosegruppen einzubeziehen, um das Potenzial der Daten auszuschöpfen und zur berichteten Kleinteiligkeit in der Komorbiditätsforschung ein Gegengewicht zu setzen. Es sollten möglichst viele potenzielle Störungskombinationen einbezogen werden, um klinisch relevante Aussagen zu begünstigen. Allerdings hatte die Wahl dieser relativ vielen Diagnosegruppen auch oft niedrige Zellbesetzungen zur Folge; um umfassendere Aussagen treffen zu können wäre daher ein noch sehr viel größerer Datensatz notwendig gewesen. Alternativ hätten weitere Diagnosezusammenfassungen getroffen werden können. Angewendet auf einen epidemiologischen Datensatz findet sich etwa eine sehr viel gröbere Einteilung von insgesamt 87 Diagnosen in fünf Hauptstörungsgruppen (Wittchen & Jacobi, 2002). Bezüglich solcher Zusammenfassungen stellt sich jedoch immer die Frage des Kriteriums: sollte eine Einteilung in Störungsgruppen etwa orientiert am Diagnosesystem erfolgen, orientiert an der Epidemiologie oder weiteren Faktoren? Weiter gefasste Diagnosegruppen führen zwar leichter zu signifikanten Ergebnissen, jedoch zum Preis weniger spezifischer Aussagen. So hätte etwa die in dieser Arbeit gefundene Unterschiedlichkeit zwischen einfachen und rezidivierenden depressiven Störungen bezüglich ihrer Komorbiditätsstruktur nicht aufgedeckt werden können, wenn eine weitere Zusammenfassung zu einer Gruppe „unipolare affektive Störungen“ getroffen worden wäre.

Bezüglich der Komorbiditätsgruppen wurde versucht, aus der Literatur klinisch relevante Muster abzuleiten, die sich gleichzeitig in ihrer Größe nicht zu stark unterscheiden sollten, da sie als Faktor in weitere Analysen aufgenommen wurden. Die Kombination „mit Persönlichkeitsstörungen“ wurde als eine Komorbiditätsgruppe gewählt, da diese häufig als

besonders schwere Komplikation beschrieben wird (u. a. Diguier et al., 1993; Mulder, 2002; Perry et al., 1999). Eine weitere Komorbiditätsgruppe „Angst- und affektiven Störungen“ ist alleine aufgrund der Auftretenshäufigkeit der Einzeldiagnosen naheliegend, wobei es auch Hinweise auf gemeinsame zugrundeliegende Mechanismen, im Sinne einer erhöhten Internalisierungstendenz, gibt (Krueger, 1999; Krueger & Finger, 2001). Daneben wurden in einer relativ großen Gruppe alle bisher nicht abgedeckten Störungskombinationen zusammengefasst („andere Komorbidität“). Explorative Analysen zeigen, dass es in weiteren Studien sinnvoll sein könnte, einen genaueren Blick auf Kombinationen mit Somatoformen Störungen und Belastungsstörungen zu werfen. Daneben war es unter der gewählten Einteilung nicht möglich, differenzielle Effekte spezifischer Persönlichkeitsstörungen (u. a. Skodol et al., 1995) oder mehrfacher Diagnosen aus einer Störungsgruppe (etwa Mineka et al., 1998) zu untersuchen. Ungeachtet dessen sprechen die Ergebnisse dieser Arbeit für die klinische Bedeutsamkeit der großen Komorbiditätsgruppen „Komorbidität mit Persönlichkeitsstörung“ und „Komorbidität aus Angst- und affektiven Störungen“ in Bezug auf Belastung, Verlauf und Behandlungsdauer.

### **6.3 Allgemeine Diskussion und Ausblick**

In dieser Arbeit wurden verschiedene Aspekte der Diagnosestellung und psychischen Komorbidität im Rahmen einer großen naturalistischen Psychotherapiestichprobe untersucht. Das TK-Modellvorhaben, aus dem die Daten stammen, stellt die erste Studie dieser Art an ambulanten Psychotherapiepatienten in Deutschland dar und umfasst Patienten und Therapeuten aus drei kassenärztlichen Versorgungsregionen. Wie in der Routineversorgung üblich, gab es – abgesehen von den bewilligten Sitzungskontingenten – keine Vorgaben bezüglich der Behandlungsdauer und durch die breite Masse an Einschlussdiagnosen ist ein großes Spektrum an Problembereichen ambulanter Psychotherapiepatienten abgedeckt. Ein weiterer Vorteil dieses Datensatzes ist in der umfangreichen Erhebung psychometrischer Instrumente aus Patienten- und Therapeutesicht zu Beginn und Ende der Behandlung sowie für die IG auch im Verlauf zu sehen. Somit ist es möglich, darin in umfassendem Maße Ausprägung und Korrelate von Komorbidität in der ambulanten Psychotherapie in Deutschland zu untersuchen. Allerdings bringt der Datensatz bezüglich der hier

angewendeten Fragestellungen und Methoden auch einige Schwächen mit sich, so dass die Generalisierbarkeit der Ergebnisse eingeschränkt ist. Mögliche Folgeuntersuchungen sollten auf einer großen Datenbasis mit sauber dokumentierter Diagnostik beruhen. Großes Potenzial für nachfolgende Untersuchungen auch im Bereich der Komorbidität psychischer Störungen bieten die verstärkten Bemühungen zur Umsetzung einer patientenorientierten Forschung im Rahmen von Wissenschaftler-Praktiker-Netzwerken, die die Sammlung großer Datenmengen unter Routinebedingungen erlaubt. Um dabei zu aussagekräftigen Ergebnissen zu gelangen sollten hohe Ansprüche an die Qualität erhobener Daten gestellt werden, welche etwa durch Therapeuteschulungen verbessert werden kann.

Bezogen auf die angewendeten Methoden zur Diagnosestellung bleibt ein direkter Vergleich von Checklisten mit strukturierten Interviews und unstrukturierter Diagnostik noch aus, wenn auch gezeigt wurde, dass unter Verwendung der IDCL mehr Diagnosen pro Patient zu erwarten sind. Sicher können die im Rahmen von Checklisten erhobenen Diagnosen nicht die gleiche Sensitivität und Spezifität erreichen wie Diagnosen, die aus strukturierten Interviews gewonnen werden. Dennoch könnten sie eine für den klinischen Alltag praktikable und wenig zeitaufwändige Alternative zu unstrukturierten Befragungen darstellen. Dabei wären sie als Erinnerungs- und Orientierungshilfen zu verstehen, um bei der Fülle an Information zu Beginn einer Therapie nichts Wichtiges zu übersehen und alle relevanten Problembereiche zu erfassen sowie in der Therapieplanung zu berücksichtigen. Dabei sollte es nicht darum gehen, von der Vergabe komorbider Diagnosen inflationären, sondern genaueren Gebrauch zu machen.

In Rahmen der Verlaufsanalysen ergibt sich eine Präferenz für loglineare Verbesserungsverläufe, was den Annahmen des Dosis-Wirkungs-Modells (Howard et al., 1986) entspricht. Allerdings kann der im Modell von Howard et al. (1986) postulierte positive Zusammenhang zwischen der Sitzungsanzahl und dem Therapieergebnis hier nicht gefunden werden. Zusätzlich zur Identifikation latenter Veränderungsmuster im Rahmen des GMM-Ansatzes und der Herstellung eines Bezuges zum Therapieergebnis wurden latente Verlaufsklassen im Rahmen dieser Studie auch genutzt, um das Therapieergebnis im Kontext der Eingangsbelastung vorherzusagen. Daneben wurde in Ergänzung zu früheren Arbeiten Komorbidität als potenziell erklärende Variable für die Zugehörigkeit zu den Veränderungsmustern untersucht und außerdem konnten -zumindest für einen Teil der Stichprobe- Katamnese-daten einbezogen werden, wodurch auch vorsichtige Aussagen zur langfristigen

Stabilität früher Veränderungen möglich sind. Wichtig wäre es, ähnliche Analysen noch einmal mit in Deutschland sitzungsweise erhobenen Daten durchzuführen und weiter nach klassenspezifischen Prädiktoren zu suchen. Durch eine Verzahnung mit Methoden der Prozessforschung wäre es möglich zu prüfen, ob etwa die früh und schnell verbesserten Patienten besonders gut auf unspezifische Wirkfaktoren ansprechen oder ob hier möglicherweise Effekte extratherapeutischer Ereignisse oder sozialer Unterstützung wirken.

Im Hinblick auf die benötigte Therapielänge weisen die Ergebnisse dieser Untersuchung darauf hin, dass globale Aussagen für alle Patienten kaum zu treffen sind. Stattdessen erscheint es sinnvoller, auch hier der Patientenspezifität Rechnung zu tragen und Abschätzungen für spezifische Subgruppen von Patienten vorzunehmen. Diese können durch verschiedene Merkmale charakterisiert sein, wobei die in strukturierter Form erfassten Diagnosen und deren gemeinsames Auftreten gut nutzbar erscheinen. Die hier vorgestellten Ergebnisse verweisen auf eine längere Therapiedauer unter Komorbidität. Dies könnte dafür sprechen, dass bei Bearbeitung eines Symptombereichs nicht automatisch mit einer Verbesserung anderer Symptome zu rechnen ist, sondern eine größere Anzahl an Problemkomplexen in einer Psychotherapie mehr oder weniger nacheinander abgearbeitet werden muss. Weitere Forschung ist notwendig, um genauer zu verstehen, warum bestimmte Komorbiditätsgruppen eine größere Anzahl an Therapiesitzungen benötigen. Bezug nehmend auf das Phasenmodell (Howard et al., 1993) ist denkbar, dass bei einer Kombination aus Angst und affektiven Störungen die Behandlung spezifischer Symptome besonders aufwändig ist, sprich die Remediationsphase vergleichsweise viel Zeit in Anspruch nimmt. Bei Persönlichkeitsstörungen spielen tief verwurzelte Aspekte des interpersonalen Funktionierens eine große Rolle, weshalb bei dieser Komorbidität eine längere Rehabilitationsphase vorstellbar ist.

### *Fazit*

Die TK-Studie und damit auch die vorliegende Arbeit sind im Kontext einer naturalistischen Forschung und der Qualitätssicherung in der ambulanten Psychotherapie angesiedelt. Die umfassende Untersuchung von Komorbidität vor diesem Forschungshintergrund stellt einen vergleichsweise neuen Ansatz dar und die Ergebnisse dieser Arbeit zeigen, dass eine gemeinsame Betrachtung für beide Forschungsstränge profitabel ist. Durch regelmäßige

Erhebungen über den gesamten Therapieverlauf hinweg ermöglicht ein patientenorientierter Ansatz ein genaueres Verständnis auch für die längerfristigen Auswirkungen von Mehrfachdiagnosen. So wurde in dieser Arbeit gezeigt, dass Komorbidität zwar mit einer im Allgemeinen größeren Anfangsbelastung einhergeht, dass aber komorbide ebenso wie monomorbide Patienten im Verlauf sehr gut profitieren können. Jedoch ist in Abhängigkeit von der Art der Komorbidität mit flacheren Verbesserungskurven beziehungsweise mit einem erhöhten Therapieaufwand zu rechnen.

Durch das TK-Modellvorhaben und einige weitere Studien konnte gezeigt werden, dass Rückmeldungen sich positiv auf Behandlungseffekte auswirken, indem Erfolge und Misserfolge frühzeitig erkannt werden und darauf reagiert werden kann. Eine solche Anpassung an die spezifischen Bedürfnisse des einzelnen Patienten erscheint insbesondere bei komorbiden Patienten wichtig, da hier aufgrund der diagnostischen Komplexität individualisierte Fallformulierungen stattfinden müssen und ein rein manualbezogener Ansatz (entsprechend einer sogenannten „evidence-based practice“; Barkham & Mellor-Clark, 2003) nicht ausreichend ist. Ein Monitoring des Therapiefortschritts ist aber nicht nur notwendig um mögliche Verschlechterungen frühzeitig zu erkennen, sondern auch um eine Behandlung dann zu beenden, wenn der individuelle Genesungszeitpunkt erreicht ist, welcher wiederum bei komorbiden Patienten später liegt als bei solchen mit nur einer Diagnose. Dementsprechend erscheint es sinnvoll, neben der psychometrisch erfassten Eingangsbelastung und Angaben zu frühen Veränderungen auch die mit vergleichsweise geringem Aufwand erfassbare Information zur Komorbidität in Vorhersage- und Rückmeldesysteme aufzunehmen. Die vorliegenden Daten liefern Hinweise darauf, dass Therapeuten unter Routinebedingungen bereits gewisse Anpassungen vornehmen (indem Patienten mit bestimmten Komorbiditätsmustern mehr Therapiesitzungen erhalten), die allerdings durch spezifische Rückmeldungen noch unterstützt werden könnten. Im Rahmen rationaler Entscheidungssysteme, in welchen häufig das Konzept klinisch signifikanter Veränderung (Jacobson & Truax, 1991) zum Tragen kommt, ist unter Komorbidität mit einer längeren Therapiedauer bis zur Cutoff-Unterschreitung in Richtung keiner oder geringer Belastung zu rechnen und Warnsignale könnten entsprechend der bei Komorbidität zu erwartenden erhöhten Eingangsbelastung angepasst werden. Auch im Rahmen empirischer Entscheidungsregeln könnte Komorbidität eine wichtige Information in Ergänzung zur Eingangsbelastung dar-

stellen, da zumindest in einigen Komorbiditätsgruppen mit einem flacheren Verlauf zu rechnen ist.

Die Ergebnisse dieser Arbeit können ihren Teil dazu beitragen, ein besseres Verständnis für die Beobachtung zu entwickeln, dass Patienten unterschiedlich schnell und gut auf Psychotherapie ansprechen. Wenn im Rahmen eines patientenorientierten Ansatzes und im Hinblick auf Kosteneinsparungen im Gesundheitswesen psychotherapeutische und zeitliche Ressourcen verstärkt denjenigen Patienten zukommen sollen, die sie besonders benötigen, sollte Komorbidität als Indikator dafür herangezogen werden. Auf Seiten der Kostenträger könnte somit eine Implikation darin liegen, die bewilligten Sitzungskontingente individueller auf die spezifischen Diagnosekombinationen und damit Bedürfnisse der Patienten anzupassen. Eine damit einhergehende Schlussfolgerung für die klinische Praxis ist, dass der Verbesserungsverlauf individuell und engmaschig beobachtet werden sollte, um jeden einzelnen Patienten so lange in der Therapie zu halten, wie es für seine Problemstellung notwendig ist. Dementsprechend sollte auch auf Ebene der Psychotherapieausbildung Wissen zu unterschiedlichen Verbesserungsverläufen und Therapiedauern vermittelt werden, sowie die praktische Kompetenz, flexible Anpassungen in der Behandlungsplanung und -durchführung vorzunehmen. Es muss jedoch darauf hingewiesen werden, dass das Vorliegen von Komorbidität nur eines von vielen möglichen Anzeichen für die intensivere Nutzung psychotherapeutischer Angebote darstellt. Potenzial für zukünftige Forschungsfragen ist vor allem hinsichtlich der Interaktion spezifischer Komorbiditätsmuster mit Merkmalen des Therapeuten, des Settings oder der Therapieform zu sehen. Daneben sind Replikationen und weitere Untersuchungen insbesondere bezüglich der Verlaufsfragestellungen notwendig und von großem Interesse für Anstrengungen zur Qualitätssicherung und besseren Ressourcensteuerung in der Psychotherapie.

## 7 LITERATURVERZEICHNIS

Abramowitz, J. S., Franklin, M. E., Street, G. P., Kozak, M. J., & Foa, E. B. (2000). Effects of comorbid depression on response to treatment for obsessive-compulsive disorder. *Behavior Therapy, 31*, 517-528.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd edition) - DSM-III-R*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th edition) - DSM IV*. Washington, DC: American Psychiatric Association.

Anderson, E. M., & Lambert, M. J. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology, 57*(7), 875-888.

Andreoli, A., Frances, A., Gex-Fabry, M., Aapro, N., Gerin, P., & Dazord, A. (1993). Crisis intervention in depressed patients with and without DSM-III-R personality disorders. *The Journal of Nervous and Mental Disease, 181*(12), 732-737.

Andrews, G., Slade, T., & Issakidis, C. (2002). Deconstructing current comorbidity: data from the Australian National Survey of Mental Health and Well-Being. *British Journal of Psychiatry, 181*, 306-314.

Angold, A., Costello, C. G., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, 40*(1), 57-87.

Angst, J. (1998). The emerging epidemiology of hypomania and bipolar II disorder. *Journal of Affective Disorders, 50*, 143-151.

Angst, J., Gamma, A., Endrass, J., Goodwin, R., Ajdacic, V., Eich, D., & Rössler, W. (2004). Obsessive-compulsive severity spectrum in the community: prevalence, comorbidity, and course. *European Archives of Psychiatry and Clinical Neuroscience, 254*, 156-164.

Audin, K., Mellor-Clark, J., Barkham, M., Margison, F., McGrath, G., Lewis, S., Cann, L., et al. (2001). Practice research networks for effective psychological therapies. *Journal of Mental Health, 10*(3), 241-251.

Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77*(2), 203-211.

Barkham, M., Connell, J., Miles, J. N. V., Evans, C., Stiles, W. B., Margison, F., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology, 74*(1), 160-167.

Barkham, M., & Mellor-Clark, J. (2003). Bridging evidence-based practice and practice-based evidence: developing a rigorous and relevant knowledge for the psychological therapies. *Clinical Psychology & Psychotherapy, 10*(6), 319-327.

Barkham, M., Rees, A., Stiles, W. B., Hardy, G. E., & Shapiro, D. A. (2002). Dose-effect relations for psychotherapy of mild depression: A quasi-experimental comparison of 2, 8, and 16 sessions. *Psychotherapy Research, 12*(4), 463-474.

- Barkham, M., Rees, A., Stiles, W. B., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996). Dose-effect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology, 64*(5), 927-35.
- Bass, C., & Murphy, M. (1995). Somatoform and personality disorders: syndromal comorbidity and overlapping developmental pathways. *Journal of Psychosomatic Research, 39*(4), 403-427.
- Bassler, M., Potratz, B., & Krauthauser, H. (1995). Der „Helping Alliance Questionnaire“ (HAQ) von Luborsky. *Psychotherapeut, 40*, 23-32.
- Bauer, D. J., & Curran, P. J. (2003). Overextraction of latent trajectory classes: Much ado about nothing? Reply to Rindskopf (2003), Muthén (2003) and Cudeck and Henly (2003). *Psychological Methods, 8*(3), 384-393.
- Becker, E., & Margraf, J. (2007). *Generalisierte Angststörung: Ein Therapieprogramm*. Weinheim: Beltz.
- Bellach, B. M., Knopf, H., & Thefeld, W. (1998). Der Bundes-Gesundheitssurvey 1997/98. *Gesundheitswesen, 60*, 59-68.
- Bengel, J., & Hubert, S. (2010). *Anpassungsstörung und Akute Belastungsstörung*. Göttingen: Hogrefe.
- Berking, M., Orth, U., & Lutz, W. (2006). Wie effektiv sind systematische Rückmeldungen des Therapieverlaufs an den Therapeuten? *Zeitschrift für Klinische Psychologie und Psychotherapie, 35*(1), 21-29.
- Blashfield, R. K. (1990). Comorbidity and classification. In J. D. Maser & Cloninger (Hrsg.), *Comorbidity of mood and anxiety disorders* (S. 61-82). Washington D.C.: American Psychiatric Press.
- Bollen, K. A. & Curran, P. J. (2006). *Latent curve models. A structural equation perspective*. Hoboken: Wiley & Sons.
- Borkovec, T. D., Abel, J. L., & Newman, H. (1995). Effects of Psychotherapy on Comorbid Conditions in Generalized Anxiety Disorder. *Journal of Consulting and Clinical Psychology, 63*(3), 497-483.
- Borkovec, T. D., Echemendia, R. J., Ragusea, S. A., & Ruiz, M. (2001). The Pennsylvania Practice Research Network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology: Science and Practice, 8*, 155-167.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.). Berlin: Springer.
- Brawman-Mintzer, O., Lydiard, R. B., Emmanuel, N., Payeur, R., Johnson, M., Roberts, J., Jarrell, M. P., et al. (1993). Psychiatric comorbidity in patients with Generalized Anxiety Disorder. *American Journal of Psychiatry, 150*(8), 1216-1218.
- Brieger, P., & Marneros, A. (2000). Komorbidität bei psychiatrischen Erkrankungen. *Nervenarzt, 71*, 525-534.
- Bronisch, T., Hiller, W., Mombour, W., & Zaudig, M. (1995). *Manual IDCL-P Internatio-*

*nale Diagnosen Checkliste für Persönlichkeitsstörungen nach ICD-10 und DSM-IV*. Bern: Huber.

Bronisch, T., & Mombour, W. (1994). Comparison of a Diagnostic Checklist with a Structured Interview for the Assessment of DSM-III-R and ICD-10 Personality Disorders. *Psychopathology*, 27(6), 312-320.

Bronisch, T., & Wittchen, H.-U. (1994). Suicidal ideation and suicide attempts: comorbidity with depression, anxiety disorders, and substance abuse disorder. *European Archives of Psychiatry and Clinical Neuroscience*, 244, 93-98.

Brown, M. B., & Benedetti, J. K. (1977). On the mean and variance of the tetrachoric correlation coefficient. *Psychometrika*, 42(3), 347-355.

Brown, T. A., Antony, M. M., & Barlow, D. H. (1995). Diagnostic comorbidity in panic disorder: effect on treatment outcome and course of comorbid diagnoses following treatment. *Journal of Consulting and Clinical Psychology*, 63(2), 408-418.

Brown, T. A., & Barlow, D. H. (2005). Dimensional Versus Categorical Classification of Mental Disorders in the Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders and Beyond: Comment on the Special Section. *Journal of Abnormal Psychology*, 114(4), 551-556.

Brunello, N., den Boer, J. A., Judd, L. L., Kasper, S., Kelsey, J. E., Lader, M., Lecrubier, Y., et al. (2000). Social phobia: diagnosis and epidemiology, neurobiology and pharmacology, comorbidity and treatment. *Journal of Affective Disorders*, 60, 61-74.

Bullinger, M., & Kirchberger, I. (1998). *SF-36. Fragebogen zum Gesundheitszustand. Handanweisung*. Göttingen: Hogrefe.

Burke, J. D., Wittchen, H.-U., Regier, D. A., & Sartorius, N. (1990). Extracting information from diagnostic interviews on co-occurrence of symptoms of anxiety and depression. In J. D. Maser & C. R. Cloninger (Hrsg.), *Comorbidity of mood and anxiety disorders* (S. 649-667). Washington D.C.: American Psychiatric Press.

Chambless, D. L., Renneberg, B., Gracely, E. J., Goldstein, A. J., & Fydrich, T. (2000). Axis I and II comorbidity in agoraphobia: prediction of psychotherapy outcome in a clinical setting. *Psychotherapy Research*, 10(3), 279-295.

Chen, Y., & Dilsaver, S. (1995). Comorbidity of panic disorder in bipolar illness: evidence from the Epidemiologic Catchment Area Survey. *American Journal of Psychiatry*, 152(2), 280-282.

Choy, Y., Fyer, A. J., & Goodwin, R. D. (2007). Specific phobia and comorbid depression: a closer look at the National Comorbidity Survey data. *Comprehensive Psychiatry*, 48(2), 132-136.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, 100, 316-336.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology*, 46, 121-153.

Clarkin, J. F., & Kendall, P. C. (1992). Comorbidity and Treatment Planning: Summary

and Future Directions. *Journal of Consulting and Clinical Psychology*, 60, 904-908.

Clarkin, J. F., & Levy, K. N. (2004). The influence of client variables on psychotherapy. In M. Lambert (Hrsg.), *Handbook of psychotherapy and behavior change* (S. 194-227). New York, NY: Wiley & Sons.

Clogg, C. C. (1995). Latent Class Models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Hrsg.), *Handbook of statistical modeling for the social and behavioral sciences* (S. 311-359). New York: Plenum Press.

Cloninger, C. R., Martin, R. L., Guze, S. B., & Clayton, P. J. (1990). The empirical structure of psychiatric comorbidity and its theoretical significance. In J. D. Maser & C. R. Cloninger (Hrsg.), *Comorbidity of Anxiety and Mood Disorders* (S. 439-462). Washington, D. C.: American Psychiatric Association.

Colder, C. R., Campbell, R. T., Ruel, E., Richardson, J. L. & Flay, B. R. (2002). A finite mixture model of growth trajectories of adolescent alcohol use: Predictors and consequences. *Journal of Consulting and Clinical Psychology*, 70(4), 976-985.

Cooper, L. D., & Balsis, S. (2009). When less is more: how fewer diagnostic criteria can indicate more personality disorder pathology. *Psychological Assessment*, 21, 285-293.

Costello, C. G. (1993). Conceptual and methodological issues in the sciences of psychopathology. In C. G. Costello (Hrsg.), *Basic Issues in Psychopathology*. New York: Guilford Press.

Cuijpers, P., van Lier, P. A. C., van Straten, A., & Donker, M. (2005). Examining differential effects of psychological treatment of depressive disorder: an application of trajectory analyses. *Journal of Affective Disorders*, 89, 137-146.

Curran, P. J. & Hussong, A. M. (2003). The Use of Latent Trajectory Models in Psychopathology Research. *Journal of Abnormal Psychology*, 112(4), 526-544.

Davila, J., Hammen, C., Burge, D., Paley, B., & Daley, S. E. (1995). Poor Interpersonal Problem Solving as a Mechanism of Stress Generation in Depression Among Adolescent Women. *Journal of Abnormal Psychology*, 104(4), 592-600.

Dayton, C. M., & Macready, G. B. (2002). Use of categorical and continuous covariates in latent class analysis. In J. A. Hagenaars & A. L. McCutcheon (Hrsg.), *Applied Latent Class Analysis* (S. 213-233). Cambridge: University Press.

de Waal, M. W. M., Arnold, I. A., Eekhof, J. A., & van Hemert, A. M. (2004). Somatoform disorders in general practice. *British Journal of Psychiatry*, 184, 470-476.

Deering, C. G., Glover, S. G., Ready, D., Eddleman, H. C., & Alarcon, R. D. (1996). Unique patterns of comorbidity in posttraumatic stress disorder from different sources of trauma. *Comprehensive Psychiatry*, 37(5), 336-346.

Depla, M. F. I. A., Have, M. L., Balkom, A. J. L. M., & Graaf, R. (2007). Specific fears and phobias in the general population: Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology*, 43(3), 200-208.

Derogatis, L. R. (1992). *SCL-90-R: Administration, scoring and procedures manual - II*. Baltimore, MD: Clinical Psychometric Research.

- Derogatis, L. R. (1993). *BSI. Brief Symptom Inventory. Administration, Scoring, and Procedures Manual*. Minneapolis, MN: National Computer Systems.
- Diaconu, G., & Turecki, G. (2007). Panic disorder and suicidality: Is comorbidity with depression the key? *Journal of Affective Disorders, 104*, 203-209.
- Diguer, L., Barber, J. P., & Luborsky, L. (1993). Three concomitants: personality disorders, psychiatric severity, and outcome of dynamic psychotherapy of major depression. *American Journal of Psychiatry, 150*(8), 1246-1248.
- Dreessen, L., Arntz, A., Luttels, C., & Sallaerts, S. (1994). Personality disorders do not influence the results of cognitive behavior therapies for anxiety disorders. *Comprehensive Psychiatry, 35*(4), 265-274.
- Dreessen, L., Hoekstra, R., & Arntz, A. (1998). Personality disorders do not influence the results of cognitive and behavior therapy for obsessive compulsive disorder. *Journal of Anxiety Disorders, 11*(5), 503-521.
- Duncan, T. E. & Duncan, S. E. (2004). An introduction to latent growth curve modeling. *Behavior Therapy, 35*(2), 333-363.
- Durham, R. C., Allan, T., & Hackett, C. A. (1997). On predicting improvement and relapse in generalized anxiety disorder following psychotherapy. *The British Journal of Clinical Psychology / the British Psychological Society, 36 ( Pt 1)*, 101-119.
- Ehlers, A., & Margraf, J. (2001). *Fragebogen zu körperbezogenen Ängsten, Kognitionen und Vermeidung (AKV)* (Bd. 2., überarbeitete und neunormierte Auflage). Weinheim: Beltz.
- Eldredge, K. L., Agras, W. S., Arnow, B., Telch, C. F., Bell, S., Castonguay, L., & Marnell, M. (1997). The effects of extending cognitive-behavioral therapy for binge eating disorder among initial treatment nonresponders. *International Journal of Eating Disorders, 21*(4), 347-352.
- Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33*, 766-771.
- Erwin, B. A., Heimberg, R. G., Juster, H., & Mindlin, M. (2002). Comorbid anxiety and mood disorders among persons with social anxiety disorder. *Behaviour Research and Therapy, 40*, 19-35.
- Essau, C. A., Conradt, J., & Petermann, F. (2000). Frequency, comorbidity, and psychosocial impairment of Specific Phobia in adolescents. *Journal of Clinical Child Psychology, 29*(2), 221-231.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.
- Feinstein, A. R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases, 23*, 455-468.
- Ferrada-Noli, M., Asberg, M., Ormstad, K., Lundin, T., & Sundbom, E. (1998). Suicidal behavior after severe trauma. Part 1: PTSD diagnoses, psychiatric comorbidity, and as-

assessments of suicidal behavior. *Journal of Traumatic Stress, 11*(1), 103-112.

Feske, U., Frank, E., Mallinger, A. G., Houck, P. R., Fagiolini, A., Shear, M. K., Grochocinski, V. J., et al. (2000). Anxiety as a correlate of response to the acute treatment of Bipolar I disorder. *American Journal of Psychiatry, 157*(6), 956-962.

Fiedler, P. (1998). *Persönlichkeitsstörungen*. Weinheim: Beltz.

First, M. D., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1994). *Structured Clinical Interview for axis I DSM-IV disorders: Patient edition (SCID-I/P, version 2.0)*. New York: Biometrics Research Department, State Psychiatric Institute.

Frances, A., Widiger, T. A., & Fyer, M. R. (1990). The influence of classification methods on comorbidity. *Comorbidity of Anxiety and Mood Disorders* (S. 41-60). Washington, D.C.: American Psychiatric Association.

Franke, G. H. (1995). *SCL-90-R. Die Symptom-Checkliste von Derogatis*. Göttingen: Beltz Test GmbH.

Franke, G. H. (2000). *Brief Symptom Inventory von L. R. Derogatis (Kurzform der SCL-90-R) - Deutsche Version*. Göttingen: Beltz Test GmbH.

Freeman, M. P., Freeman, S. A., & McElroy, S. L. (2002). The comorbidity of bipolar and anxiety disorders: prevalence, psychobiology, and treatment issues. *Journal of Affective Disorders, 68*(1), 1-23.

Fydreich, T., Nagel, A., Lutz, W., & Richter, R. (2003). Qualitätsmonitoring in der ambulanten Psychotherapie: Modellprojekt der Techniker Krankenkasse (TK). *Verhaltenstherapie, 13*, 291-295.

Galbaud du Fort, G., Newman, S. C., & Bland, R. C. (1993). Psychiatric Comorbidity and Treatment Seeking: Sources of Selection Bias in the Study of Clinical Populations. *The Journal of Nervous and Mental Disease, 181*, 467-474.

Geiser, C. (2010). *Datenanalyse mit Mplus*. Wiesbaden: Verlag für Sozialwissenschaften.

Gershuny, B. S., Baer, L., Jenike, M. A., Minichiello, W. E., & Wilhelm, S. (2002). Comorbid Posttraumatic Stress Disorder: Impact on Treatment Outcome for Obsessive-Compulsive Disorder. *Am J Psychiatry, 159*(5), 852-854.

Godard, N. T., Perdereau, F., Rein, Z., Berthoz, S., Wallier, J., Jeammet, P., & Flament, M. F. (2007). Comorbidity studies of eating disorders and mood disorders. Critical review of the literature. *Journal of Affective Disorders, 97*, 37-49.

Grawe, K. (2004). *Psychoneurotherapie*. Göttingen: Hogrefe.

Haas, E., Hill, R. D., Lambert, M. J., & Morell, B. (2002). Do early responders to psychotherapy maintain treatment gains? *Journal of Clinical Psychology, 58*(9), 1157-1172.

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and a algorithmus for identifying patients at risk for treatment failure. *Journal of Clinical Psychology, 61*(2), 155-163.

Hansen, N. B., & Lambert, M. J. (2003). An Evaluation of the Dose-Response Relationship in Naturalistic Treatment Settings Using Survival Analysis. *Mental Health Services*

*Research*, 5(1), 1-12.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implication for treatment delivery systems. *Clinical Psychology: Science and Practice*, 9, 329-343.

Hardy, G. E., Barkham, M., Shapiro, D. A., Stiles, W. B., Rees, A., & Reynolds, S. (1995). Impact of Cluster C Personality Disorders on Outcomes of Contrasting Brief Psychotherapies for Depression. *Journal of Consulting and Clinical Psychology*, 63(6), 997-1004.

Hautzinger, M., Bailer, M., Worall, H., & Keller, F. (1995). *Beck-Depressions-Inventar (BDI)*. Testhandbuch. Bern: Hans Huber.

Hayes, A. M., Feldman, G. C., Beevers, C. G., Laurenceau, J.-P., Cardaciotto, L., & Lewis-Smith, J. (2007). Discontinuities and Cognitive Changes in an Exposure-Based Cognitive Therapy for Depression. *Journal of Consulting and Clinical Psychology*, 75(3), 409-421.

Helou, A., Schwartz, F. W., & Ollenschläger, G. (2002). Qualitätsmanagement und Qualitätssicherung in Deutschland. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 45, 205-214.

Hiller, W., Rief, W., & Fichter, M. M. (1997). How disabled are patients with somatoform disorders? *General Hospital Psychiatry*, 19, 432-438.

Hiller, W., Zaudig, M., & Mombour, W. (2004). *IDCL - International Diagnostic Checklists for ICD-10 and DSM-IV*. Bern: Hans Huber.

Hiller, W., Zaudig, M., Mombour, W., & Bronisch, T. (1993). Routine psychiatric examinations guided by ICD-10 Diagnostic Checklists (International Diagnostic Checklists). *European Archives of Psychiatry and Clinical Neuroscience*, 242, 218-223.

Hiller, W., Dichtl, G., Hecht, H., Hundt, W., & Zerssen, D. (1993). An empirical comparison of diagnoses and reliabilities in ICD-10 and DSM-III-R. *European Archives of Psychiatry and Clinical Neuroscience*, 242(4), 209-217.

Hirschfeld, R. M. A. (1999). Personality disorders and depression: comorbidity. *Depression and anxiety*, 10, 142-146.

Horowitz, L. M., Strauss, B., & Kordy, H. (2000). *Inventar zur Erfassung Interpersonaler Probleme (IIP-D)*. Göttingen: Beltz Test GmbH.

Howard, K. I., Kopta, M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist*, 41, 159-164.

Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: causal mediation of change. *Journal of Consulting and Clinical Psychology*, 61(4), 678-685.

Howard, K. I., Moras, K., Brill, B. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of Psychotherapy. Efficacy, Effectiveness and Patient Progress. *American Psychologist*, 51, 1059-1064.

Hudson, J. I., Hiripi, E., Pope, H. G., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological Psy-*

*chiatry*, 61, 348-358.

Ilardi, S. S., Craighead, W. E., & Evans, D. D. (1997). Modeling Relapse in Unipolar Depression: The Effects of Dysfunctional Cognitions and Personality Disorders. *Journal of Consulting and Clinical Psychology*, 65(3), 381-391.

Jacobi, F., Klose, M., & Wittchen, H.-U. (2004). Psychische Störungen in der deutschen Allgemeinbevölkerung: Inanspruchnahme von Gesundheitsleistungen und Ausfalltage. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 47(8), 736-744.

Jacobi, F., Wittchen, H.-U., Höltling, C., Höfler, M., Pfister, H., Müller, N., & Lieb, R. (2004). Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychological Medicine*, 34, 1-15.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12-9.

Joormann, J., Kosfelder, J., & Schulte, D. (2005). The Impact of Comorbidity of Depression on the Course of Anxiety Treatments. *Cognitive Therapy and Research*, 29(5), 569-591.

Joyce, A. S., Ogrodniczuk, J., Piper, W. E., & McCallum, M. (2002). A test of the phase model of psychotherapy change. *Canadian Journal of Psychiatry*, 47(8), 759-766.

Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Practice and Research*, 5, 132-151.

Kaplan, M. H., & Feinstein, A. R. (1974). The importance of classifying initial comorbidity in evaluating the outcome of diabetes mellitus. *Journal of Chronic Diseases*, 27(7-8), 387-404.

Kastenholz, H., & Both, B. (2002). Qualitätssicherung in der medizinischen Versorgung aus Sicht des Bundesministeriums für Gesundheit. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 45(3), 215-222.

Kaye, W. H., Bulik, C. M., Thornton, L., Barbarich, N., & Masters, K. (2004). Comorbidity of anxiety disorders with anorexia and bulimia nervosa. *American Journal of Psychiatry*, 161(12), 2215-2221.

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*; *American Psychologist*, 63(3), 146-159.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., Rush, A. J., et al. (2003). The epidemiology of major depressive disorder. Results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association*, 289(23), 3095-3105.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2004). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 617-627.

- Kessler, R. C., Chiu, W. T., Jin, R., Ruscio, A. M., Shear, M. K., & Walters, E. E. (2006). The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *63*, 415-424.
- Kessler, R. C., DuPont, R. L., Berglund, P., & Wittchen, H.-U. (1999). Impairment in pure and comorbid General Anxiety Disorder and Major Depression at 12 months in two national surveys. *American Journal of Psychiatry*, *156*(12), 1915-1923.
- Kessler, R. C., Keller, M. B., & Wittchen, H.-U. (2001). The epidemiology of Generalized Anxiety Disorder. *The Psychiatric Clinics of North America*, *24*(1), 19-39.
- Kessler, R. C., Nelson, C. B., McGonagle, K. A., Liu, J., Swartz, M., & Blazer, D. G. (1996). Comorbidity of DSM-III-R Major Depressive Disorder in the general population: Results from the US National Comorbidity Study. *British Journal of Psychiatry*, *168*(supp. 30), 17-30.
- Kessler, R. C., Stang, P., Wittchen, H.-U., Stein, M., & Walters, E. E. (1999). Lifetime comorbidities between social phobia and mood disorders in the US National Comorbidity Survey. *Psychological Medicine*, *29*, 555-567.
- Kessler, R. C., Stein, M. B., & Berglund, P. (1998). Social phobia subtypes in the National Comorbidity Survey. *American Journal of Psychiatry*, *155*, 613-619.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*, *62*(6), 593-602.
- Klein, D. N., & Riso, L. P. (1993). Psychiatric Disorders: Problems of Boundaries and Comorbidity. In C. G. Costello (Hrsg.), *Basic Issues in Psychopathology* (S. 19-66). New York: Guilford Press.
- Klepsch, R., Zaworka, W., Hand, I., Lünenschloss, K., & Jauering, G. (1993). *Hamburger Zwangsinventar - Kurzform (HZI-K). Manual*. Göttingen: Beltz.
- Klerman, G. L. (1990). Approaches to the phenomena of comorbidity. In J. D. Maser & C. R. Cloninger (Hrsg.), *Comorbidity of mood and anxiety disorders* (S. 13-37). Washington, D.C.: American Psychiatric Press.
- Kopta, S. M. (2003). The dose-effect relationship in psychotherapy: A defining achievement for Dr. Kenneth Howard. *Journal of Clinical Psychology*, *59*, 727-733.
- Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, *62*, 1009-1016.
- Korobkin, S. B., Herron, W. G., & Ramirez, S. M. (1998). Severity of symptoms of depression and anxiety as predictors of duration of psychotherapy. *Psychological Reports*, *82*(2), 427-433.
- Kraft, D., Koellner, V., & Lutz, W. (2008). Patientenorientierte Psychotherapieforschung in der stationären Routineversorgung. 1. Forschungswerkstatt Stationäre Psychotherapie.
- Krause, M. S., Howard, K. I., & Lutz, W. (1998). Exploring individual change. *Journal of Consulting and Clinical Psychology*, *66*(5), 838-45.

- Krueger, R. F. (1999). The structure of common mental disorders. *Archives of General Psychiatry*, *56*, 921-926.
- Krueger, R. F. (2002). Psychometric Perspectives on Comorbidity. In J. E. Helzer & J. J. Hudziak (Hrsg.), *Defining psychopathology in the 21st century* (S. 41). Washington, D. C.: American Psychiatric Publishing, Inc.
- Krueger, R. F., & Finger, M. S. (2001). Using item response theory to understand comorbidity among anxiety and unipolar mood disorders. *Psychological Assessment*, *13*, 140-151.
- Krueger, R. F., & Markon, K. E. (2006). Reinterpreting Comorbidity: A Model-Based Approach to Understanding and Classifying Psychopathology. *Annual review of clinical psychology*, *2*, 111-133.
- Kruger, S., Cooke, R. G., Hasey, G. M., Jorna, T., & Persad, E. (1995). Comorbidity of obsessive compulsive disorder in bipolar disorder. *Journal of Affective Disorders*, *34*, 117-120.
- Kuyken, W. (2004). Cognitive therapy outcome: the effects of hopelessness in a naturalistic outcome study. *Behaviour Research and Therapy*, *42*, 631-646.
- Lambert, M. J. (2001). Psychotherapy outcome and quality improvement: Introduction to the special section on patient-focused research. *Journal of Consulting and Clinical Psychology*, *69*, 159-172.
- Lambert, M. J. (2007). What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, *17*(1), 1-14.
- Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, *69*(2), 159-72.
- Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R. C., Shimokawa, K., et al. (2004). *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire)*. Orem, UT: American Professional Credentialing Services L.L.C.
- Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Hrsg.), *Handbook of psychotherapy and behavior change* (Bd. 5th, S. 139-193). New York, NY: Wiley & Sons.
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice*, *10*, 288-301.
- Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*, *11*, 49-68.
- Lambert, M. J., Whipple, J. M., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology and Psychotherapy*, *9*, 91-103.
- Leary, T. (1957). *Interpersonal diagnosis of personality*. New York: Ronald Press.

- Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*(3), 767-778.
- Lueger, R. J., Howard, K. I., Martinovich, Z., Lutz, W., Anderson, E. E., & Grissom, G. (2001). Assessing treatment progress of individual patients using expected treatment response models. *Journal of Consulting and Clinical Psychology*, *69*, 150-158.
- Lueger, R. J., Lutz, W., & Howard, K. I. (2000). The predicted and observed course of psychotherapy for anxiety and mood disorders. *The Journal of Nervous and Mental Disease*, *188*, 127-134.
- Lutz, Köck, K., & Böhnke, J. (2009a). Die Wirkung von Rückmeldesystemen aus ambulanten Settings: Das Modellvorhaben der Techniker Krankenkasse und Wege in die stationäre Praxis. *Praxis Klinische Verhaltensmedizin und Rehabilitation*, *22*(84), 118-125.
- Lutz, W. (2002). Patient-focused psychotherapy research and individual treatment progress as scientific groundwork for an empirical based clinical practice. *Psychotherapy Research*, *12*, 251-273.
- Lutz, W., & Bittermann, A. (2010). Wie, wann und warum verändern sich Menschen in der Psychotherapie? *Psychotherapie im Dialog*, *11*, 80-85.
- Lutz, W., Böhnke, J. R., Köck, K., & Bittermann, A. (2011). Diagnostik und psychometrische Verlaufsrückmeldungen im Rahmen eines Modellprojektes zur Qualitätssicherung in der ambulanten Psychotherapie. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *4*, 289-297.
- Lutz, W., Leach, C., Barkham, M., Lucock, M., Stiles, W. B., Evans, C., Noble, R., et al. (2005). Predicting change for individual psychotherapy clients based on their nearest neighbors. *Journal of Consulting and Clinical Psychology*, *73*(5), 904-913.
- Lutz, W., Lowry, J., Kopta, M., Einstein, A., & Howard, K. I. (2001). Prediction of dose-response relations based on patient characteristics. *Journal of Clinical Psychology*, *57*, 1-12.
- Lutz, W., Martinovich, Z., & Howard, K. I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology*, *67*, 571-577.
- Lutz, W., Rafaeli, E., Howard, K. I., & Martinovich, Z. (2002). Adaptive modeling of progress in outpatient psychotherapy. *Psychotherapy Research*, *12*(4), 427-443.
- Lutz, W., Saunders, S. M., Leon, S. C., Martinovich, Z., Kosfelder, J., Schulte, D., Grawe, K., et al. (2006). Empirically and clinically useful decision making in psychotherapy: Differential predictions with treatment response models. *Psychological Assessment*, *18*(2), 133-141.
- Lutz, W., & Stulz, N. (2008). Development and application of feedback-systems to support clinical decision making in outpatient psychotherapy. In G. R. Burthold (Hrsg.), *Psychology of Decision Making in Legal, Health Care and Science Settings* (S. 203-218). Hauppauge, NY: Nova Science Publishers.
- Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affec-*

*tive Disorders*, 118, 60-68.

Lutz, W., Stulz, N., Martinovich, Z., Leon, S., & Saunders, S. M. (2009). Methodological Background of Decision Rules and Feedback Tools for Outcomes Management in Psychotherapy. *Psychotherapy Research*, 19(4-5), 502-510.

Lutz, W., Stulz, N., Smart, D. W., & Lambert, M. J. (2007). Die Identifikation früher Veränderungsmuster in der ambulanten Psychotherapie. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 36(2), 93-104.

Lutz, W., & Tschitsaz, A. (2007). Plötzliche Gewinne und Verluste im Behandlungsverlauf von Angststörungen, depressiven und komorbiden Störungen [Sudden Gains and Losses in the treatment of patients with anxiety, depressive and comorbid disorders]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 36(4), 298-308.

Maj, M. (2005). Psychiatric Comorbidity: an artefact of current diagnostic systems? *British Journal of Psychiatry*, 186, 182-184.

Maling, M. S., Gurtman, M. B., & Howard, K. I. (1995). The response of interpersonal problems to varying doses of psychotherapy. *Psychotherapy Research*, 5(1), 63-75.

Margraf, J. (Hrsg.). (2003). *Lehrbuch der Verhaltenstherapie. Band 1*. Berlin: Springer.

Markowitz, J. C., Skodol, A. E., Petkova, E., Cheng, J., Sanislow, C. A., Grilo, C. M., Gunderson, J. G., et al. (2007). Longitudinal effects of personality disorders on psychosocial functioning of patients with major depressive disorder. *Journal of Clinical Psychiatry*, 68(2), 186-193.

Marshall, R. D., Olfson, M., Hellman, F., Blanco, C., Guardino, M., & Struening, E. L. (2001). Comorbidity, impairment, and suicidality in subthreshold PTSD. *American Journal of Psychiatry*, 158(9), 1467-1473.

Maser, J. D., & Cloninger, C. R. (1990). Comorbidity of Anxiety and Mood Disorders: Introduction and Overview. In J. D. Maser & C. R. Cloninger (Hrsg.), *Comorbidity of Anxiety and Mood Disorders* (S. 3-12). Washington, D. C.: American Psychiatric Association.

McCutcheon, A. L. (2002). Basic concepts and procedures in single-and multiple-group latent class analysis. In J. A. Hagenaars & A. L. McCutcheon (Hrsg.), *Applied Latent Class Analysis* (S. 56-85). Cambridge: University Press.

McFarlane, A. C., & Papay, P. (1992). Multiple diagnoses in posttraumatic stress disorder in the victims of a natural disaster. *Journal of Nervous and Mental Disease*, 180(8), 498-504.

McGlinchey, J. B., & Zimmerman, M. (2007). Examining a dimensional representation of depression and anxiety disorders' comorbidity in psychiatric outpatients with item response modeling. *Journal of Abnormal Psychology*, 116(3), 464-474.

Mergl, R., Seidscheck, I., Allgaier, A.-K., Möller, H.-J., Hegerl, U., & Henkel, V. (2007). Depressive, anxiety, and somatoform disorders in primary care: prevalence and recognition. *Depression and anxiety*, 24, 185-195.

Miller, P. R., Dasher, R., Collins, R., Griffiths, P., & Brown, F. (2001). Inpatient diagnostic assessments: 1. Accuracy of structured and unstructured interviews. *Psychiatry Research*, 105, 255-264.

- Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology*, *49*, 377-412.
- Moggi, F. (2002). *Doppeldiagnosen. Komorbidität psychischer Störungen und Sucht*. Bern: Huber.
- Mulder, R. T. (2002). Personality Pathology and Treatment Outcome in Major Depression: A Review. *Am J Psychiatry*, *159*(3), 359-371.
- Muthén, B. O. (2002). Beyond SEM: General Latent Variable Modeling. *Behaviormetrika*, *29*(1), 81-117.
- Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. A. Kaplan, *Handbook of quantitative methodology for social sciences*. Newbury Park: Sage.
- Muthén, B. O., Brown, C. H., Maysn, K., Jo, B., Khoo, S.-T., Yang, C.-C., Wang, C.-P., et al. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics*, *3*(4), 459-475.
- Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth Mixture Modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*(6), 882-891.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide. Sixth edition* (Bd. 6th). Los Angeles, CA: Muthén & Muthén.
- Nagin, D. S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychological Methods*, *4*(2), 139-157.
- Neale, B. (2005). Liability Threshold Models. In B. S. Everitt & D. C. Howell (Hrsg.), *Encyclopedia of Statistics in Behavioral Science*. Chichester, UK: John Wiley & Sons.
- Newman, D. L., Moffitt, T. E., Caspi, A., & Silva, P. A. (1998). Comorbid mental disorders: Implications for treatment and sample selection. *Journal of Abnormal Psychology*, *107*(2), 305-311.
- Newman, M. G., Przeworski, A., Fisher, A. J., & Borkovec, T. D. (2010). Diagnostic Comorbidity in Adults With Generalized Anxiety Disorder: Impact of Comorbidity on Psychotherapy Outcome and Impact of Psychotherapy on Comorbid Diagnoses. *Behavior Therapy*, *41*, 59-72.
- Newton-Howes, G., Tyrer, P., & Johnson, T. (2006). Personality disorder and the outcome of depression: meta-analysis of published studies. *The British Journal of Psychiatry*, *188*(1), 13-20.
- Noyes, R., Reich, J., Christiansen, J., Suelzer, M., Pfohl, B., & Coryell, W. A. (1990). Outcome of Panic Disorder. *Archives of General Psychiatry*, *47*, 809-818.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling*, *14*(4), 535-569.
- O'Brien, K. M., & Vincent, N. K. (2003). Psychiatric comorbidity in anorexia and bulimia nervosa: nature, prevalence, and causal relationships. *Clinical Psychology Review*, *23*, 57-

74.

O'Donnell, M. L., Creamer, M., & Pattison, P. (2004). Posttraumatic stress disorder and depression following trauma: understanding comorbidity. *American Journal of Psychiatry*, *161*(8), 1390-1396.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy - noch einmal. In A. E. Bergin & S. L. Garfield (Hrsg.), *Handbook of psychotherapy and behavior change* (Bd. 4th, S. 270-376). New York, NY: Wiley & Sons.

Pallister, E., & Waller, G. (2008). Anxiety in the eating disorders: understanding the overlap. *Clinical Psychology Review*, *28*, 366-386.

Paul, T., & Thiel, A. (2005). *Eating Disorder Inventory-2. Deutsche Version*. Göttingen: Hogrefe.

Perry, J. C., Banon, E., & Ianni, F. (1999). Effectiveness of psychotherapy for personality disorders. *American Journal of Psychiatry*, *156*(9), 1312-1221.

Perugi, G., Akiskal, H. S., Pfanner, C., Presta, S., Gemignani, A., Milanfranchi, A., Lensi, P., et al. (1997). The clinical impact of bipolar and unipolar affective comorbidity on obsessive-compulsive disorder. *Journal of Affective Disorders*, *46*, 15-23.

Pilkonis, P. A., & Frank, E. (1988). Personality pathology in recurrent depression: Nature, prevalence, and relationship to treatment response. *American Journal of Psychiatry*, *145*, 435-441.

Preacher, C. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent Growth Curve Modeling*. Thousand Oaks: Sage.

Puschner, B., Kraft, S., & Bauer, S. (2004). Interpersonal problems and outcome in outpatient psychotherapy: Findings from a long-term longitudinal study. *Journal of Personality Assessment*, *83*(3), 223-234.

Ramaswamy, V., DeSabro, W., Reibstein, D., & Robinson, W. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, *12*, 103-124.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. (Bd. 2nd edition). Newbury Park: Sage Publications.

Reardon, M. L., Cukrowicz, K. C., Reeves, M. D., & Joiner, T. E. (2002). Duration and Regularity of Therapy Attendance as Predictors of Treatment Outcome in an Adult Outpatient Population. *Psychotherapy Research*, *12*, 273-285.

Reich, J. (2003). The effects of axis II disorders on the outcome of treatment of anxiety and unipolar depressive disorders: A review. *Journal of Personality Disorders*, *17*, 387-405.

Reich, J. H., & Green, A. I. (1991). Effect of personality disorders on outcome of treatment. *The Journal of Nervous and Mental Disease*, *79*(2), 74-82.

Renaud, J., Brent, D. A., Baugher, M., Birmaher, B., Kolko, D. J., & Bridge, J. (1998). Rapid response to psychological treatment for adolescent depression: A two-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 1184-1191.

- Rhee, S. H., Hewitt, J. K., Corley, R. P., Willcutt, E. G., & Pennington, B. F. (2005). Testing Hypotheses Regarding the Causes of Comorbidity: Examining the Underlying Deficits of Comorbid Disorders. *Journal of Abnormal Psychology, 114*(3), 346-362.
- Rief, W., & Hiller, W. (2011). *Somatisierungsstörung*. Göttingen: Hogrefe.
- Rief, W., Hiller, W., & Heuser, J. (1997). *SOMS - Das Screening für somatoforme Störungen (Manual zum Fragebogen)*. Bern: Huber.
- Rief, W., Schäfer, S., Hiller, W., & Fichter, M. M. (1992). Lifetime diagnoses in patients with somatoform disorders: which came first? *European Archives of Psychiatry and Clinical Neuroscience, 241*, 236-240.
- Rief, W., Trenkamp, S., Auer, C., & Fichter, M. M. (2000). Cognitive Behavior Therapy in Panic Disorder and Comorbid Major Depression. *Psychotherapy and Psychosomatics, 69*, 70-78.
- Rodebaugh, T. L., Holaway, R. M., & Heimberg, R. G. (2004). The treatment of social anxiety disorder. *Clinical Psychology Review, 24*(7), 883-908.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley & Sons.
- Rush, A. J., Zimmerman, M., Wisniewski, S. R., Fava, M., Hollon, S. D., Warden, D., Biggs, M. M., et al. (2005). Comorbid psychiatric disorders in depressed outpatients: Demographic and clinical features. *Journal of Affective Disorders, 87*, 43-55.
- Sanderson, W. C., Beck, A. T., & Beck, J. (1990). Syndrome comorbidity in patients with major depression or dysthymia: prevalence and temporal relationship. *American Journal of Psychiatry, 147*(8), 1025-1028.
- Sartorius, N., Üstün, T. B., Lecrubier, Y., & Wittchen, H.-U. (1996). Depression comorbid with anxiety: results from the WHO Study on Psychological Disorders in Primary Health Care. *British Journal of Psychiatry, 168*(suppl. 30), 38-43.
- Schepank, H. (1995). *BSS. Der Beeinträchtigungs-Schwere-Score. Ein Instrument zur Bestimmung der Schwere einer psychogenen Erkrankung. Manual und Testmappe*. Göttingen: Beltz Test.
- Schneier, F. R., Johnson, J., Hornig, C. D., Liebowitz, M. R., & Weissman, M. M. (1992). Social phobia: Comorbidity and morbidity in an epidemiologic sample. *Archives of General Psychiatry, 49*, 282-288.
- Schonfeld, W. H., Verboncoeur, C. J., Fifer, S. K., Lipschutz, R. C., Lubeck, D. P., & Buesching, D. P. (1997). The functioning and well-being of patients with unrecognized anxiety disorders and major depressive disorder. *Journal of Affective Disorders, 43*, 105-119.
- Schulz, H., Barghaan, D., Harfst, T., & Koch, U. (2008). *Psychotherapeutische Versorgung* (No. 41). Gesundheitsberichterstattung des Bundes. Berlin: Robert-Koch-Institut.
- Schwartz, G. (1978). Estimating dimensions of a model. *Annals of Statistics, 6*, 461-464.
- Seligman, M. E. P. (1995). The effectiveness of psychotherapy. The consumer reports study. *American Psychologist, 50*(12), 965-974.

- Shalev, A. Y., Freedman, S., Peri, T., Brandes, D., Sahar, T., Orr, S. P., & Pitman, R. K. (1998). Prospective study of Posttraumatic Stress Disorder and Depression following trauma. *American Journal of Psychiatry*, *155*(5), 630-637.
- Shapiro, D. A., Barkham, M., Rees, A., Hardy, G. E., Reynolds, S., & Startup, M. (1994). Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology*, *62*, 522-534.
- Shapiro, D. A., Barkham, M., Stiles, W. B., Hardy, G. E., Rees, A., Reynolds, S., & Startup, M. (2003). Time is of the essence: A selective review of the fall and rise of brief therapy research. *Psychology and Psychotherapy: Theory, Research and Practice*, *76*(3), 211-235.
- Shea, M. T., Pilkonis, P. A., Beckham, E., Clollins, J. C., Elkin, I., Sotsky, S. M., & Docherty, J. P. (1990). Personality disorders and treatment outcome in the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry*, *147*(6), 711-718.
- Shea, M. T., Widiger, T. A., & Klein, M. H. (1992). Comorbidity of Personality Disorders and Depression: Implications for treatment. *Journal of Consulting and Clinical Psychology*, *60*(6), 857-868.
- Shear, M. K., Greeno, C., Kang, J., Ludewig, D., Frank, E., Swartz, H. A., & Hanekamp, M. (2000). Diagnosis of nonpsychotic patients in community clinics. *American Journal of Psychiatry*, *157*(4), 581-587.
- Sher, K. J., & Trull, T. J. (1996). Methodological issues in psychopathological research. *Annual Review of Psychology*, *47*, 371-400.
- Sherbourne, C. D., & Wells, K. B. (1997). Course of depression in patients with comorbid anxiety disorders. *Journal of Affective Disorders*, *43*(3), 245-250.
- Skodol, A. E., Grilo, C. M., Pagano, M. E., Bender, D. S., Gunderson, J. G., Shea, M. T., Yen, S., et al. (2005). Effects of personality disorders on functioning and well-being in Major Depressive Disorder. *Journal of Psychiatric Practice*, *11*(6), 363-368.
- Skodol, A. E., Oldham, J. M., Hyler, S. E., Kellman, H. D., Doidge, N., & Davies, M. (1993). Comorbidity of DSM-III-R eating disorders and personality disorders. *International Journal of Eating Disorders*, *14*(4), 403-416.
- Skodol, A. E., Oldham, J. M., Hyler, S. E., Stein, D. J., Hollander, E., Gallaher, P. E., & Lopez, A. E. (1995). Patterns of anxiety and personality disorder comorbidity. *Journal of Psychiatric Research*, *29*(5), 361-374.
- Stein, M. B., & Kean, Y. M. (2000). Disability and quality of life in social phobia: epidemiologic findings. *American Journal of Psychiatry*, *157*(10), 1606-1613.
- Steketee, G., Chambless, D. L., & Tran, G. Q. (2001). Effects of Axis I and II comorbidity on behavior therapy outcome for obsessive-compulsive disorder and agoraphobia. *Comprehensive Psychiatry*, *42*(1), 76-86.
- Stiles, W. B., Leach, C., Barkham, M., Lucock, M., Iveson, S., Iveson, M., Shapiro, D. A., et al. (2003). Early sudden gains in psychotherapy under routine clinic conditions: Prac-

tice-based evidence. *Journal of Consulting and Clinical Psychology*, 71, 14-21.

Stiles, William B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in Psychotherapy. *Clinical Psychology: Science and Practice*, 5(4), 439-458.

Strain, J. J., Smith, G. C., Hammer, J. S., McKenzie, D. P., Blumenfield, M., Muskin, P., Newstadt, G., et al. (1998). Adjustment disorder: a multisite study of its utilization and interventions in the consultation-liaison psychiatry setting. *General Hospital Psychiatry*, 20(3), 139-149.

Stuart, S., Pfohl, B., Battaglia, M., Bellodi, L., Grove, W., & Cadoret, R. (1998). The cooccurrence of DSM-III-R personality disorders. *Journal of Personality Disorders*, 12(4), 302-315.

Stulz, N., & Lutz, W. (2007). Multidimensional patterns of change in outpatient psychotherapy: The Phase Model revisited. *Journal of Clinical Psychology*, 63(9), 817-833.

Stulz, N., Lutz, W., Kopta, S. M., & Saunders, S. M. (in Vorbereitung). The Dose-Effect Relationship in Routine Outpatient Psychotherapy: Does Treatment Duration Matter?

Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology*, 75(6), 864-874.

Stulz, Niklaus, Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology*, 75(6), 864-874.

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 67, 894-904.

Tholen, S. (2006). *Die Komorbidität psychischer Störungen. Eine metaanalytische Untersuchung und die Entwicklung eines Fragebogens zur Erfassung der strukturellen Komorbidität psychischer Störungen*. Universität Bern, Bern.

Tohen, M., Calabrese, J., Vieta, E., Bowden, C., Gonzalez-Pinto, A., Lin, D., Xu, W., et al. (2007). Effect of comorbid anxiety on treatment response in bipolar depression. *Journal of Affective Disorders*, 104(1-3), 137-146.

Torres, A. R., Prince, M. J., Bebbington, P. E., Bhugra, D., Bhugra, T. S., Farrell, M., Jenkins, R., et al. (2006). Obsessive-Compulsive Disorder: Prevalence, comorbidity, impact, and help-seeking in the British National Psychiatric Morbidity Survey of 2000. *American Journal of Psychiatry*, 163(11), 1978-1985.

Tsao, J. C., Lewin, M. R., & Craske, M. G. (1998). The Effects of Cognitive-Behavior Therapy for Panic Disorder on Comorbid Conditions. *Journal of Anxiety Disorders*, 12(4), 357-371.

Tschitsaz-Stucki, A., & Lutz, W. (2009). Identifikation und Aufklärung von Veränderungssprüngen im individuellen Therapieverlauf. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 38(1), 13-23.

Tükel, R., Polat, A., Özdemir, Ö., Aksüt, D., & Türksöy, N. (2002). Comorbid conditions in obsessive-compulsive disorder. *Comprehensive Psychiatry*, 43(3), 204-209.

- Urban, D. & Mayerl, J. (2006). *Regressionsanalyse: Theorie, Technik und Anwendung*. Wiesbaden: Verlag für Sozialwissenschaften.
- Vella, G., Aragona, M., & Alliani, D. (2000). The Complexity of Psychiatric Comorbidity: A Conceptual and Methodological Discussion. *Psychopathology*, 33(1), 25-30.
- Viinamäki, H., Hintikka, J., Honkalampi, K., Koivumaa-Honkanen, H., Kuisma, S., Antikainen, R., Tanskanen, A., et al. (2002). Cluster C personality disorder impedes alleviation of symptoms in major depression. *Journal of Affective Disorders*, 71(1-3), 35-41.
- Weltgesundheitsorganisation. (2005). *Internationale Klassifikation psychischer Störungen. ICD-10 Kapitel V (F)*. Bern: Huber.
- Wetherell, J. L., Hopko, D. R., Diefenbach, G. J., Averill, P. M., Beck, J. G., Craske, M. G., Gatz, M., et al. (2005). Cognitive-behavioral therapy for late-life generalized anxiety disorder: Who gets better? *Behavior Therapy*, 36(2), 147-156.
- Widiger, T. A., & Samuel, D. B. (2005). Diagnostic Categories or Dimensions? A Question for the Diagnostic and Statistical Manual of Mental Disorders—Fifth Edition. *Journal of Abnormal Psychology*, 114(4), 494-504.
- Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology, Research and Practice*, 24(2), 190-195.
- Wirtz, M. A. (1998). *Statistische Methoden für Psychologen: Deskriptive Statistik*. Weinheim: Juventa.
- Wittchen, H.-U. (1996). Critical issues in the evaluation of comorbidity of psychiatric disorders. *British Journal of Psychiatry*, 168(supp. 30), 9-16.
- Wittchen, H.-U., Carter, R. M., Pfister, H., Montgomery, S. A., & Kessler, R. C. (2000). Disabilities and quality of life in pure and comorbid generalized anxiety disorder and major depression in a national survey. *International Clinical Psychopharmacology*, 15, 319-328.
- Wittchen, H.-U., & Hoyer, J. (2011). *Klinische Psychologie & Psychotherapie*. Berlin: Springer.
- Wittchen, H.-U., & Jacobi, F. (2002). Die Versorgungssituation psychischer Störungen in Deutschland. Eine klinisch-epidemiologische Abschätzung anhand des Bundes-Gesundheitssurveys 1998. *Psychotherapeutenjournal*, 6-15.
- Wittchen, H.-U., Lieb, R., Wunderlich, U., & Schuster, P. (1999). Comorbidity in primary care: presentation and consequences. *Journal of Clinical Psychiatry*, 60, 29-36.
- Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). *SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen*. Göttingen: Hogrefe.
- Wittmann, W. W., Lutz, W., Steffanowski, A., Kriz, D., Glahn, E. M., Völkle, M. C., Böhnke, J. R., et al. (2011). *Qualitätsmonitoring in der ambulanten Psychotherapie: Modellprojekt der Techniker Krankenkasse - Abschlussbericht*. Hamburg: Techniker Krankenkasse.

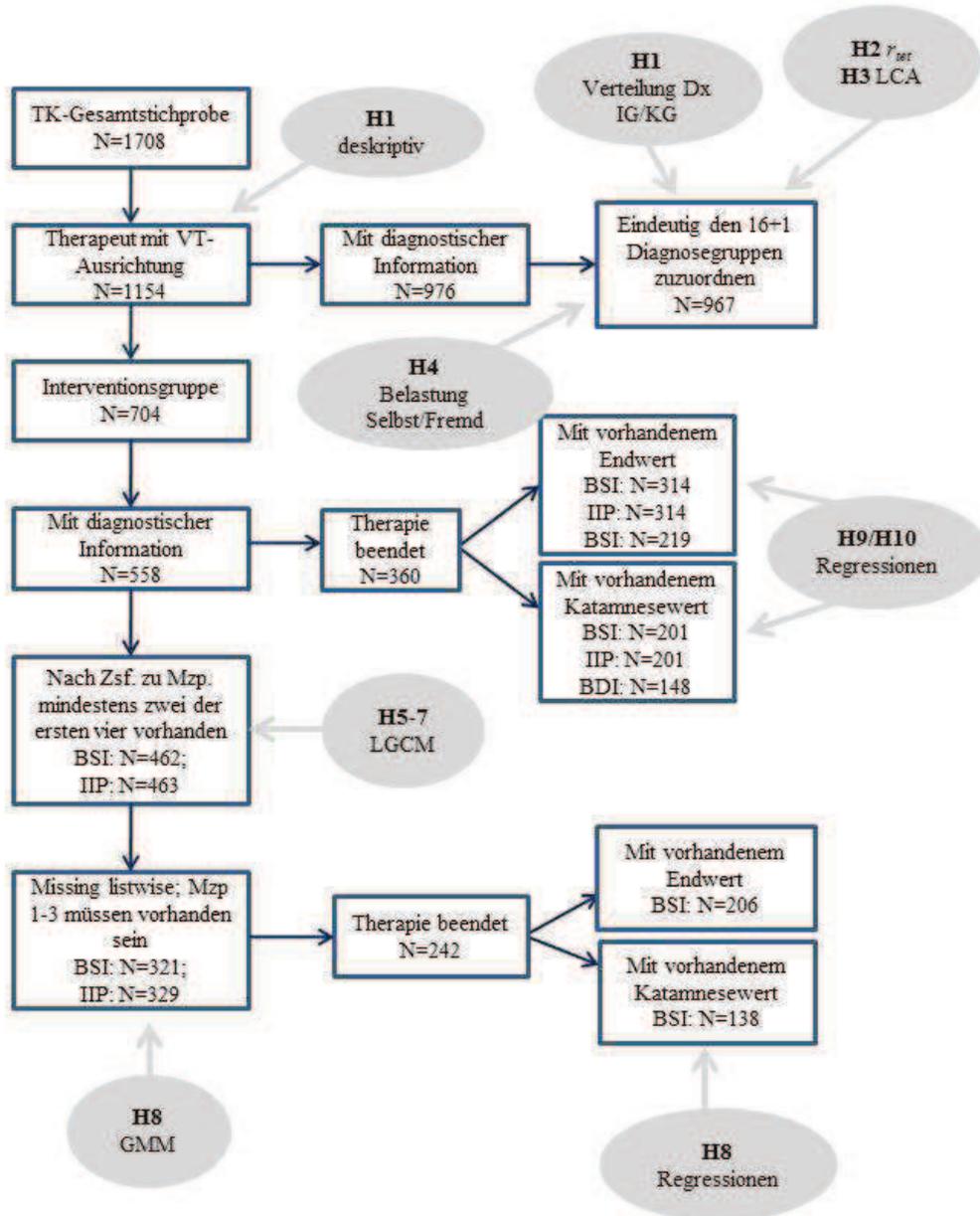
- Wonderlich, S. A., Fullerton, D., Swift, W. J., & Klein, M. H. (1994). Five-year outcome from eating disorders: Relevance of personality disorders. *International Journal of Eating Disorders*, *15*(3), 233-243.
- Yonkers, K. A., Dyck, I. R., Warshaw, M., & Keller, M. B. (2000). Factors predicting the clinical course of generalised anxiety disorder. *The British Journal of Psychiatry*, *176*(6), 544-549.
- Young, L. T., Cooke, R. G., Robb, J. C., Levitt, A. J., & Joffe, R. T. (1993). Anxious and non-anxious bipolar disorder. *Journal of Affective Disorders*, *29*, 49-53.
- Yuan, Y. C. *Multiple Imputation for Missing Data: Concepts and New Development*. Online verfügbar unter <http://www.ats.ucla.edu/stat/sas/library/multipleimputation.pdf>.
- Zimmerman, M. (1994). Diagnosing personality disorders. *Archives of General Psychiatry*, *51*, 225-245.
- Zimmerman, M., & Mattia, J. I. (1999a). Differences between clinical and research practices in diagnosing Borderline Personality Disorder. *American Journal of Psychiatry*, *156*(10), 1570-1574.
- Zimmerman, M., & Mattia, J. I. (1999b). Psychiatric diagnosis in clinical practice: is comorbidity being missed? *Comprehensive Psychiatry*, *40*, 182-191.
- Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The prevalence of DSM-IV personality disorders in psychiatric outpatients. *American Journal of Psychiatry*, *162*, 1911-1918.



## 8 ANHANG

Übersicht über Schritte zur Datenreduktion .....	171
Tabelle 1: Häufigkeit der Diagnosen in den Komorbiditätsgruppen, sowie relativer Anteil an der Gesamtdiagnoseanzahl .....	172
Tabelle 2: Deskriptive Statistiken zu den vier Komorbiditätsmustern .....	173
Verteilungskennwerte der zusammengesetzten Beeinträchtigungsmaße im Selbst- und Fremdbereich zu Beginn der Therapie. ....	174
Tabelle 3: Übersicht über die Belastung nach Komorbiditätsgruppen.....	175
Tabelle 4: Übersicht über die Belastung nach Latenten Diagnoseklassen .....	177
Verteilungskennwerte des BSI zu den vier in die LGM eingehenden Messzeitpunkte .....	179
Verteilungskennwerte des IIP zu den vier in die LGM eingehenden Messzeitpunkte .....	180
Prüfung der Voraussetzungen für die Regressionsanalysen erster Block.....	181
Prüfung der Voraussetzungen für die Regressionsanalysen zweiter Block.....	183
Prüfung der Voraussetzungen für die Vorhersage der zweiten und dritten Zwischenmessung .....	187
Tabelle 5: Anzahl (relativer Anteil) im BSI reliabel verbesserter Patienten in den identifizierten Klassen zum Therapieende und Katamnesezeitpunkt .....	189
Tabelle 6: Regressionsanalysen (schrittweise) zur Vorhersage des Status in BSI, IIP und BDI zum zweiten und dritten Messzeitpunkt unter Berücksichtigung von Komorbidität.....	190
Tabelle 7: Anzahl Sitzungen bis zum Therapieende nach Komorbiditätsmuster und Anzahl Diagnosen.....	191
Tabelle 8: Deskriptive Statistiken der Hauptbelastungsmaße zum Therapieende, differenziert nach Komorbiditätsgruppen .....	191
Tabelle 9: Anzahl (relativer Anteil) im BSI, IIP und BDI reliabel verbesserter nach Komorbiditätsmuster zum Therapieende und Katamnesezeitpunkt .....	192

Übersicht über Schritte zur Datenreduktion



**Tabelle 1: Häufigkeit der Diagnosen in den Komorbiditätsgruppen, sowie relativer Anteil an der Gesamtdiagnoseanzahl**

Diagnose nach ICD-10	keine Komorbidität		Komorbidität mit PS		Komorbidität Angst/Affekt. St.		andere Komorbidität	
	(n=842)		(n=115)		(n=162)		(n=332)	
Organische psychische Störungen	3	0.3%	-	-	-	-	-	-
Psychische Störungen durch psychotrope Substanzen	4	0.4%	9	3.2%	6	1.5%	24	3.4%
Schizophrenie und wahnhafte Störungen	-	-	-	-	-	-	5	0.7%
Manische Episode	1	0.1%	-	-	1	0.2%	-	-
Bipolare affektive Störung	7	0.8%	2	0.7%	1	0.2%	2	0.2%
Depressive Episode	211	25.1%	21	7.5%	77	19.1%	81	11.5%
Rezidivierende depressive Störung	112	13.3%	43	15.0%	55	13.6%	73	10.3%
Zyklothymia	65	7.7%	6	2.1%	11	2.7%	21	3.0%
Dysthymia	-	-	10	3.6%	21	5.2%	26	3.7%
andere affektive Störung	-	-	-	-	-	-	1	0.1%
Agoraphobie ohne Panikstörung	58	6.9%	8	2.9%	41	10.2%	30	4.2%
Agoraphobie mit Panikstörung	-	-	-	-	-	-	-	-
soziale Phobie	-	-	4	1.4%	37	9.2%	18	2.5%
spezifische Phobie	-	-	3	1.1%	16	4.0%	9	1.3%
andere phobische Störungen	-	-	-	-	3	0.7%	1	0.1%
Panikstörung	93	11.0%	8	2.9%	48	11.9%	48	6.8%
GAS	-	-	6	2.1%	27	6.7%	7	0.9%
Angst und depressive Störung gemischt	-	-	1	0.4%	5	1.2%	5	0.7%
Andere Angststörung	-	-	-	-	7	1.7%	1	0.1%
Zwangsstörung	15	1.8%	2	0.7%	11	2.7%	29	4.1%
Akute Belastungsreaktion	203	24.1%	12	4.3%	7	1.7%	84	11.9%
PTBS	-	-	2	0.7%	4	0.9%	21	3.0%
Anpassungsstörung	-	-	5	1.8%	1	0.2%	46	6.5%

Andere Reaktion auf schwere Belastungen	-	-	-	-	-	-	1	0.1%
Dissoziative Störungen	2	0.2%	1	0.4%	1	0.2%	11	1.6%
Somatoforme Störung	26	3.1%	10	3.6%	13	3.2%	85	12.0%
Sonstige neurotische Störung	4	0.4%	1	0.4%	-	-	4	0.6%
Anorexia nervosa	19	2.3%	4	1.4%	2	0.4%	17	2.4%
nichtorganische Schlafstörungen	1	0.1%	-	-	2	0.4%	7	0.9%
nichtorganische sexuelle Funktionsstörungen	1	0.1%	-	-	-	-	3	0.4%
Andere psychische Störungen	6	0.7%	2	0.7%	2	0.4%	31	4.4%
Persönlichkeitsstörung	11	1.3%	117	41.8%	-	-	-	-
Bulimia nervosa	-	-	2	0.7%	2	0.4%	4	0.6%
sonstige Essstörung	-	-	1	0.4%	2	0.4%	11	1.6%
Anzahl Diagnosen gesamt	842		280		403		706	

Anmerkung: die Prozentangaben beziehen sich auf die Gesamtzahl vergebener Diagnosen in der jeweiligen Komorbiditätsgruppe

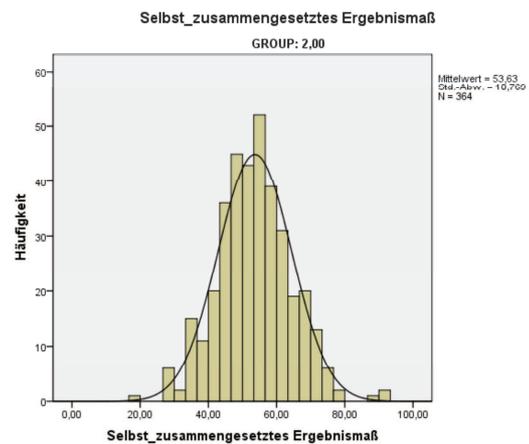
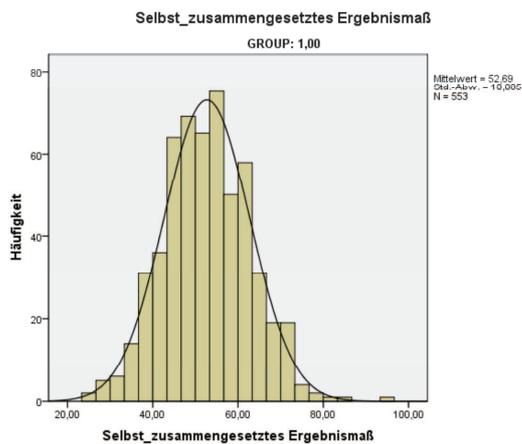
**Tabelle 2: Deskriptive Statistiken zu den vier Komorbiditätsmustern**

	Alter (Monate)	Geschlecht (N, % weiblich)	BSI-GSI	IIP Gesamt	BDI Gesamt	GAF	beendete Therapien (N, %)
keine Komorbidität (N=842)	N=838 M=533 SD=135	580 (68.9%)	N=777 M=1.17 SD=.66	N=781 M=1.48 SD=.54	N=600 M=20.01 SD=9.34	N=779 M=57.43 SD=10.27	468 (55.6%)
Komorbidität mit Persönlichkeitsstörung (N=115)	N=115 M=524 SD=127	64 (55.7%)	N=105 M=1.37 SD=.64	N=108 M=1.75 SD=.55	N=88 M=10.03 SD=8.47	N=109 M=52.45 SD=10.63	62 (53.9%)
Komorbidität Angst/Affektive Störungen (N=162)	N=161 M=539 SD=124	113 (69.8%)	N=154 M=1.35 SD=.70	N=155 M=1.64 SD=.51	N=97 M=22.79 SD=9.31	N=154 M=54.44 SD=9.74	83 (51.2%)
Komorbidität andere (N=332)	N=328 M=524 SD=138	230 (69.3%)	N=317 M=1.22 SD=.65	N=315 M=1.55 SD=.55	N=206 M=20.97 SD=8.89	N=307 M=58.64 SD=41.58	166 (50.0%)

### Verteilungskennwerte der zusammengesetzten Beeinträchtigungsmaße im Selbst- und Fremdbesicht zu Beginn der Therapie.

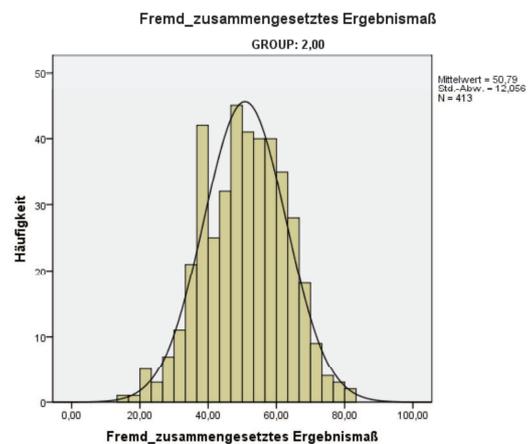
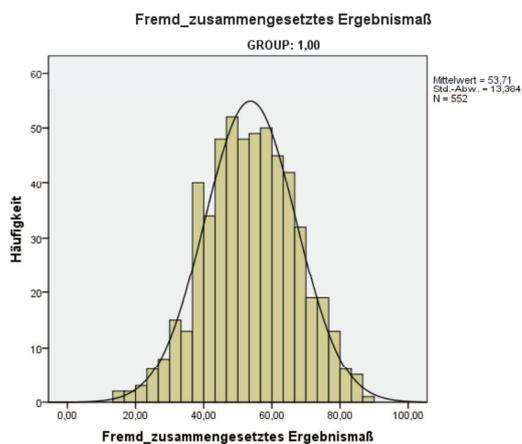
Selbstbericht, Interventionsgruppe: N=553; M=52.7; SD=10.1; Range: Min=25; Max=96; Skewness=.174; Kurtosis=.343; K-S-Test auf Normalverteilung: z=0.021, p=.200

Selbstbericht, Kontrollgruppe: N=364; M=53.6; SD=10.8; Range: Min=17; Max=92; Skewness=-.060; Kurtosis=-.237; K-S-Test auf Normalverteilung: z=0.030, p=.200



Fremdbesicht, Interventionsgruppe: N=552; M=53.7; SD=13.4; Range: Min=15; Max=89; Skewness=.170; Kurtosis=.671; K-S-Test auf Normalverteilung: z=0.022, p=.200

Fremdbesicht, Kontrollgruppe: N=413; M=50.1; SD=12.1; Range: Min=14; Max=82; Skewness=-.193; Kurtosis=-.285; K-S-Test auf Normalverteilung: z=0.046, p=.058



**Tabelle 3: Übersicht über die Belastung nach Komorbiditätsgruppen**

	keine Komorbidität (1)	Kom. mit Per- sönlichkeits- störung (2)	Kom. Angst- /Affektive Stö- rungen (3)	Kom., andere (4)	ANOVA F (ges) / sign. Einzelvgl.	
BSI	N=510 M=1.17 SD=.66	N=73 M=1.41 SD=.62	N=107 M=1.43 SD=.72	N=218 M=1.23 SD=.68	6.090 (1-2), (1-3)	
IIP	N=510 M=1.46 SD=.55	N=76 M=1.81 SD=.56	N=108 M=1.64 SD=.54	N=218 M=1.52 SD=.56	10.278 (1-2), (1-3), (2-4)	
SF-12 körperlich	N=484 M=46.71 SD=10.23	N=76 M=47.40 SD=10.40	N=101 M=45.75 SD=11.01	N=204 M=44.93 SD=11.60	1.722	
SF-12 psychisch	N=484 M=33.85 SD=11.04	N=76 M=29.56 SD=9.11	N=101 M=31.28 SD=10.50	N=206 M=32.99 SD=10.67	3.404 (1-2)	
GAF	N=508 M=57.43 SD=11.18	N=79 M=52.00 SD=11.67	N=111 M=53.35 SD=9.36	N=208 M=56.80 SD=28.14	3.806 (1-2)	
BSS Beeinträchtigungs-Schwere-Score zu Beginn der Therapie	letzte 7 Tage körperlich	N=540 M=1.59 SD=1.03	N=83 M=1.58 SD=1.03	N=113 M=1.78 SD=1.08	N=219 M=1.75 SD=1.09	1.879
	letzte 7 Tage psychisch	N=540 M=2.65 SD=0.74	N=83 M=2.99 SD=0.72	N=113 M=2.78 SD=0.74	N=219 M=2.83 SD=0.71	7.120 (1-2); (1-4)
	letzte 7 Tage sozial- kommunikativ	N=540 M=1.79 SD=1.01	N=83 M=2.41 SD=0.81	N=113 M=2.07 SD=0.94	N=218 M=1.89 SD=1.01	10.917 (1-2); (1-3); (2-4)
	letzte 12 Mo- nate körperlich	N=536 M=1.83 SD=1.08	N=83 M=1.81 SD=1.09	N=113 M=2.03 SD=1.01	N=217 M=2.10 SD=1.13	3.811 (1-4)
	letzte 12 Mo- nate psychisch	N=538 M=2.74 SD=0.76	N=83 M=2.94 SD=0.82	N=113 M=2.93 SD=0.66	N=217 M=2.83 SD=0.81	3.165
	letzte 12 Mo- nate sozial- kommunikativ	N=536 M=1.95 SD=1.03	N=83 M=2.57 SD=0.86	N=113 M=2.33 SD=0.85	N=216 M=1.98 SD=1.05	12.260 (1-2); (1-3); (2-4); (3-4)

Fortsetzung Tabelle 3: Übersicht über die Belastung nach Komorbiditätsgruppen

	keine Komorbidität (1)	Kom. mit Per- sönlichkeits- störung (2)	Kom. Angst- /Affektive Stö- rungen (3)	Kom., andere (4)	ANOVA F (ges) / sign. Einzelvgl.	
Problemrelevante Lebensbereiche	Beruf/ Ausbildung	N=542 M=4.55 SD=3.41	N=82 M=6.02 SD=2.98	N=113 M=5.60 SD=3.19	N=222 M=5.15 SD=3.25	7.270 (1-2); (1-3)
	Partnerschaft	N=541 M=5.73 SD=3.22	N=82 M=7.16 SD=2.84	N=112 M=5.67 SD=2.91	N=220 M=5.45 SD=3.20	6.138 (1-2); (2-3); (2-4)
	Familie	N=543 M=5.50 SD=2.88	N=82 M=6.43 SD=2.61	N=114 M=5.31 SD=2.74	N=221 M=5.52 SD=2.85	2.953 (1-2); (2-3)
	Freizeit	N=533 M=4.17 SD=2.64	N=83 M=4.96 SD=2.31	N=111 M=4.79 SD=2.65	N=222 M=4.32 SD=2.77	3.374
	Soziale Be- dingungen	N=542 M=3.18 SD=2.79	N=83 M=4.87 SD=2.88	N=113 M=3.52 SD=2.67	N=221 M=3.15 SD=2.78	9.444 (1-2); (2-3); (2-4)
	Selbstverant- wortung	N=541 M=3.94 SD=2.61	N=83 M=5.49 SD=2.69	N=113 M=4.38 SD=2.85	N=221 M=4.26 SD=2.79	8.288 (1-2); (2-3); (2-4)
	Gesundheits- verhalten	N=537 M=3.77 SD=2.61	N=83 M=4.95 SD=2.62	N=110 M=4.04 SD=2.71	N=218 M=4.19 SD=2.80	5.186 (1-2)

**Tabelle 4: Übersicht über die Belastung nach Latenten Diagnoseklassen**

	Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	ANOVA F (ges) / sign. Einzelvgl.	
BSI	N=165 M=1.41 SD=0.65	N=298 M=1.18 SD=0.71	N=237 M=1.31 SD=0.67	N=29 M=1.27 SD=0.69	N=179 M=1.08 SD=0.60	6.376 (1-2); (1-5); (3-5)	
IIP	N=166 M=1.70 SD=0.54	N=298 M=1.47 SD=0.59	N=237 M=1.55 SD=0.55	N=30 M=1.58 SD=0.56	N=181 M=1.42 SD=0.51	6.410 (1-2); (1-5)	
SF-12 körperlich	N=167 M=45.09 SD=10.75	N=299 M=47.44 SD=10.30	N=239 M=45.74 SD=10.81	N=30 M=44.36 SD=12.42	N=182 M=46.59 SD=11.29	1.571	
SF-12 psychisch	N=167 M=30.12 SD=10.13	N=299 M=34.45 SD=11.14	N=239 M=30.61 SD=10.11	N=30 M=32.33 SD=12.08	N=182 M=34.024 SD=10.85	6.704 (1-2); (1-5); (2-3); (3-5)	
GAF	N=170 M=51.61 SD=11.14	N=290 M=58.21 SD=24.39	N=238 M=54.27 SD=9.47	N=26 M=55.27 SD=11.95	N=182 M=60.48 SD=11.31	8.405 (1-2); (1-5); (2-3); (3-5)	
BSS Beeinträchtigungs-Schwere-Score zu Beginn der Therapie	letzte 7 Tage körperlich	N=177 M=1.74 SD=1.02	N=307 M=1.61 SD=1.10	N=254 M=1.63 SD=1.07	N=30 M=1.87 SD=1.04	N=187 M=1.63 SD=0.98	.826
	letzte 7 Tage psychisch	N=177 M=2.83 SD=.72	N=307 M=2.64 SD=0.74	N=254 M=2.89 SD=0.72	N=30 M=2.70 SD=0.66	N=187 M=2.60 SD=0.75	6.413 (1-2); (1-5); (2-3); (3-5)
	letzte 7 Tage sozial- kommunikativ	N=177 M=2.09 SD=.94	N=307 M=1.77 SD=1.08	N=253 M=2.10 SD=.93	N=30 M=1.90 SD=.89	N=187 M=1.65 SD=.95	8.661 (1-2); (1-5); (2-3); (3-5)
	letzte 12 Mo- nate körperlich	N=177 M=1.97 SD=1.08	N=306 M=1.88 SD=1.10	N=251 M=2.01 SD=1.07	N=30 M=1.97 SD=1.16	N=185 M=1.79 SD=1.09	1.260
	letzte 12 Mo- nate psychisch	N=177 M=2.97 SD=0.72	N=306 M=2.79 SD=0.72	N=252 M=2.88 SD=0.77	N=30 M=2.57 SD=0.82	N=186 M=2.59 SD=0.82	7.039 (1-5); (2-5); (3-5)
	letzte 12 Mo- nate sozial- kommunikativ	N=177 M=2.32 SD=0.94	N=305 M=1.97 SD=1.07	N=252 M=2.20 SD=0.97	N=30 M=1.90 SD=1.06	N=184 M=1.78 SD=0.99	8.780 (1-2); (1-5); (2-3); (3-5)

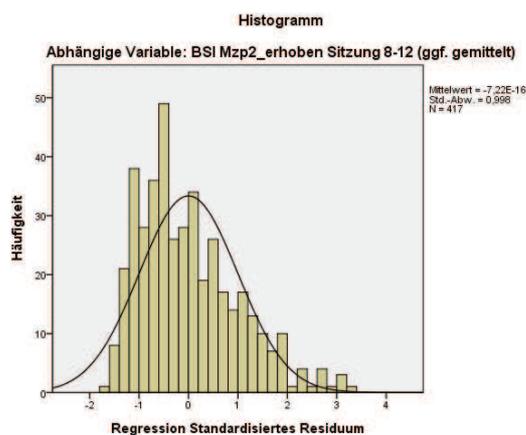
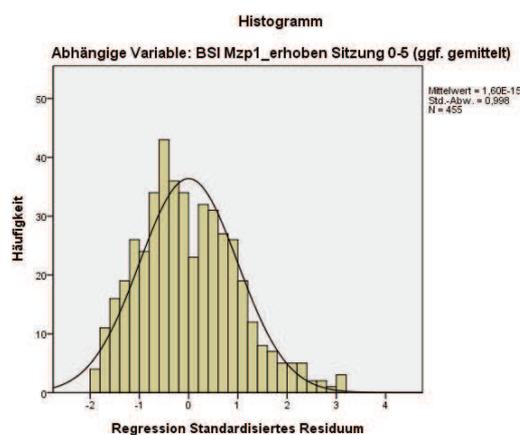
**Fortsetzung Tabelle 4: Übersicht über die Belastung nach Latenten Diagnoseklassen**

		Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	ANOVA F (ges) / sign. Einzelvgl.
Problemrelevante Lebensbereiche	Beruf/ Ausbildung	N=177 M=5.56 SD=3.23	N=307 M=4.53 SD=3.20	N=255 M=5.41 SD=3.37	N=32 M=5.25 SD=3.81	N=188 M=4.32 SD=3.36	5.750 (1-2); (1-5); (2-3); (3-5)
	Partnerschaft	N=175 M=6.15 SD=2.99	N=309 M=5.10 SD=3.07	N=253 M=6.24 SD=3.15	N=32 M=5.47 SD=3.57	N=186 M=6.00 SD=3.31	5.945 (1-2); (2-3); (2-5)
	Familie	N=177 M=5.85 SD=2.82	N=308 M=5.10 SD=2.87	N=254 M=6.09 SD=2.72	N=32 M=5.44 SD=3.04	N=189 M=5.34 SD=2.82	5.053 (1-2); (2-3); (3-5)
	Freizeit	N=175 M=5.06 SD=2.55	N=308 M=4.16 SD=2.67	N=250 M=4.56 SD=2.63	N=32 M=3.59 SD=2.66	N=184 M=3.83 SD=2.59	6.519 (1-2); (1-4); (1-5); (3-5)
	Soziale Be- dingungen	N=176 M=4.13 SD=2.99	N=310 M=3.02 SD=2.72	N=252 M=3.50 SD=2.75	N=32 M=2.56 SD=2.78	N=189 M=3.15 SD=2.76	5.575 (1-2); (1-4); (1-5)
	Selbstverant- wortung	N=177 M=4.84 SD=2.77	N=308 M=4.00 SD=2.65	N=252 M=4.26 SD=2.75	N=32 M=4.41 SD=2.63	N=189 M=3.81 SD=2.68	3.895 (1-2); (1-5)
	Gesundheits- verhalten	N=176 M=4.49 SD=2.76	N=301 M=3.92 SD=2.75	N=251 M=4.12 SD=2.70	N=31 M=4.00 SD=3.16	N=189 M=3.52 SD=2.32	3.146 (1-5)

## Verteilungskennwerte des BSI zu den vier in die LGM eingehenden Messzeitpunkte

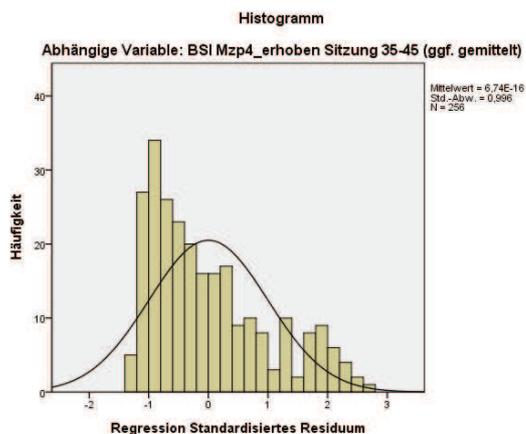
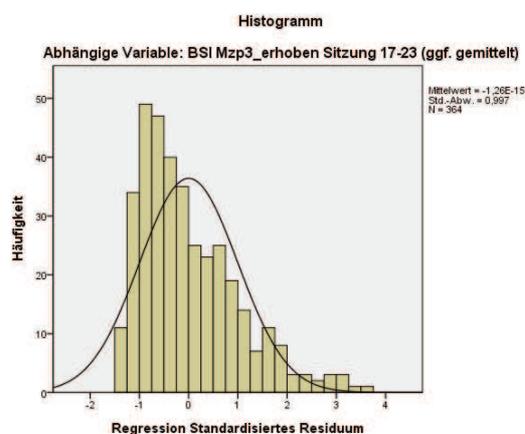
Zusammengesetzter Messzeitpunkt 1: N=455; M=1.30; SD=.65; Range: Min=.04; Max=3.32; Skewness=.518; Kurtosis=.031; K-S-Test auf Normalverteilung:  $z=0.066$ ,  $p=.000$

Zusammengesetzter Messzeitpunkt 2: N=417; M=1.01; SD=.64; Range: Min=.00; Max=3.17; Skewness=.826; Kurtosis=.273; K-S-Test auf Normalverteilung:  $z=0.099$ ,  $p=.000$



Zusammengesetzter Messzeitpunkt 3: N=364; M=.94; SD=.67; Range: Min=.04; Max=3.38; Skewness=1.085; Kurtosis=.935; K-S-Test auf Normalverteilung:  $z=0.108$ ,  $p=.000$

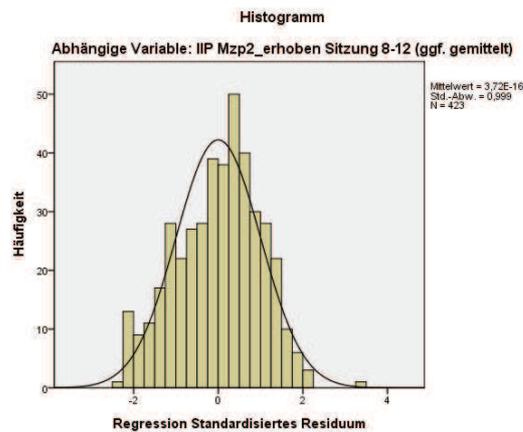
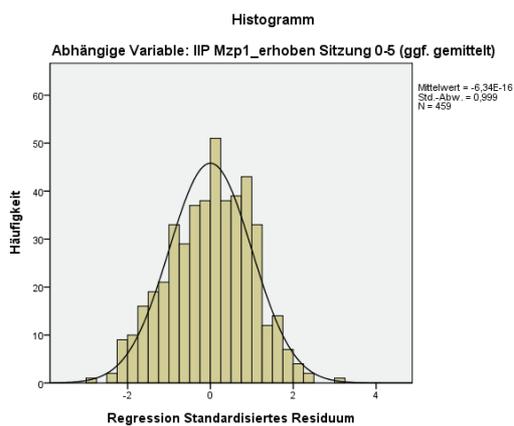
Zusammengesetzter Messzeitpunkt 4: N=265; M=.81; SD=.63; Range: Min=.00; Max=2.47; Skewness=.865; Kurtosis=-.244; K-S-Test auf Normalverteilung:  $z=0.122$ ,  $p=.000$



## Verteilungskennwerte des IIP zu den vier in die LGM eingehenden Messzeitpunkte

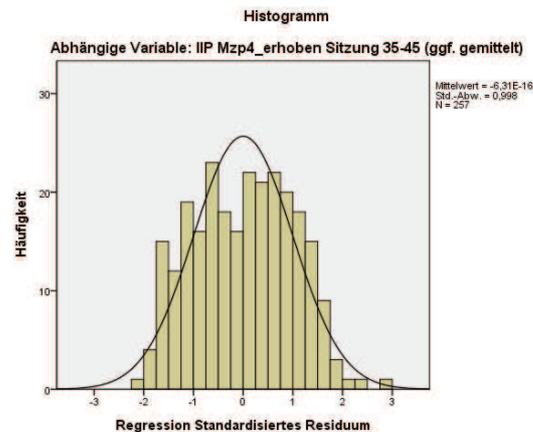
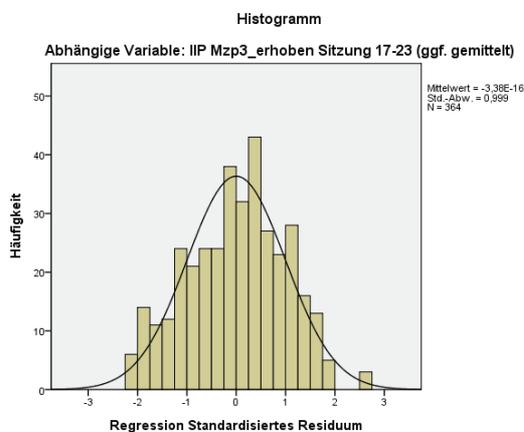
Zusammengesetzter Messzeitpunkt 1: N=459; M=1.58; SD=.55; Range: Min=.06; Max=3.28; Skewness=-.133; Kurtosis=-.396; K-S-Test auf Normalverteilung:  $z=0.040$ ,  $p=.080$

Zusammengesetzter Messzeitpunkt 2: N=423; M=1.48; SD=.57; Range: Min=.05; Max=3.41; Skewness=-.179; Kurtosis=-.361; K-S-Test auf Normalverteilung:  $z=0.057$ ,  $p=.002$



Zusammengesetzter Messzeitpunkt 3: N=364; M=1.36; SD=.60; Range: Min=.05; Max=2.95; Skewness=-.107; Kurtosis=-.581; K-S-Test auf Normalverteilung:  $z=0.041$ ,  $p=.200$

Zusammengesetzter Messzeitpunkt 4: N=257; M=1.25; SD=.61; Range: Min=.02; Max=2.84; Skewness=.019; Kurtosis=-.874; K-S-Test auf Normalverteilung:  $z=0.068$ ,  $p=.006$



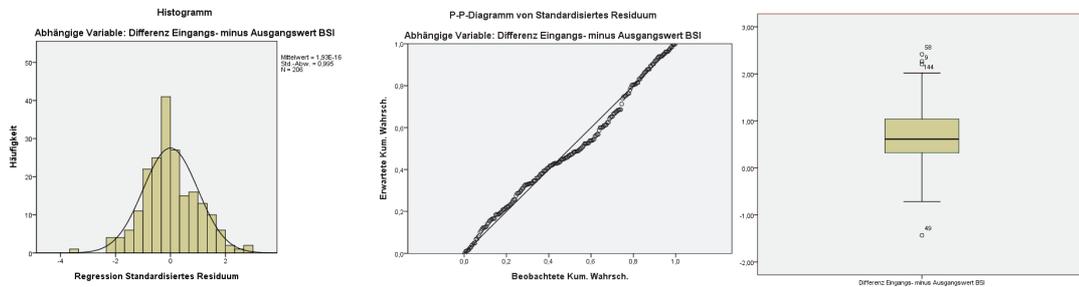
## Prüfung der Voraussetzungen für die Regressionsanalysen erster Block

### Korrelationsmatrix

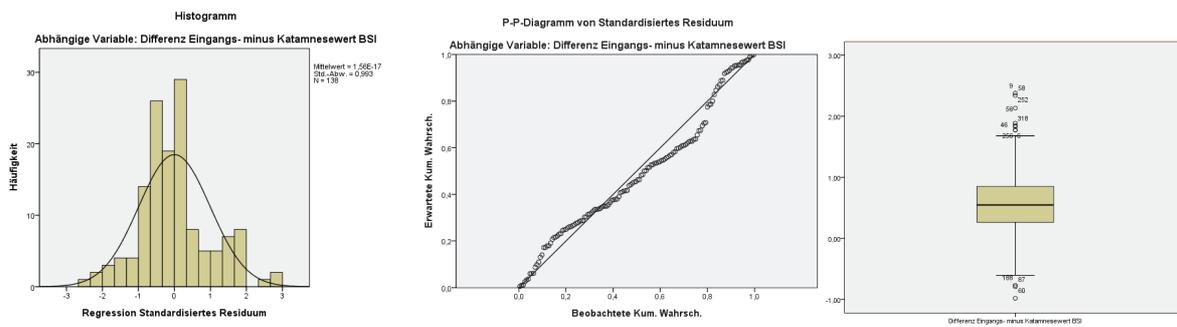
	1	2	3	4	5	6	7	8	9	10	11	12
1 BSI-Differenz Ende												
2 BSI-Differenz Kat.	.082**											
3 BSI-GSI Eingang	.568**	.552**										
4 Alter	-.072	.059	-.010									
5 Geschlecht	.029	.025	.065	-.118*								
6 Anzahl Sitzungen	-.004	-.067	.097	-.033	-.008							
7 Klasse 1	.283**	-.099	-.200*	.011	.030	-.009						
8 Klasse 2	.268**	-.194**	-.040	.021	-.017	.177**	-.125*					
9 Klasse 3	-.524**	-.019	-.200*	-.011	-.052	-.095	-.284**	.779**				
10 KOM PS	.029	.012	.078	-.012	-.027	.264**	.021	.086	-.061			
11 KOM Angst/Affekt.	.082	.021	.088	-.006	.073	.041	-.038	.048	-.093	.138**		
12 KOM andere	.096	-.037	-.113	-.080	.008	.065	.087	.170**	.175**	.206**	.225**	
13 Anzahl Diagnosen	.041	.113	.160**	-.106	.063	.288**	.023	.209**	.224**	.388**	.452**	.366**

Histogramme, Streudiagramme sowie Boxplots der in die Regressionsanalysen eingehenden abhängigen Variablen BSI-Differenz Ende, BSI-Differenz Katamnese und Anzahl Sitzungen (bzw. deren Transformationen).

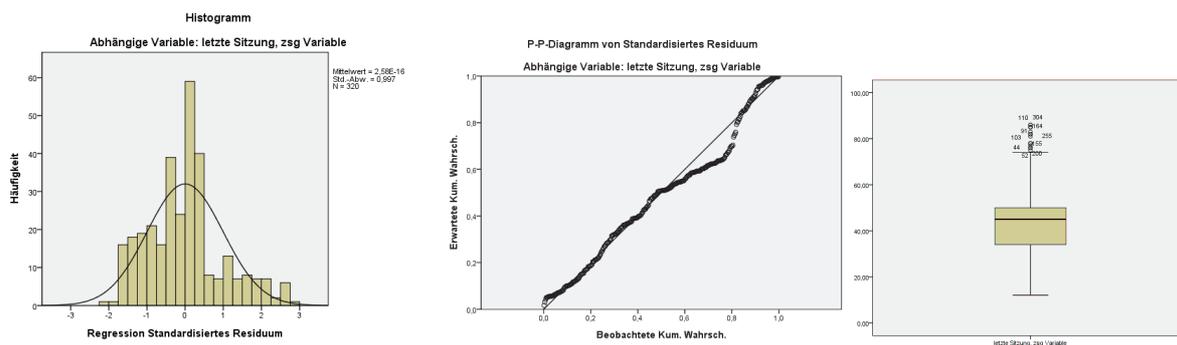
a) BSI-Differenz Ende



b) BSI-Differenz Katamnese



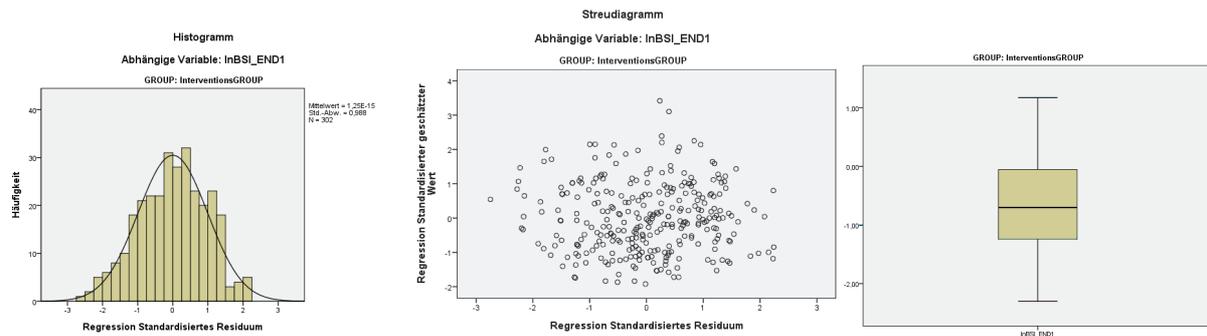
c) Anzahl Sitzungen



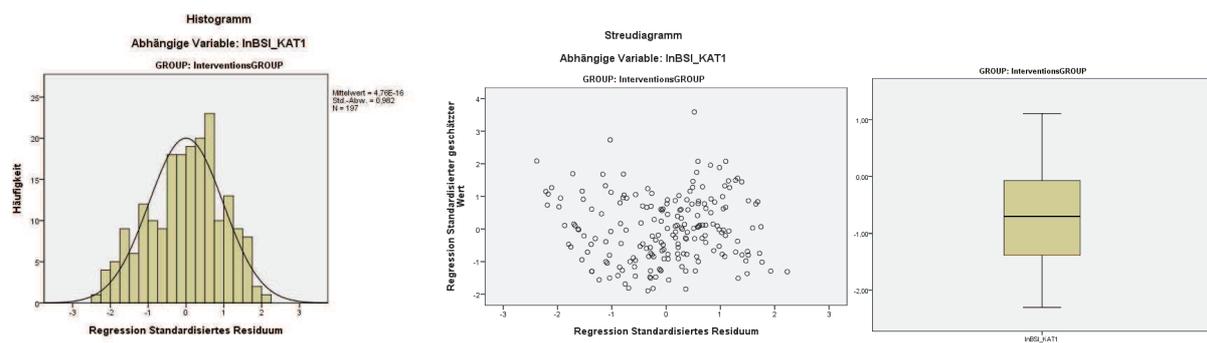
### Prüfung der Voraussetzungen für die Regressionsanalysen zweiter Block

Histogramme, Streudiagramme sowie Boxplots der in die Regressionsanalysen eingehenden abhängigen Variablen BSI, IIP, BDI und Anzahl Sitzungen (bzw. deren Transformationen) zum Therapieende und zur Katamnese.

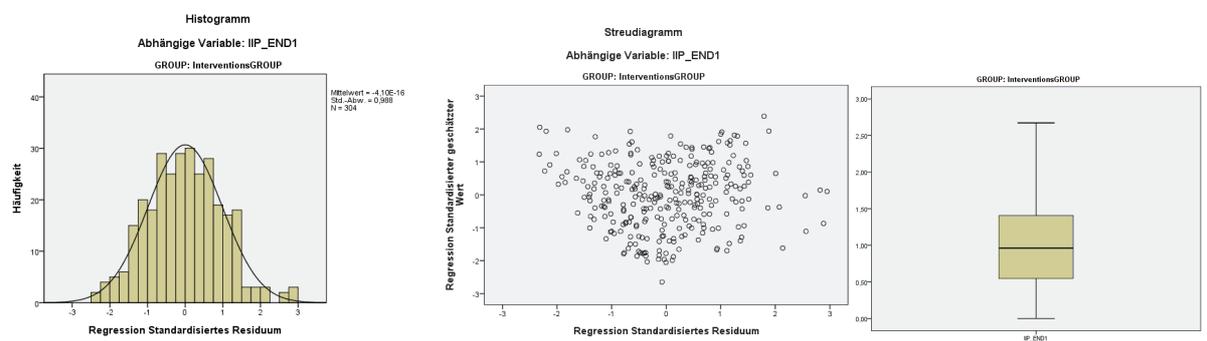
#### a) lnBSI, Therapieende



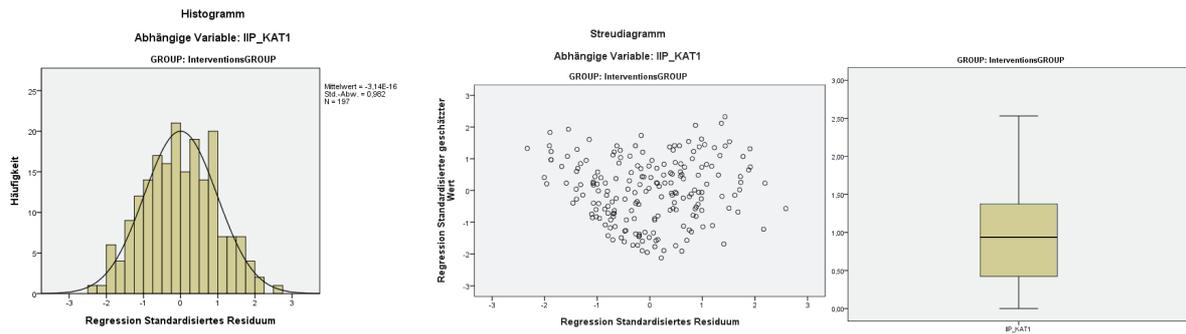
#### b) lnBSI, Katamnese



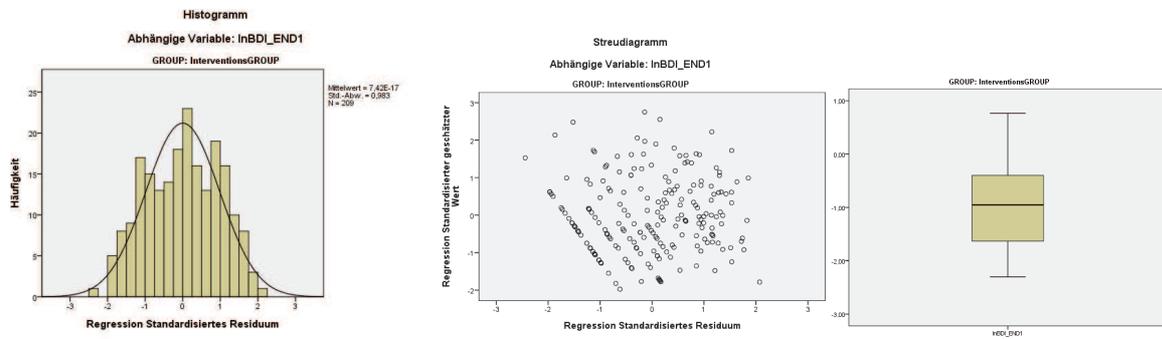
#### c) IIP, Therapieende



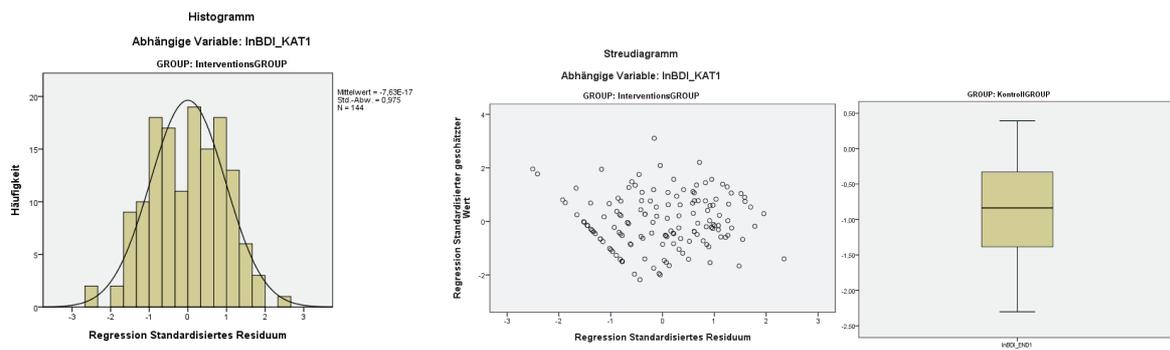
d) IIP, Katamnese



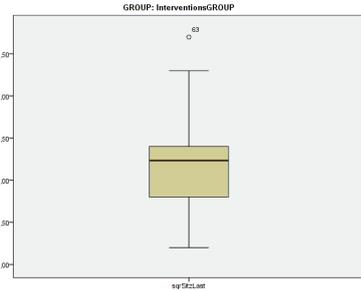
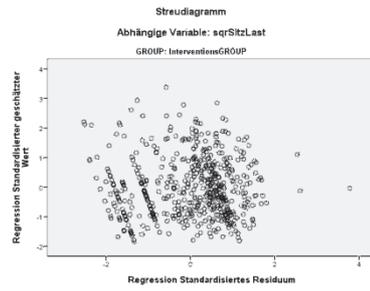
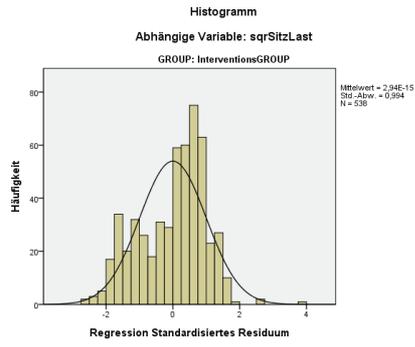
e) InBDI, Therapieende



f) InBDI, Katamnese



g) sqrDauer (Anzahl Sitzungen)



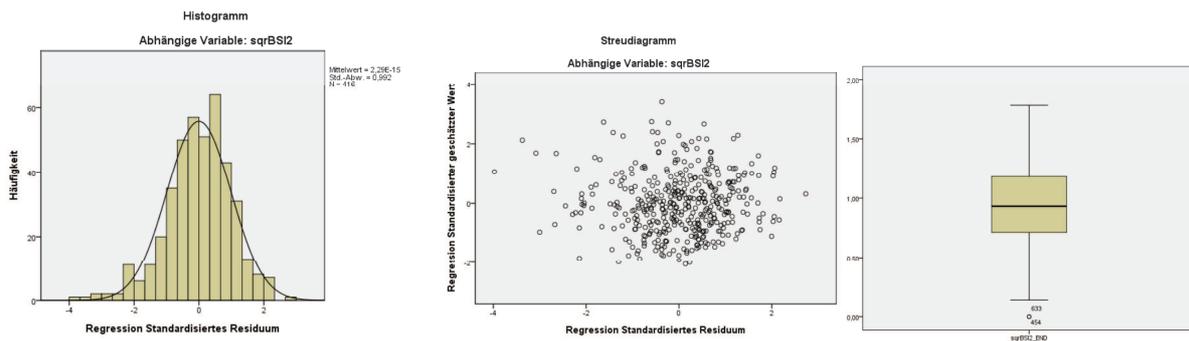
## Korrelationsmatrix, als Voraussetzung für die Regressionsanalysen zweiter Block

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 BSI End																	
2 BSI Kat	.729**																
3 IIP End	.733**	.616**															
4 IIP Kat	.595**	.785**	.729**														
5 BDI End	.806**	.641**	.644*	.559**													
6 BDI Kat	.656**	.853**	.594**	.731**	.675**												
7 Anzahl Sitzungen	.202**	.075	.166**	.096	.150*	.111											
8 Alter	.020	.048	-.005	.062	-.023	.051	-.053										
9 Geschlecht	-.010	-.007	.039	.082	-.006	.011	-.058	.035									
10 BSI-GSI Eingang	.467**	.430**	.388**	.406**	.322**	.329**	.136**	.018	.098**								
11 KOM PS	.091	.052	.117	.092	.105	.026	.111	-.025	-.054	.077							
12 KOM Angst/Affekt.	.152**	.108	.105*	.121*	.179**	.147*	.100**	-.022	.015	.106**	-.111**						
13 KOM andere	.033	.026	.042	.046	.004	-.028	.022	.053	.005	-.001	-.166**	-.198**					
14 Anzahl Diagnosen	.170**	.144*	.175**	.198**	.162**	.102	.200**	.014	-.005	.128**	.355**	.472**	.447**				
15 KOM PS * Dauer	.064	.026	.090	.076	.056	.005	.280**	-.013	-.020	.057	.814**	-.082*	-.123**	.328**			
16 KOM Angst/Affekt. * Dauer	.163**	.131*	.101*	.139*	.164**	.147*	.278**	-.016	-.010	.108**	-.080*	.809**	-.143**	.369**	-.059		
17 KOM andere * Dauer	.092	.045	.046	.014	-.001	.027	.388**	-.027	.014	.069*	-.107**	-.128**	.699**	.320**	-.079*	-.092**	

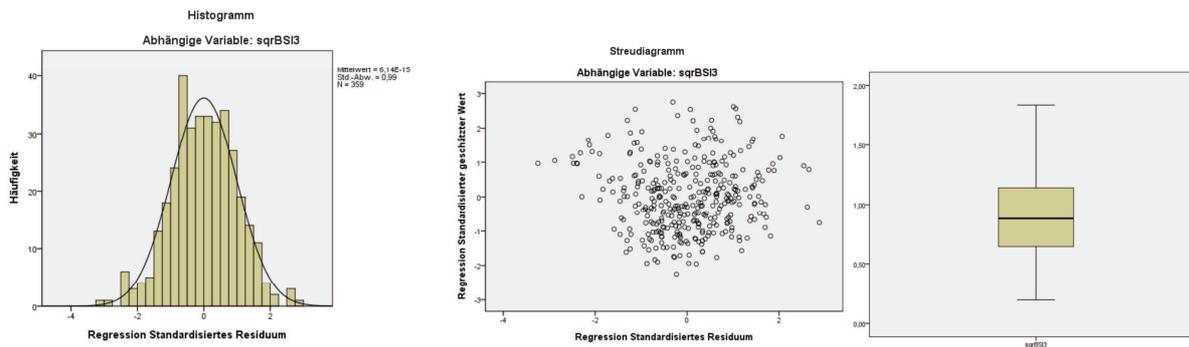
## Prüfung der Voraussetzungen für die Vorhersage der zweiten und dritten Zwischenmessung

Histogramme, Streudiagramme sowie Boxplots der in die Regressionsanalysen eingehenden abhängigen Variablen BSI, IIP und BDI (bzw. deren Transformationen) zur zweiten und dritten Zwischenerhebung (Sitzung 10 bzw. Sitzung 20).

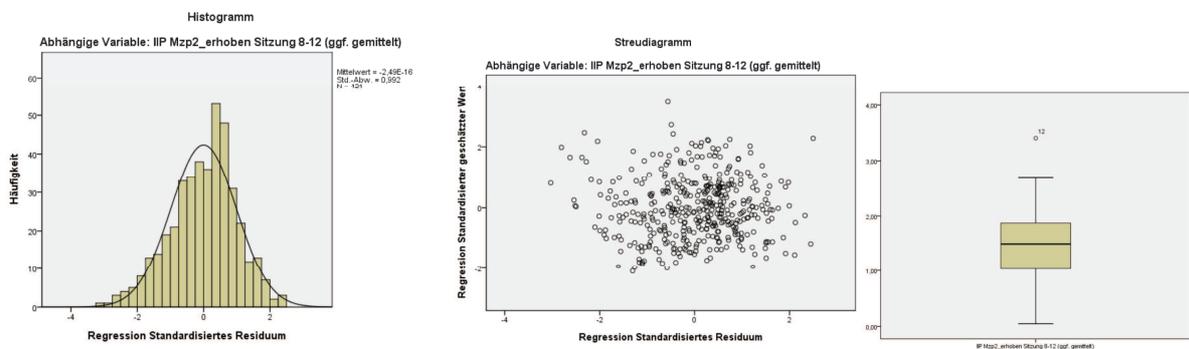
### a) sqrBSI2



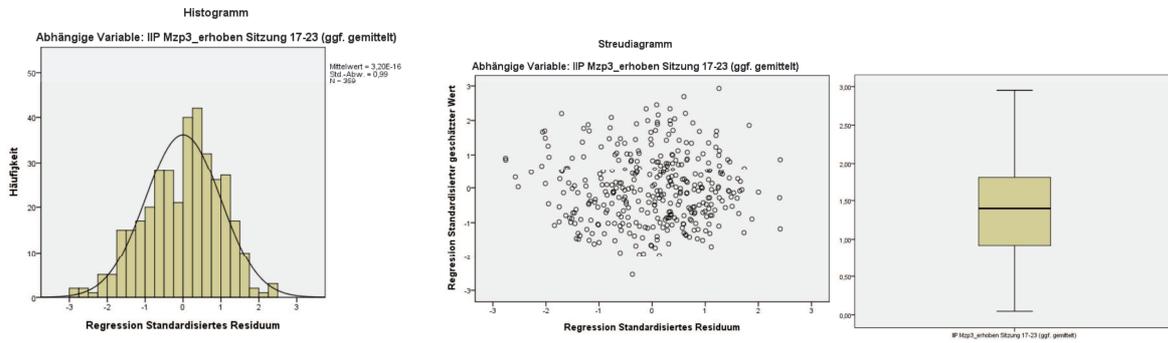
### b) sqrBSI3



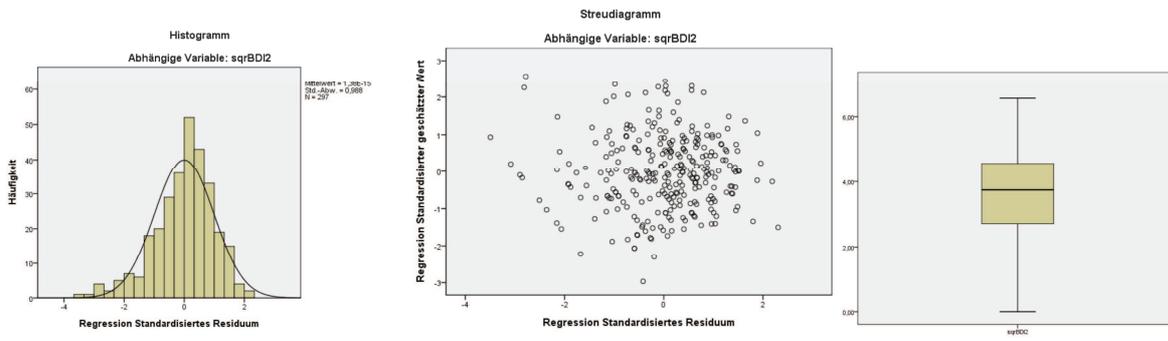
### c) IIP\_Mzp2



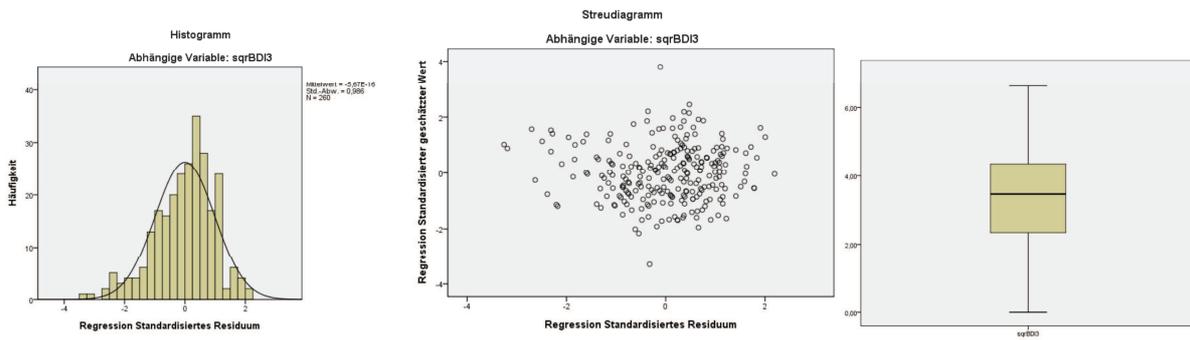
d) IIP\_Mzp3



e) Sqr\_BDI\_Mzp2



f) Sqr\_BDI\_Mzp3



**Tabelle 5: Anzahl (relativer Anteil) im BSI reliabel verbesserter Patienten in den identifizierten Klassen zum Therapieende und Katamnesezeitpunkt**

		wahrscheinlichste Klasse				Gesamt
		Klasse 1 leicht verbessert	Klasse 2 unverändert	Klasse 3 verschlechtert	Klasse 4 schnell verbessert	
Therapieende	reliable Verschlechterung	3 (2.3%)	7 (13.5%)	1 (14.3%)	0	11 (5.3%)
	keine Veränderung	21 (15.8%)	9 (17.3%)	1 (14.3%)	0	31 (15.0%)
	reliable Verbesserung	109 (82.0%)	36 (69.2%)	6 (71.4%)	14 (100%)	164 (79.6%)
	klinisch signifikante Verbesserung	23 (17.3%)	25 (48.1%)	4 (57.1%)	2 (14.3%)	54 (26.2%)
	Gesamt	133 (64.6%)	52 (25.2%)	7 (3.4%)	14 (6.8%)	206
		wahrscheinlichste Klasse				Gesamt
		Klasse 1 leicht verbessert	Klasse 2 unverändert	Klasse 3 verschlechtert	Klasse 4 schnell verbessert	
Katamnese	reliable Verschlechterung	6 (6.7%)	2 (6.3%)	2 (50.0%)	0	10 (7.2%)
	keine Veränderung	14 (15.6%)	2 (6.3%)	1 (25.0%)	0	17 (12.3%)
	reliable Verbesserung	70 (77.8%)	28 (87.5%)	1 (25.0%)	12 (100%)	111 (80.4%)
	klinisch signifikante Verbesserung	18 (20.0%)	17 (53.1%)	1 (25.0%)	3 (25.0%)	39 (28.3%)
	Gesamt	90 (65.2%)	32 (23.2%)	4 (2.9%)	12 (8.7%)	138



**Tabelle 7: Anzahl Sitzungen bis zum Therapieende nach Komorbiditätsmuster und Anzahl Diagnosen**

Komorbiditätsmuster	Anzahl Sitzungen	Anzahl Diagnosen	Anzahl Sitzungen
keine Komorbidität	N=182 M=35.87 SD=16.52	1	N=182 M=35.87 SD=16.52
Komorbidität mit Persönlichkeitsstörung	N=39 M=51.21 SD=20.79	2	N=110 M=41.91 SD=18.00
Komorbidität Angst/Affektive Störungen	N=46 M=46.67 SD=20.14	3	N=51 M=49.98 SD=25.74
Komorbidität andere	N=83 M=40.96 SD=21.09	4	N=7 M=55.00 SD=21.24

**Tabelle 8: Deskriptive Statistiken der Hauptbelastungsmaße zum Therapieende, differenziert nach Komorbiditätsgruppen**

	BSI-GSI Differenzwert	IIP Differenzwert	BDI Differenzwert	BSI-GSI Ende	IIP Ende	BDI Ende
keine Komorbidität (N=322)	N=273 M=.66 SD=.57	N=271 M=.46 SD=.50	N=204 M=.56 SD=.42	N=273 M=.49 SD=.46	N=272 M=.97 SD=.55	N=207 M=7.17 SD=6.84
Komorbidität mit Persönlichkeitsstörung (N=45)	N=40 M=.58 SD=.59	N=43 M=.47 SD=.62	N=37 M=.46 SD=.51	N=44 M=.73 SD=.63	N=44 M=1.27 SD=.58	N=38 M=11.45 SD=10.91
Komorbidität Angst/Affektive Störungen (N=62)	N=50 M=.74 SD=.67	N=50 M=.42 SD=.58	N=30 M=.56 SD=.53	N=50 M=.80 SD=.63	N=50 M=1.23 SD=.52	N=30 M=12.30 SD=7.27
Komorbidität andere (N=114)	N=101 M=.55 SD=.69	N=101 M=.41 SD=.56	N=59 M=.56 SD=.45	N=101 M=.63 SD=.63	N=101 M=1.10 SD=.64	N=60 M=8.61 SD=8.38

**Tabelle 9: Anzahl (relativer Anteil) im BSI, IIP und BDI reliabel verbesserter nach Komorbiditätsmuster zum Therapieende und Katamnesezeitpunkt**

a) BSI

		Komorbiditätsmuster				Gesamt
		keine Komorbidität	Komorbidität mit PS	Komorbidität Angst/Affektive	Komorbidität andere	
BSI Therapieende	reliable Verschlechterung	6 (2.2%)	3 (7.5%)	3 (6.0%)	9 (8.9%)	21 (4.5%)
	keine Veränderung	45 (16.5%)	7 (17.5%)	8 (16.0%)	23 (22.8%)	83 (17.9%)
	reliable Verbesserung	222 (81.3%)	30 (75.0%)	39 (78.0%)	69 (68.3%)	360 (77.6%)
	klinisch signifikante Verbesserung	153 (56.0%)	17 (40.5%)	19 (38.0%)	45 (44.6%)	234 (50.2%)
Gesamt		273	40	50	101	464

		Komorbiditätsmuster				Gesamt
		keine Komorbidität	Komorbidität mit PS	Komorbidität Angst/Affektive	Komorbidität andere	
BSI Katamnese	reliable Verschlechterung	13 (6.8%)	1 (5.0%)	1 (2.8%)	10 (16.7%)	25 (8.1%)
	keine Veränderung	30 (15.6%)	3 (15.0%)	2 (5.6%)	8 (13.3%)	43 (14.0%)
	reliable Verbesserung	149 (77.6%)	16 (80.0%)	33 (91.7%)	42 (70.0%)	240 (77.9%)
	klinisch signifikante Verbesserung	110 (57.3%)	8 (38.1%)	18 (50.0%)	27 (45.0%)	163 (52.8%)
Gesamt		192	20	36	60	308

b) IIP

		Komorbiditätsmuster				Gesamt
		keine Komorbidität	Komorbidität mit PS	Komorbidität Angst/Affektive	Komorbidität andere	
IIP Therapieende	reliable Verschlechterung	10 (3.7%)	2 (4.7%)	4 (8.0%)	6 (5.9%)	22 (4.7%)
	keine Veränderung	110 (40.6%)	21 (48.8%)	22 (44.0%)	51 (50.5%)	204 (43.9%)
	reliable Verbesserung	151 (55.7%)	20 (46.5%)	24 (48.0%)	44 (43.6%)	231 (51.4%)
	klinisch signifikante Verbesserung	139 (51.3%)	20 (46.5%)	20 (40.0%)	38 (37.6%)	217 (46.7%)
Gesamt		271	43	50	101	465

		Korbiditätsmuster				Gesamt
		keine Korbidität	Korbidität mit PS	Korbidität Angst/ Affektive	Korbidität andere	
IIP Katamnese	reliable Verschlechterung	9 (4.7%)	2 (10.0%)	4 (10.8%)	3 (5.0%)	18 (5.9%)
	keine Veränderung	79 (41.6%)	6 (30.0%)	19 (51.4%)	33 (55.0%)	137 (44.6%)
	reliable Verbesserung	102 (53.7%)	12 (60.0%)	14 (37.8%)	24 (40.0%)	152 (49.5%)
	klinisch signifikante Verbesserung	98 (51.6%)	10 (50.0%)	13 (35.1%)	24 (40.0%)	145 (47.2%)
Gesamt		190	20	37	60	307

## c) BDI

		Korbiditätsmuster				Gesamt
		keine Korbidität	Korbidität mit PS	Korbidität Angst/ Affektive	Korbidität andere	
BDI Therapieende	reliable Verschlechterung	/	/	/	/	/
	keine Veränderung	100 (36.6%)	10 (22.7%)	7 (14.0%)	31 (30.7%)	148 (31.6%)
	reliable Verbesserung	173 (63.4%)	34 (77.3%)	43 (86.0%)	70 (69.3%)	320 (68.4%)
	klinisch signifikante Verbesserung	85 (31.1%)	13 (29.5%)	16 (32.0%)	29 (28.7%)	143 (30.6%)
Gesamt		273	44	50	101	468

		Korbiditätsmuster				Gesamt
		keine Korbidität	Korbidität mit PS	Korbidität Angst/ Affektive	Korbidität andere	
BDI Katamnese	reliable Verschlechterung	/	/	/	/	/
	keine Veränderung	66 (34.2%)	5 (23.8%)	7 (18.9%)	22 (36.7%)	100 (32.2%)
	reliable Verbesserung	127 (65.8%)	16 (76.2%)	30 (81.1%)	38 (63.3%)	211 (67.8%)
	klinisch signifikante Verbesserung	64 (33.2%)	4 (19.0%)	14 (37.8%)	12 (20.0%)	94 (30.2%)
Gesamt		193	21	37	60	311

