

EVALUATION AMBULANTER PSYCHOTHERAPIEN IN EINER HOCHSCHULAMBULANZ

SUSAN RICHTER, FELICITAS GUTT, ALFONS O. HAMM

Universität Greifswald

ZUSAMMENFASSUNG: Das Fehlen einheitlicher Maßnahmen zur Sicherung der Ergebnisqualität bei ambulanten Psychotherapien macht es schwierig, die Wirksamkeit von Therapien in der Routineversorgung realistisch einzuschätzen. Erstrebenswert ist zukünftig eine QM-bezogene Evaluation mit dem Ziel der kontinuierlichen Qualitätsverbesserung, welche die Erfassung von Daten aus langjährigen internen Bewertungsprozessen sowie externen Vergleichen mit anderen Behandlern ermöglicht. Diese Arbeit evaluiert die Ergebnisse ambulanter Therapien am Zentrum für Psychologische Psychotherapie (ZPP) der Universität Greifswald (2009-2011) und vergleicht sie mit Daten aus einer vorhergehenden internen Evaluation (Richter, 2010), um festzustellen, wie wirksam die Therapien am ZPP sind, ob sich diese Ergebnisse über die Zeit bestätigen lassen und um nicht zuletzt Überlegungen anzuregen, wie Effekte ggf. noch vergrößert werden können. In einem naturalistischen Eingruppen-Design wird der Behandlungserfolg (Mittelwertveränderungen, Effektstärken, Response- und Remissionsanalysen) erfasst, wobei verschiedene Arten der Berechnung kritisch gegenübergestellt werden. Ergänzend werden Daten zur Patientenzufriedenheit analysiert. Es wird deutlich, dass bei der Einschätzung der Therapieergebnisse Berechnungen zur klinischen Signifikanz unerlässlich sind.

SCHLÜSSELWÖRTER: Psychotherapie, Hochschulambulanz, Evaluation, Ergebnisqualität, Qualitätsmanagement

EVALUATION OF PSYCHOTHERAPIES AT AN OUTPATIENT PSYCHOTHERAPY UNIT AT AN UNIVERSITY

ABSTRACT: This article presents questions regarding evaluation of psychotherapies and quality assurance at the Outpatient Psychotherapy Clinic of the Psychological Institute, University of Greifswald (Germany), providing a contribution to health services research. In order to evaluate treatment outcomes mean differences together with effect sizes as well as different kinds of response and remission rates were analyzed. In addition there are ratings of patients who evaluate their treatments. In order to develop a realistic picture of treatment quality it seems to be important to analyse clinical significance when psychotherapies are evaluated.

Further evaluations within the frame of quality management are supposed to enable the Psychotherapeutic Outpatient Clinic of the University of Greifswald to develop a realistic picture of treatment quality by keeping track of quality indices over longer periods of time. This again is supposed to serve as a source for continuous improvements.

KEYWORDS: Psychotherapy, Outpatient University Psychotherapy Clinic, Evaluation, Quality of Treatment Outcomes, Quality Management

1 EINFÜHRUNG

Ambulante Psychotherapie als Leistung der gesetzlichen Krankenkassen findet aktuell in der vertragsärztlichen Versorgung statt, deren gesetzliche Grundlage im Fünften Sozialgesetzbuch (SGB V) definiert ist. Dabei ist im § 135 a festgelegt, dass auch psychotherapeutische Praxisinhaber zur Qualitätssicherung und zum Aufbau eines internen Qualitätsmanagements verpflichtet sind (Sulz, 2007). Qualitätssicherung in der vertragsärztlichen Versorgung findet aktuell v.a. in den Bereichen der Struktur- und der Prozessqualität statt, muss aber permanent weiterentwickelt werden. Maßnahmen bzgl. der Qualitätssicherung im Bereich der Ergebnisqualität bedürfen eines gezielten Ausbaus (vgl. Schmidt-Bodenstein, 1998).

Das *Phasenmodell der Therapieevaluation* gliedert den Prozess der Psychotherapieprüfung in vier Phasen (Buchkremer & Klingberg, 2001). Dabei stellen Effectiveness-Studien Phase-IV-Forschung dar, bei der im Vergleich zu den kontrollierten Therapiestudien der Phase III keine Randomisierung vorliegt (Hiller, Schindler, Andor & Rist, 2011). Aktuell gewinnen zunehmend naturalistische Designs im Rahmen von Qualitätssicherung als Praxisfeld der Evaluationsforschung an Bedeutung (Lutz & Böhnke, 2010), insbesondere an psychotherapeutischen Hochschulambulanzen universitärer Institute. Ziele der *Effective-*

ness-Studien sind die Überprüfung der klinischen Brauchbarkeit therapeutischer Interventionen in der Alltagspraxis und der Vergleich von Effektstärken aus randomisiert-kontrollierten Therapien mit denen in der ambulanten Versorgung.

In einigen Hochschulambulanzen sind in den letzten Jahren bereits systematische, leider aber auch heterogene Evaluationssysteme zur Qualitätssicherung eingerichtet worden. Es gibt für den Bereich der psychotherapeutischen Behandlung in Ambulanzen bisher keine in den klinischen Alltag implementierten übergreifenden Qualitätssicherungsprogramme (Kawski, Koch, Lubecki & Schulz, 2005) und es fehlen Standards zur Erfassung von Struktur-, Prozess- und Ergebnisqualität.

Es ist bislang kaum möglich, realistische Einschätzungen darüber zu treffen, ob Therapieergebnisse in der Versorgung von Patienten in einer Hochschulambulanz ausreichend gut sind und ob eine Orientierung an den in Forschungstherapien erhobenen Effekten realistisch ist, da bislang Vergleichsdaten fehlen. Das Beispiel der Evaluierung der Verhaltenstherapie-Ambulanz der Technischen Universität Braunschweig zeigt, „daß kognitiv-verhaltenstherapeutische Psychotherapie sich nicht nur in randomisierten und kontrollierten Therapiestudien als wirksam erwiesen hat, sondern auch im klinischen Alltag als sehr effektiv und brauchbar zu bewerten ist“ (vgl. Schulz, Hoyer & Hahlweg, 1998, S. 154). Insgesamt bewirken die therapeutischen Inter-

ventionen in dieser Untersuchung eine deutliche Reduktion der Störungsintensität und eine hohe Patientenzufriedenheit mit der Behandlung. Schulz et al. (1998) berichten von Effektstärken zwischen 0,77 (BSQ) für Angstpatienten, 0,40 (GSI der SCL-90R) für die Gesamtstichprobe und bis zu 1,66 (BDI) in der Gruppe der depressiven Patienten. Hiller, Bleichhardt & Schindler (2009) von der psychotherapeutischen Hochschulambulanz der Universität Mainz berichten beim Vergleich von Prä- und Post-Messungen bei Therapien zwischen 2005 und 2007 von Effektstärken zwischen 0,60 und 0,92 (Intention-to-Treat) bzw. 0,98 und 1,07 (Completer) für den Gesamtwert GSI des Brief Symptom Inventory BSI. In der Gruppe der depressiven Patienten (BDI) liegen die Effektstärken für diesen Zeitraum zwischen 1,00 und 1,29 (ITT) bzw. 1,37 und 1,52 (Completer). Die Hochschulambulanz der Universität Mainz verfügt darüber hinaus seit 2005 als erste Hochschulambulanz in Deutschland über ein geprüftes Qualitätsmanagement nach DIN EN ISO 9001 (Hiller, Bleichhardt, Haaf, Legenbauer, Mauer-Matzen & Rübler, 2006). Eine QM-bezogene Evaluation ambulanter Psychotherapien eröffnet die Chance, durch langjährige Vergleiche interner Therapieergebnisse ein realistisches Bild der Therapiequalität zu erhalten. Darüber hinaus eröffnet die Entwicklung an den verschiedenen Hochschulambulanzen die Chance, über Benchmarking auch Vergleiche zwischen den Einrichtungen zu ermöglichen und dadurch wichtige Diskussionen z.B. zu realistischen Therapieeffekten im klinischen Alltag, aber auch zu sinnvollen übergreifenden Qualitätskriterien und Qualitätssicherungsmaßnahmen anzustoßen. Diese Arbeit soll einen Beitrag leisten zum internen Vergleich von Therapieergebnissen im Verlauf mehrerer Jahre, da die Daten aus dem Evaluati-

onszeitraum 2009 bis 2011 mit Daten aus einer vorhergehenden Evaluation am ZPP verglichen werden können. Das scheint insbesondere vor dem Hintergrund struktureller Besonderheiten einer universitären Einrichtung und der existierenden Mitarbeiterfluktuation v.a. unter den Therapeuten und wissenschaftlichen Hilfskräften in der Ausbildungsambulanz interessant und wichtig, um langfristig erfolgreiche Therapien durchführen und ggf. optimieren zu können. Bei Richter (2010) finden sich Effektstärken von 0,83 (BSQ) und 1,06 (MI) für Angstpatienten, 0,59 (GSI der SCL-90R) für die Gesamtstichprobe und von 0,70 (BDI) in der Gruppe der depressiven Patienten. Es finden sich sogar noch höhere Effektstärken ($d = 0,98$ im BSQ, $d = 1,14$ im MI) in der Gruppe der Patienten mit Agoraphobie und Panikstörung, die im Rahmen einer Multicenterstudie in Deutschland (Paniknetz) behandelt wurden. Diese Patienten erhielten eine Verhaltenstherapie mit dem Behandlungsschwerpunkt der Exposition in vivo in einem deutlich kürzeren Zeitraum (6 Wochen), je zweimal wöchentlich, nach einem festgelegten Manual zur Therapie (vgl. Lang, Helbig-Lang, Westphal, Gloster & Wittchen, 2012). Die Studenttherapeuten wurden besonders geschult und supervidiert. Diese Ergebnisse weisen darauf hin, dass auch „normale“ Patienten, die im Rahmen der ambulanten Versorgung nach einer Verhaltenstherapie suchen, erfolgreich in einem Studiendesign behandelt werden können. Diese Arbeit soll ferner einen Beitrag leisten zum möglichen Vergleich mit Therapieergebnissen anderer Hochschulambulanzen. Dabei werden nicht nur Daten zu indirekten und direkten Veränderungsmessungen und zur Aufrechterhaltung des Behandlungserfolgs analysiert, sondern auch zum Therapieerfolg auf Einzelfallebene. Letztere unterstützen die Bedeutsamkeit

der klinischen Signifikanz, deren Berechnung Voraussetzung ist zur realistischen Beurteilung von Therapieerfolgen in der ambulanten psychotherapeutischen Versorgung.

2 FRAGESTELLUNG

Die Ergebnisse dieser Arbeit sollen mit der zurückliegenden Evaluation der Ergebnisqualität der Institutsambulanz am Zentrum für Psychologische Psychotherapie (ZPP) der Universität Greifswald (Richter, 2010) verglichen werden. Dieser interne Vergleich sowie die methodische Ergänzung um eine Unterscheidung zwischen Intention-to-Treat- und Completer-Analysen sollen es ermöglichen, eine realistische Einschätzung der Ergebnisqualität der ambulanten Therapien in unserer Einrichtung vornehmen zu können und ggf. Hinweise zu gewinnen, wie diese langfristig zu sichern und zu optimieren ist, indem dann bei zukünftigen Evaluationen zu intern festgelegten Erwartungen im Sinne von Qualitätszielen Bezug genommen werden kann.

Es wird eine annähernd vergleichbare Ergebnisqualität erwartet (therapiebedingte statistisch signifikante Verbesserungen in der globalen Psychopathologie der Gesamtstichprobe sowie der störungsspezifischen Symptomatik der Patienten mit depressiven Störungen sowie Agoraphobie und/oder Panikstörung, mindestens mittelhohe Prä-Post-Effektstärken; hohe Patientenzufriedenheit zum Post-Zeitpunkt; Aufrechterhaltung des Therapieerfolgs bis zum Katamnese-Zeitpunkt).

Ferner soll der Anteil der statistisch und klinisch signifikant veränderten Patienten der Gesamtstichprobe sowie spezifischer Diagnosegruppen bezüglich globaler und spezifischer Maße unter-

sucht werden, da es in der ambulanten psychotherapeutischen Versorgung von hoher Bedeutung ist, ob ein einzelner Patient auf eine Behandlung anspricht. Erwartet werden Response- und Remissionsquoten von jeweils mindestens 30%. Dabei werden zwei Methoden zur Berechnung der klinischen Signifikanz eingesetzt und kritisch gegenübergestellt. Die Ergebnisse sollen Auskunft darüber geben, ob der Einsatz einer bestimmten Methode bei der Einschätzung der Ergebnisqualität von ambulanten Psychotherapien in einer Hochschulambulanz geeigneter scheint. Durch einen daraus möglich werdenden Vergleich mit anderen publizierten Daten (z.B. Hiller et al., 2009) erhoffen wir uns die Anregung einer Diskussion über sinnvolle übergreifende Maßnahmen zur Sicherung und realistischen Einschätzung der Ergebnisqualität von Psychotherapien in einer Hochschulambulanz.

3 METHODISCHES VORGEHEN

3.1 STICHPROBE

Es wurden alle Patienten in die Berechnungen mit einbezogen, die ihre Therapie am Zentrum für Psychologische Psychotherapie (ZPP) der Universität Greifswald, unabhängig von deren Beginn, zwischen Juli 2009 und Oktober 2011 beendet haben ($n = 167$). Für die statistischen Analysen wurden alle Patienten mit vollständigen Datensätzen, d.h. vorhandenen Prä- und Post-Fragebögen (acht Wochen nach der Therapie), berücksichtigt. Zusätzlich wurden, sofern vorliegend, die Katamnesefragebögen (ein Jahr nach der Therapie) in die Berechnungen mit einbezogen.

3.1.1 SOZIODEMOGRAFISCHE DATEN

Tabelle 1 beschreibt die soziodemografische Zusammensetzung der untersuchten Stichprobe.

Tabelle 1: *Soziodemografische Zusammensetzung der ITT-Stichprobe (n = 167)*

	Anzahl der Pat. (%)
Geschlecht	
männlich	26
weiblich	74
Alter (in Jahren)	
18-29	37
30-39	16
40-49	19
50-59	20
≥ 60	8
Höchster Schulabschluss	
kein Abschluss	2
Hauptschule	4
Realschule	42
Abitur	33
Hochschule	19
Berufliche Situation	
berufstätig	44
Student/Azubi	30
arbeitslos	9
Rente	12
Sonstige	5
Familienstand	
ledig	53
Partnerschaft/verheiratet	38
geschieden	5
verwitwet	4

3.1.2 DIAGNOSEN UND VERTEILUNG DER STÖRUNGSGRUPPEN

Die Verteilung der Störungsgruppen der ITT-Stichprobe (n = 167) ist in Tabelle 2 abgebildet. Dabei sind Angststörungen sowie affektive Erkrankungen am häufigsten vertreten (> 80%). Die Behandlung erfolgte in der Mehrzahl der Fälle im Rahmen einer Langzeittherapie (73%) mit durchschnittlich 40 Behandlungseinheiten. Ferner ist das Auftreten komorbider und komplizierter Störungsbilder am ZPP-Patientengut nicht ungewöhnlich. Der Anteil komorbider Patienten liegt bei über 50%, wobei knapp 24% mehr als zwei Diagnosen erhielten (Multimorbidität).

Tabelle 2: *Verteilung der Störungsgruppen der ITT-Stichprobe*

Störungsgruppe	Häufigkeit in %
Angststörungen	61,1
Affektive Störungen	27,5
Somatoforme Störungen	4,8
Persönlichkeitsstörungen	3,6
Essstörungen	1,8
Sonstige	1,2

3.1.3 QUALIFIKATION DER THERAPEUT(INN)EN UND ART DER BEHANDLUNGEN

In der Beschaffenheit der Hochschulambulanz als Ausbildungsstätte wird die Mehrzahl der ambulanten Psychotherapien von Therapeuten durchgeführt, die sich in der Ausbildung befinden. Sie weisen die folgenden Qualifikationsmerkmale auf: Diplom in Psychologie, Teilnehmer des APPVT (Ausbildungsstu-

diengang Psychologische Psychotherapie Verhaltenstherapie), regelmäßige externe und interne Einzel- sowie Gruppensupervisionen im Rahmen der Weiterbildung durch anerkannte Supervisoren, die verpflichtend ist.

Bezug nehmend auf die in der Ambulanz am häufigsten diagnostizierten psychischen Störungen ist an dieser Stelle zu erwähnen, dass die Behandlung von Depressionen mit kognitiver Verhaltenstherapie in Anlehnung an Hautzinger (2010) erfolgt. Die Therapien zur Behandlung von Angststörungen erfolgen v.a. expositionsbasiert, zur Behandlung von Panikstörung mit Agoraphobie z.B. nach Lang et al. (2012).

3.2 DESIGN

In einem naturalistischen Eingruppen-Design wird der *Behandlungserfolg* mit Hilfe von Mittelwertsveränderungen im Therapieverlauf, Effektstärken sowie Response- und Remissionsanalysen in unterschiedlichen Störungsgruppen untersucht. Dies erfolgt unter Einbezug von drei Messzeitpunkten (vor Therapiebeginn, 8 Wochen sowie ein Jahr nach Therapieende). Der Einsatz der verschiedenen Messinstrumente erfolgt nach Literaturrecherchen und in Anlehnung an deren Einsatz auch in anderen Hochschulambulanzen.

3.2.1 MESSINSTRUMENTE

Zur Erfassung der allgemeinen Psychopathologie setzen wir die Symptom-Checkliste von Derogatis in ihrer deutschen Version von Franke (1995) ein. Die Auswahl der störungsspezifischen Fragebögen erfolgt anhand der gestellten ICD-10-Diagnosen in der Eingangsdiagnostik. Es handelt sich hierbei z.B. um

die deutsche Version des Beck-Depressions-Inventars (BDI; Hautzinger, Bailer, Worall & Keller, 1994) und den Fragebogen zu körperbezogenen Ängsten, Kognitionen und Vermeidung in seiner deutschen Version (AKV; Ehlers, Margraf & Chambless, 2001). Zur Erfassung der Patientenzufriedenheit setzen wir (nach Literaturrecherchen und nach dem Kriterium der Ökonomie) ein eigenständig entwickeltes Messinstrument ein.

3.2.2 DURCHFÜHRUNG

Die Datenerhebung dieser Untersuchung fand im Rahmen der kontinuierlichen Dokumentations- und Qualitätssicherungsprozesse in der Institutsambulanz des Zentrums für Psychologische Psychotherapie (ZPP) statt. Die Hochschulambulanz verfügt über eine standardisierte Eingangs-, Verlaufs- und Ergebnisdiagnostik, die eine systematische Behandlung und Evaluation erlaubt. Dabei erfolgt die klassifikatorische Eingangsdiagnostik im Rahmen der probatorischen Sitzungen nach den Kriterien der Internationalen Klassifikation psychischer Störungen (ICD-10, Kapitel V (F); Dilling, Mombour & Schmidt, 2011) mit Hilfe eines klinischen Interviews (Diagnostisches Expertensystem für psychische Störungen, DIA-X; Wittchen & Pfister, 1997). Dieses wird nur zu Beginn der Therapie durchgeführt. Alle anderen Instrumente, abgesehen von dem nur zum Post-Zeitpunkt ausgehändigten Fragebogen zur Patientenzufriedenheit, werden zum Prä-, Post- und Katamnese-Zeitpunkt eingesetzt.

3.3 STATISTISCHE AUSWERTUNG

3.3.1 ALLGEMEINE BERECHNUNGEN ZUM THERAPIEERFOLG

Die Berechnung von Mittelwertsdifferenzen und Effektstärken in dieser Untersuchung erfolgte mittels SPSS 20.0 für Windows. Bei allen statistischen Analysen wurde ein Signifikanzniveau von 5% vorausgesetzt. Es wurden Tests auf statistische Unterschiede auf Gruppenebene, t-Tests bei verbundenen Stichproben, zwischen den Messzeitpunkten Prä, Post und Post-Katamnese durchgeführt, um zu untersuchen, ob sich die jeweiligen Mittelwerte der verschiedenen Maße/Skalen statistisch signifikant unterscheiden (vgl. z.B. Bortz, 2005). Aufgrund geringer Stichprobengrößen und nicht normalverteilter Daten (Differenzen) wurde zum Post-Katamnese-Mittelwertvergleich anstatt des t-Tests teilweise der nichtparametrische Wilcoxon-Test durchgeführt. Zur Bewertung der Größe der statistischen Differenzen wurden zusätzlich Effektstärken berechnet. Es werden zum Vergleich zwei unterschiedliche Effektstärkenvarianten (ES_{diff} , ES_{pool}) berechnet (vgl. Maier-Riehle & Zwingmann, 2000). Zur Bewertung des Therapieerfolgs dienen einerseits die Grenzwerte von Cohen (1988), andererseits wurden vergleichbare Evaluationsuntersuchungen zur Bewertung der Größe der Effekte herangezogen.

3.3.2 BERECHNUNGEN ZUR KLINISCHEN SIGNIFIKANZ

Zur Analyse von Therapieergebnissen auf Einzelfallebene wurden in dieser Untersuchung zusätzlich klinische Signifikanzen berechnet. Die Berechnung erfolgte einerseits über den *Reliable Change Index* (RCI) nach Jacobson und

Truax (1991), andererseits über die *Methode der prozentualen Wertereduktion* in Anlehnung an Hiller et al. (2009). Zur Berechnung der jeweiligen Cut-Offs der eingesetzten Messinstrumente wurden die Mittelwerte (M) und Standardabweichungen (SD) der gesunden (norm) sowie klinischen (klin) Normstichproben den jeweiligen Manualen (Franke, 1995) bzw. Neunormierungen (Ehlers et al., 2001; Seggar, Lambert & Hansen, 2002) entnommen.

4

ERGEBNISSE

Im Folgenden werden die Ergebnisse in Bezug auf die wichtigsten Fragen und Ziele dieser Untersuchung (siehe Abschnitt 2) dargestellt.

4.1 PRÄ-POST-ANALYSEN

Die allgemeine Psychopathologie der Patienten verringert sich im Therapieverlauf. Die Mittelwertsdifferenz des GSI-Gesamtwertes der SCL-90-R ist statistisch signifikant. Beide Effektstärkenvarianten ES_{diff} und ES_{pool} liegen vergleichbar im mittleren Bereich (vgl. Konventionen nach Cohen).

Auch bei den Patienten mit depressiven Störungen verringert sich die Symptomatik deutlich. Es zeigen sich sowohl in der Completer- als auch der ITT-Stichprobe signifikante Prä-Post-Mittelwertsdifferenzen ($p < 0,01$) sowie mittelgroße Effektstärken von $ES_{diff} = 0,63$ und $ES_{pool} = 0,56$ (ITT) bzw. $ES_{diff} = 0,69$ und $ES_{pool} = 0,65$ (Completer). Die Effektstärkevariante ES_{diff} fällt hier vermutlich aufgrund einer deutlichen Erhöhung der Post-Streuung (SD_{post}) im Vergleich zur Prä-Streuung ($SD_{prä}$) für die ITT- und Completer-Stichprobe größer aus als die Variante ES_{pool} .

Ferner berichten auch die Patienten mit den Diagnosen Agoraphobie und/oder Panikstörung nach der Therapie von einer deutlich geringeren Angstsymptomatik als vor Therapiebeginn. Alle Prä-Post-Gruppenvergleiche sind statistisch signifikant ($p < 0,01$). Die Effektgrößen beider Effektgrößenvarianten liegen bzgl. aller Skalen im mittleren bis hohen Bereich, wobei das MI-A bei der Variante ES_{diff} mit $ES_{diff} = 0,93$ (ITT) bzw. $ES_{diff} = 0,99$ (Completer) die höchsten Effektstärken verzeichnet. Bei Anwendung der Effektstärkeberechnung ES_{pool} weist der BSQ die höchsten Effektstärken mit $ES_{pool} = 0,84$ (ITT) bzw. $ES_{pool} = 0,92$ (Completer) auf. Insgesamt überschreiten alle Effektstärken der Completer-Ana-

lysen diejenigen der ITT-Analysen. Betrachtet man beide Effektstärkevarianten im Vergleich, ergeben sich meist unterschiedliche ES-Werte innerhalb derselben Skalen, jedoch kein systematischer Trend, d.h. dass die Anwendung der einen oder der anderen Effektstärkevariante zu keinem durchgängig besseren oder schlechteren Ergebnis führt, sondern skalenabhängig variiert. Die Ergebnisse sind in der Tabelle 3 zusammengefasst.

4.2 PATIENTENZUFRIEDENHEIT

Bezüglich der Untersuchung der Patientenzufriedenheit mit dem Behandlungsergebnis wurden ausschließlich die Da-

Tabelle 3: Statistische Signifikanzen der Mittelwertsdifferenzen und berechnete Effektstärken der Prä-Post-Vergleiche in der Gesamtstichprobe (GSI der SCL-90-R), von Patienten mit depressiven Störungen (BDI) sowie von Patienten mit Agoraphobie und/oder Panikstörung (BSQ, ACQ, MI-A, BDI)

	M (SD) prä	M (SD) post	p-Wert	ES_{diff}	ES_{pool}
GSI (SCL-90-R)					
ITT (n = 153)	1,08 (0,62)	0,67 (0,66)	< 0,01*	0,69	0,64
Completer (n = 140)	1,06 (0,60)	0,61 (0,60)	< 0,01*	0,77	0,74
BDI					
ITT (n = 44)	19,27 (9,30)	13,29 (11,85)	< 0,01*	0,63	0,56
Completer (n = 39)	18,26 (8,37)	11,92 (11,00)	< 0,01*	0,69	0,65
BSQ					
ITT (n = 43)	2,69 (0,77)	2,06 (0,73)	< 0,01*	0,72	0,84
Completer (n = 39)	2,65 (0,79)	1,99 (0,64)	< 0,01*	0,75	0,92
ACQ					
ITT (n = 44)	2,26 (0,69)	1,73 (0,70)	< 0,01*	0,73	0,76
Completer (n = 40)	2,27 (0,70)	1,70 (0,65)	< 0,01*	0,80	0,85
MI-A					
ITT (n = 42)	2,74 (1,12)	1,90 (1,11)	< 0,01*	0,93	0,76
Completer (n = 39)	2,70 (1,06)	1,80 (1,02)	< 0,01*	0,99	0,85
BDI (Angstpat.)					
ITT (n = 44)	17,11 (12,43)	9,70 (11,20)	< 0,01*	0,83	0,63
Completer (n = 40)	16,10 (10,51)	8,10 (8,87)	< 0,01*	0,88	0,82

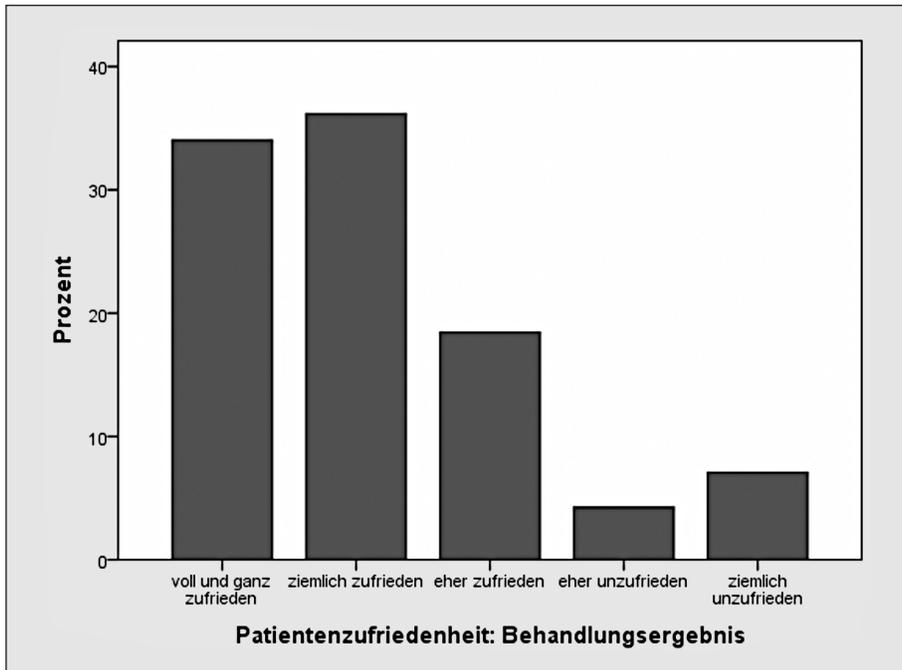


Abbildung 1: Patientenzufriedenheit mit dem Behandlungsergebnis (n = 141)

ten der Completer-Stichprobe (n = 152) betrachtet. Aufgrund von elf fehlenden Angaben wurden n = 141 Patienten analysiert. Die Ergebnisse sprechen für eine hohe Patientenzufriedenheit und sind in Abbildung 1 dargestellt.

4.3 AUFRECHTERHALTUNG DES THERAPIEERFOLGS

Zur Evaluation der Aufrechterhaltung des Prä-Post-Therapieerfolgs wurden Post-Katamnese-Mittelwertsdifferenzen sowie Effektgrößen der Completer-Stichprobe berechnet. Dies erfolgte sowohl für die Gesamtstichprobe als auch für spezifische Störungsgruppen. In allen Erhebungsmaßen wird die kontinuierliche Reduktion der Symptomatik über die drei Messzeitpunkte (Prä-, Post- und Jahreskatamnese) ersichtlich und im Sinne eines

stabilen Therapieerfolgs interpretiert. Die Abbildung 2 zeigt beispielhaft den Verlauf der Mittelwerte des GSI über die Zeit.

4.4 RESPONSE- UND REMISSIONSANALYSEN

Es wurden die klinischen Signifikanzen der Completer-Stichprobe nach der RCI-Methode sowie nach der Methode der prozentualen Wertereduktion sowohl für die Gesamtstichprobe als auch für spezifische Störungsgruppen berechnet. In Abbildung 3 wird exemplarisch die Verteilung der statistischen/reliablen Veränderungen und klinischen Signifikanzen nach der RCI-Methode dargestellt. Die roten Linien der beiden Achsen stellen die jeweiligen Cut-Off-Werte der Prä- und Post-Skala dar, anhand derer sich

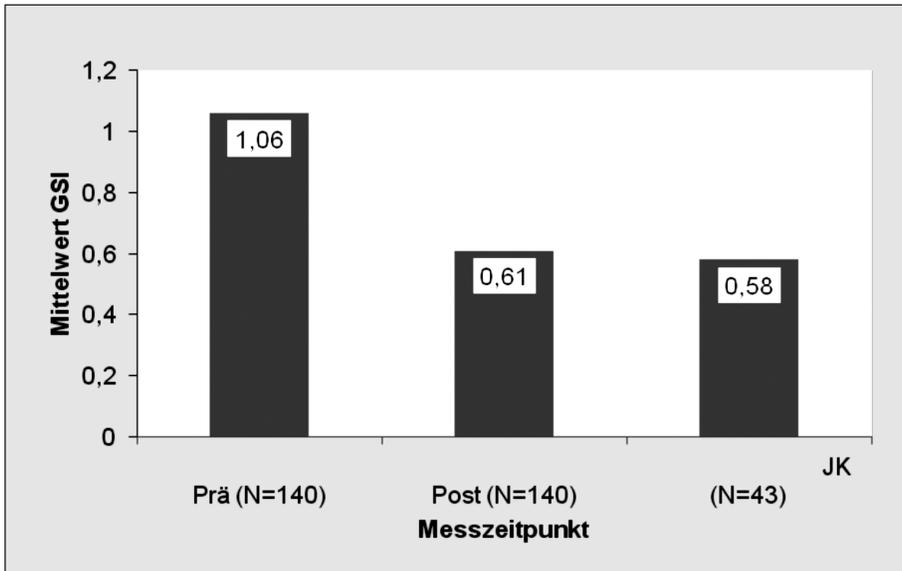


Abbildung 2: Zeit-Verlauf der Mittelwerte des GSI (Global Severity Index) der SCL-90-R der Gesamtstichprobe (Completer)

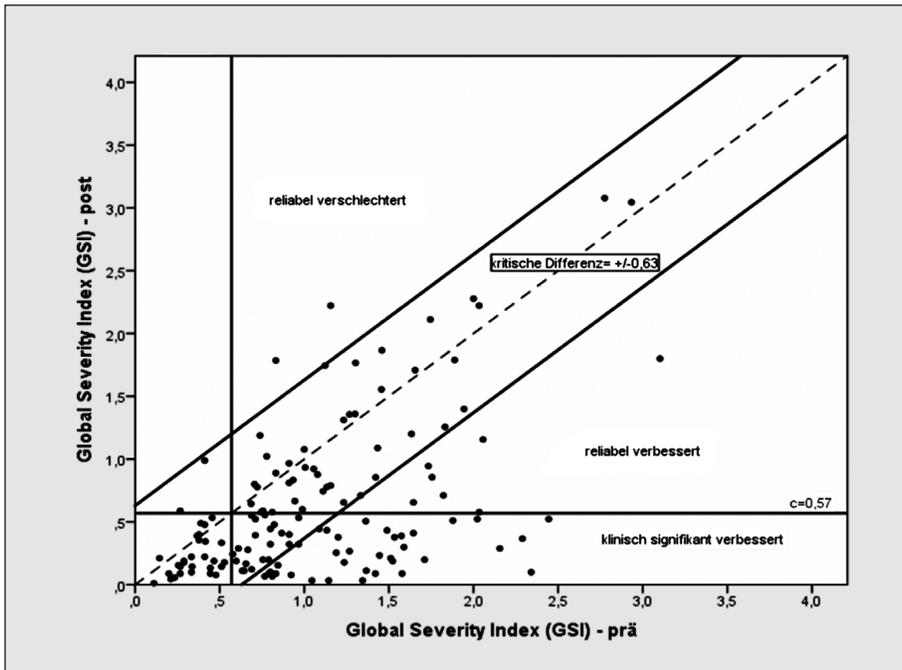


Abbildung 3: Statistische und klinische Signifikanzen der Gesamtstichprobe (Completer) vom Prä- zum Post-Zeitpunkt bzgl. des GSI gemäß dem RCI und Trennpunkt c

die einzelnen Patienten den Kategorien vorher krank – vorher gesund sowie nachher krank – nachher gesund zuteilen lassen. Die beiden schwarzen diagonal verlaufenden Linien umreißen den Bereich der kritischen Differenz, innerhalb dessen die individuellen Prä-Post-Veränderungen als nicht reliabel gemäß des RCI gelten. Jenseits dieses kritischen Bereichs liegen die Bereiche der reliab-

len Verbesserung oder Verschlechterung. Ein Patient hat sich bezüglich einer Skala klinisch signifikant verbessert, wenn er außerhalb der kritischen Differenz liegt, einen Prä-Wert oberhalb des Cut-Offs c und einen Post-Wert unterhalb dieses Cut-Offs aufweist.

Tabelle 4 zeigt die prozentuale Häufigkeit der Veränderungskategorien der Gesamtstichprobe bzgl. des GSI, der

Tabelle 4: Prozentuale Häufigkeit der Veränderungskategorien der Gesamtstichprobe (GSI), von depressiven Patienten (BDI) und Angstpatienten (BSQ, ACQ, MI-A) vom Prä- zum Post-Zeitpunkt gemäß dem RCI und Trennpunkt c

	reliabel verschlechtert	test-normal*	unverändert**	reliabel verbessert***	klinisch signifikant verbessert
GSI					
Completer (n = 140)	1,42%	19,28%	47,85%	31,42%	26,42%
Completer (Path. A.) (n = 110)	1,81%	–	58,18%	40,00%	33,63%
BDI					
Completer (n = 39)	7,69%	33,33%	33,33%	25,64%	23,07%
Completer (Path. A.) (n = 26)	11,53%	–	50,00%	38,46%	34,61%
BSQ					
Completer (n = 39)	7,69%	15,38%	30,76%	46,15%	28,20%
Completer (Path. A.) (n = 28)	–	–	35,71%	64,28%	39,28%
ACQ					
Completer (n = 40)	–	7,50%	55,00%	37,50%	25,00%
Completer (Path. A.) (n = 32)	–	–	53,12%	46,87%	31,25%
MI-A					
Completer (n = 39)	5,12%	25,64%	15,38%	53,84%	38,46%
Completer (Path. A.) (n = 29)	6,89%	–	20,68%	72,41%	51,72%

* Werte der Patienten lagen zum Prä- und Postzeitpunkt im gesunden Bereich · ** nicht reliabel verbessert oder verschlechtert · *** gemäß dem RCI

Gruppe der depressiven Patienten bzgl. des BDI sowie der Patienten mit Agoraphobie und/oder Panikstörung bzgl. BSQ, ACQ und MI-A. Dabei werden alle Berechnungen jeweils sowohl für die gesamte Completer-Stichprobe als auch für die Teilstichprobe der Completer-Patienten mit pathologischem Ausgangswert (Path. A.) durchgeführt. In der Kategorie „reliabel verbessert“ wird der Gesamtprozentsatz aller Patienten mit reliabler Verbesserung aufgeführt. Es zeigt sich, dass die Response- und Remissionsquoten bei ausschließlicher Berücksichtigung von Patienten mit pathologischem Ausgangswert durchweg höher ausfallen als bei Einbezug aller Completer-Patienten. Dies ist auf eine Verschiebung der prozentualen Verhältnisse zurückzuführen, die entsteht, wenn alle Patienten

mit nicht-klinischem Prä-Wert aus den Berechnungen entfernt werden. Diese Quoten liegen ferner alle über dem erwarteten Wert von 30%.

Tabelle 5 zeigt die Response- und Remissionsquoten derselben Stichproben und Fragebogenverfahren gemäß der Methode der prozentualen Wertereduktion. Auch hier zeigt sich, dass die Response- und Remissionsquoten bei ausschließlicher Berücksichtigung von Patienten mit pathologischem Ausgangswert durchweg höher ausfallen als bei Einbezug aller Completer-Patienten. Darüber hinaus fallen die Response- und Remissionsquoten (mit Ausnahme des MI-A) bei dieser Methode zur Berechnung klinischer Signifikanzen deutlich höher aus als bei den Berechnungen gemäß dem RCI.

Tabelle 5: Response- und Remissionsquoten der Gesamtstichprobe (GSI), von depressiven Patienten (BDI) und von Angstpatienten (BSQ, ACQ, MI-A) vom Prä- zum Post-Zeitpunkt gemäß der Methode der prozentualen Wertereduktion

	Response*	Remission**
GSI alle Ausgangswerte (n = 140) pathology. Prä-Wert (n = 110)	52,14% 66,36%	40,00% 50,90%
BDI alle Ausgangswerte (n = 39) pathology. Prä-Wert (n = 26)	41,02% 61,53%	38,46% 57,69%
BSQ alle Ausgangswerte (n = 39) pathology. Prä-Wert (n = 28)	51,28% 71,42%	33,33% 46,42%
ACQ alle Ausgangswerte (n = 40) pathology. Prä-Wert (n = 32)	50,00% 62,50%	37,50% 46,87%
MI-A alle Ausgangswerte (n = 39) pathology. Prä-Wert (n = 29)	51,28% 68,96%	41,02% 55,17%

* mindestens 50%ige Wertereduktion im pathologischen Skalenbereich sowie mindestens 25%ige Wertereduktion auf der Gesamtskala (Responsekriterium)

** Erfüllung des Responsekriteriums sowie Unterschreitung des Cut-Offs c

4.5 INTERNER VERGLEICH ZUR ERGEBNISQUALITÄT

4.5.1 VERGLEICH DES THERAPIEERFOLGS

Nachfolgend werden die Prä-Post-Effektgrößen der jeweiligen Completer-Stichproben der zu vergleichenden Evaluationszeiträume 2004 bis 2009 (Richter, 2010) und 2009 bis 2011 bzgl. der globalen Psychopathologie sowie störungsspezifischer Symptomatik gegenübergestellt (siehe Tabelle 6). Es findet sich eine annähernd vergleichbare Ergebnisqualität über die Zeit.

Im Zeitraum 2004-2009 liegt die durchschnittliche Effektstärke aller SCL-90-R-Skalen plus Global Severity Index

bei $d = 0,48$, im Zeitraum 2009-2011 bei durchschnittlich $d = 0,61$. Für die Patientengruppe mit Agoraphobie und/oder Panikstörung ergibt sich im Zeitraum 2004-2009 eine durchschnittliche Effektstärke von $d = 0,86$, im Zeitraum von 2009-2011 liegt die durchschnittliche Effektstärke über die drei Skalen (BSQ, ACQ, MI-A) hinweg bei $d = 0,87$. Bei den Patienten mit depressiven Störungen wird bzgl. des BDI eine etwas bessere Effektstärke im Zeitraum 2004-2009 verzeichnet ($d = 0,70$). Alle Effektstärken liegen im mittleren bis hohen Bereich. Alle Prä-Post-Mittelwertsdifferenzen sind in beiden Zeiträumen statistisch signifikant ($p < 0,01$).

4.5.2 VERGLEICH DER KLINISCHEN SIGNIFIKANZEN

Tabelle 6: Vergleich der Effektstärken (ES_{pool}) der Completer-Gesamtstichproben bzgl. der globalen Psychopathologie (SCL-90-R), der Completer-Stichproben von Patienten mit depressiven Störungen bzgl. des BDI sowie der Completer-Stichproben von Patienten mit Agoraphobie und/oder Panikstörung bzgl. BSQ, ACQ und MI-A von jeweils 2004-2009 und 2009-2011

	ES_{pool}	
	(2004-09)	(2009-11)
SCL-90-R: GSI	0,59 (n = 137)	0,74 (n = 138)
BDI-Gesamtwert	0,70 (n = 33)	0,65 (n = 39)
BSQ	0,83 (n = 63)	0,92 (n = 39)
ACQ	0,70 (n = 66)	0,85 (n = 40)
MI-A	1,06 (n = 58)	0,85 (n = 39)

In Tabelle 7 ist der Vergleich der Ergebnisse zu den Berechnungen der klinischen Signifikanzen verschiedener Stichproben und Messverfahren dargestellt. Es werden die Ergebnisse nach der RCI-Methode zu beiden Zeiträumen (2004-2009 und 2009-2011) sowie nach der Methode der prozentualen Wertereduktion für den Zeitraum 2009-2011 gegenübergestellt. Vergleicht man die Prozentsätze der jeweiligen Teilstichproben mit pathologischem Prä-Wert in beiden Zeiträumen, ergeben sich im Zeitraum 2009-2011 fast durchgängig höhere Response- und Remissionsquoten (nach RCI-Methode) als im Zeitraum 2004-2009. Des Weiteren führte die Anwendung der Prozentmethode zu fast durchgängig besseren Response- und Remissionsquoten als die Berechnung des RCI (durchschnittliche Responsequote von knapp 70%).

Tabelle 7: Übersicht der Response- und Remissionsquoten verschiedener Messinstrumente nach dem Reliable Change Index (RCI) und der Methode der prozentualen Wertereduktion im Zeitvergleich

	reliabel verbessert (RCI)	klinisch signifikant verbessert (RCI)	Response nach %-Methode	Remission nach %-Methode
GSI* 2004-2009 (c = 0,58) 2009-2011 (c = 0,57)	31,00% 40,00%*	23,90% 33,63%*	– 66,36%*	– 50,90%*
BSQ* 2004-2009 (c = 2,08) 2009-2011 (c = 2,10)	34,00% 64,28%*	28,00% 39,28%*	– 71,42%*	– 46,42%*
ACQ* 2004-2009 (c = 1,59) 2009-2011 (c = 1,61)	30,90% 46,87%*	21,80% 31,25%*	– 62,50%*	– 46,87%*
MI-A* 2004-2009 (c = 1,82) 2009-2011 (c = 1,93)	73,30% 72,41%*	60,00% 51,72%*	– 68,96%*	– 55,17%*

* alle Stichproben mit pathologischem Prä-Wert

5 DISKUSSION

Im Folgenden sollen die Ergebnisse dieser Arbeit und einige kritische Anmerkungen diskutiert werden.

5.1 PRÄ-POST-ANALYSEN UND INTERNER VERGLEICH VON EVALUATIONSDATEN

Bezüglich der globalen Psychopathologie der Gesamtstichproben (ITT und Completer) zeigte sich eine Verbesserung der Gesamtsymptomatik vom Prä- zum Post-Zeitpunkt (statistisch signifikante Mittelwertsdifferenz ($p < 0,01$) für den GSI-Gesamtwert der SCL-90-R). Es wurden zwei Effektstärkevarianten berechnet, die zu vergleichbaren Ergebnissen im mittleren Bereich führen. Die Betrachtung der ES_{pool} erfolgte aus zweierlei Gründen. Ei-

nerseits gibt es Interpretationsschwierigkeiten und eine Korrelationsabhängigkeit der ES_{diff} (hohe Korrelationen der beiden Messwertreihen führen zu höheren ES_{diff} -Effektstärken; vgl. Maier-Riehle & Zwingmann, 2000). Andererseits existieren bereits publizierte Daten von Therapien in der ambulanten Versorgung, die Mittelwertsdifferenzen in Relation zu gepoolten Standardabweichung beider Messzeitpunkte angeben (Hiller et al., 2009). Es gilt die Frage, ob dies nicht auch bei zukünftigen Berechnungen sinnvoll erscheint, v.a. vor dem Hintergrund noch fehlender methodischer Standards für die Analyse von ambulanten Psychotherapien. Die ITT-Stichprobe bzgl. des GSI-Gesamtwertes weist eine niedrigere Effektstärke auf als die Completer-Stichprobe, sodass die ausschließliche Betrachtung der Completer-Stichprobe zu optimistischeren Einschätzungen der Ergebnisqualität führt. Da allerdings Prä-Post-Effekt-

stärken (Eingruppen-Design) verglichen mit Effektstärken aus kontrollierten Therapiestudien (Zweigruppen-Design) – aufgrund der Nichtberücksichtigung von Veränderungen der Kontrollgruppe und des Phänomens der Regression zur Mitte – meist höher ausfallen (vgl. Maier-Riehle & Zwingmann, 2000), sollte eine Prä-Post-Effektstärke ggf. zurückhaltend bewertet bzw. nach unten korrigiert werden. Im Vergleich fällt die Effektstärke des GSI-Wertes der Completer-Stichprobe 2004-2009 niedriger aus als 2009-2011. Dies könnte auf eine Verbesserung des Therapieerfolgs, aber auch auf eine abweichende Verteilung der verschiedenen Diagnosegruppen und den Einbezug eines größeren Evaluationszeitraums zurückzuführen sein. Letzteres könnte zukünftig ausgeschlossen werden, wenn in den internen Evaluationsansatz spezielle Rahmenbedingungen des Qualitätsmanagements mit berücksichtigt werden würden, z.B. die kontinuierliche (meist jährliche) Berechnung von Effektstärken und Response- sowie Remissionsanalysen nach vergleichbaren Standards. Interne Ergebnisse zur SCL-90-R können bislang nicht mit publizierten Daten verglichen werden, da dort meist auf den GSI-Wert des Brief Symptom Inventory BSI, der Kurzform der SCL-90-Skala, eingegangen wird, die wir inzwischen auch zur Erhebung der allgemeinen Psychopathologie unserer Patienten eingeführt haben, sodass zukünftig ein Vergleich auch im Sinne eines Benchmarking möglich wird.

Eine Besserung der Symptomatik findet sich auch bei den Patienten spezifischer klinischer Gruppen. Die Prä-Post-Effektstärken in der Gruppe der depressiven Patienten liegen im mittleren Bereich und sind vergleichbar für beide Evaluationszeiträume. Die Completer-Stichprobe erreicht erwartungsgemäß höhere Effektstärken als die ITT-Stich-

probe. Patienten mit Agoraphobie und/oder Panikstörung weisen bzgl. der störungsspezifischen Maße (BSQ, ACQ, MI) sowie bzgl. des BDI statistisch signifikante Prä-Post-Mittelwertsdifferenzen ($p < 0,01$) auf. Letzteres zeigt, dass auch Panikpatienten mit verringerter Depressivität (BDI) auf die Therapie ansprechen. Die Effektstärken liegen insgesamt und vergleichbar für beide Evaluationszeiträume im mittleren bis hohen Bereich (Completer $>$ ITT). Auch für die Evaluation der spezifischen Störungsgruppen gelten die oben erwähnten Interpretationseinschränkungen bzgl. der Höhe der errechneten Effektstärken. Dennoch zeigen sich bei Anwendung störungsspezifischer Fragebögen deutlich höhere durchschnittliche Effektstärken als bei der Verwendung des Breitbandinstruments SCL-90-R. Dies weist auf einen guten Therapieerfolg v.a. hinsichtlich störungsspezifischer Symptomatik hin. Zu bedenken ist, dass die ITT-Stichprobe in dieser Arbeit (zur Vermeidung von zu pessimistischen Urteilen bei Einsatz des Prä-Wertes als letzten verfügbaren Messwert der Abbrecher) nur diejenigen Therapieabbrecher umfasst, die über vollständige Post-Fragebögen verfügen. Somit ist die ITT-Stichprobe nur geringfügig größer als die Completer-Stichprobe und die Ergebnisse der ITT-Untersuchung vermutlich in positiver Richtung verzerrt. Dies mag jedoch auch die Bedeutsamkeit einer genaueren Erfassung und Analyse von Daten der Therapieabbrecher verdeutlichen.

Ferner hat die Analyse der erhobenen Daten gezeigt, dass das diagnostische Spektrum der Ambulanz recht groß ist, wenngleich Angststörungen sowie depressive Erkrankungen in der Ambulanz die am häufigsten diagnostizierten Störungsbilder sind. Auch Komorbidität ist eher die Regel. Der besseren Versorgung von Angstpatienten (mit Panikstörung

und Agoraphobie) in der Region trägt die Ambulanz mittlerweile seit Frühjahr 2007 erfolgreich Rechnung, indem sie durch die Teilnahme an einer Multicenterstudie in Deutschland eine für diese Patientengruppe zugeschnittene und nach Manual durchgeführte Verhaltenstherapie mit dem Behandlungsschwerpunkt der Exposition in vivo anbieten kann.

So könnten Unterschiede in den Effektivitätsstärken zwischen einzelnen Diagnosegruppen auch einen Zusammenhang zur Kompetenz der einzelnen Therapeuten aufweisen. Die Schwerpunktbehandlung von Angststörungen führt möglicherweise dazu, in diesem Störungsbereich bessere Therapieangebote machen zu können, die mitunter auch schneller angeboten werden können. Im Allgemeinen warten Patienten am ZPP 9-12 Monate auf einen Therapieplatz. Panikpatienten hingegen können oft nach nur wenigen Wochen Wartezeit einem Therapeuten zugeordnet werden. Es stellt sich hier die Frage, ob ähnliche spezifische Behandlungsformen nicht auch für andere Patientengruppen angeboten werden könnten (z.B. für den Bereich der Prüfungsangst), um die Versorgungslage mit ambulanter Psychotherapie in der Region weiter zu verbessern.

Auch mag die Stichprobe im Vergleich zu niedergelassenen Praxen relativ jung und hochgebildet erscheinen. Der Eindruck aus den Erstkontakten zur Ambulanz spricht jedoch dafür, dass dies auch das Klientel ist, das sich an die Kollegen in der Niederlassung wendet, zumindest in der Stadt Greifswald als Universitätsstandort. In umliegenden, eher ländlichen Regionen mag das anders sein. Dies ist bei der Generalisierbarkeit der Ergebnisse ggf. zu berücksichtigen. Nähere Untersuchungen stehen noch aus.

Für zukünftige Evaluationen wären darüber hinaus die genauere Erfassung von Parametern zur Chronifizierung der psychischen Störungen bei den Patienten und deren Zusammenhang zum Behandlungserfolg sowie nähere Untersuchungen zur „Dosis-Wirkungs-Relation“ wünschenswert. Aussagen hierzu konnten anhand des vorliegenden Datensatzes bisher nicht getroffen werden. Aus Sicht der Patienten, aber auch unter dem Gesichtspunkt der knappen finanziellen Mittel im Gesundheitswesen sind weitere Analysen zur Adäquatheit von Behandlungen und zur Qualität von Psychotherapien in der ambulanten Versorgung bedeutsam.

5.2 PATIENTENZUFRIEDENHEIT

Die Patientenzufriedenheit mit dem Behandlungsergebnis fiel wie erwartet recht hoch aus. Aus Qualitätssicht und im Hinblick auf eine angestrebte kontinuierliche Qualitätsverbesserung im Rahmen eines internen QM-Systems ist zu überlegen, zukünftig die Gründe für Unzufriedenheit näher zu explorieren. Außerdem ist erneut zu berücksichtigen, dass diese Urteile der Completer-Stichprobe entstammen und somit potenziell unzufriedene Therapieabbrecher nicht mit einbezogen und somit positive Verzerrungen wahrscheinlich sind.

5.3 AUFRECHTERHALTUNG DES THERAPIERFOLGS

Die Ergebnisse sprechen für eine Stabilität des erreichten Therapieerfolgs zum Katamnese-Zeitpunkt. Zu bemerken sind allerdings die geringeren Stichprobengrößen zum Katamnese-Zeitpunkt, wodurch die Aussagekraft bzw. die Generalisierbarkeit der Ergebnisse eingeschränkt

ist. Aus Sicht einer kontinuierlichen Qualitätsverbesserung sind Überlegungen zu treffen, wie eine höhere Rücklaufquote an Fragebögen v.a. zum Katamnesezeitpunkt erreicht werden könnte. Auch hier wären mehr publizierte Vergleichsdaten und fachlicher Austausch wünschenswert.

5.4 RESPONSE- UND REMISSIONSANALYSEN

Insgesamt fallen die Response- und Remissionsquoten nach der Methode der prozentualen Wertereduktion für fast alle Fragebögen deutlich höher aus als die nach der RCI-Methode berechneten. Betrachtet man die Prä-Mittelwerte der verschiedenen Instrumente, wird erkennbar, dass sich die Werte der Patienten durchschnittlich eher im unteren Bereich der jeweiligen Skala der Fragebögen SCL-90-R, BDI, BSQ, ACQ befinden (Ausnahme: MI-A). Der Anteil der als „klinisch signifikant verbessert“ bewerteten Patienten und damit auch die Bewertung der Ergebnisqualität fällt somit je nach eingesetzter Methode und je nach Ausmaß der Psychopathologie der untersuchten Stichprobe bzgl. des eingesetzten Instruments unterschiedlich aus. Aufgrund der vergleichsweise geringen bzw. spezifischeren Psychopathologie der Patienten am ZPP der Universität Greifswald erscheint zur Evaluation der Ergebnisqualität insgesamt die prozentuale Methode (besonders bei der Anwendung von störungsunspezifischen oder Breitbandinstrumenten) geeigneter bzw. günstiger.

Im Zeitvergleich wird ersichtlich, dass die Anteile klinisch signifikant verbesserter Patienten (nach der RCI-Methode) im Zeitraum 2009-2011 meist höher ausfallen als im Zeitraum 2004-2009. Als Ursachen könnten der Einsatz unterschiedlicher Parameter in der RCI-Berechnung, leicht abweichende Cut-Off-c-Werte

aufgrund des Einsatzes unterschiedlicher Normwerte, verschieden große Evaluationszeiträume und Stichproben sowie eine mögliche Verbesserung der Therapiequalität diskutiert werden. Insgesamt ist trotz z.T. hoher Responsequoten der Anteil an nicht veränderten/verschlechterten Patienten noch recht hoch. Aufgrund fehlender Standards bzgl. Qualitätskriterien und -maße ist allerdings eine institutionsübergreifende Bewertung der Ergebnisse nicht bzw. nur eingeschränkt möglich. Betrachten wir die Response- und Remissionsquoten für den GSI-Wert der Patienten am ZPP (66,4% und 50,9%), ist die Höhe vergleichbar mit Angaben von Hiller et al. (2009) aus der Hochschulambulanz Mainz (65,8% und 49,4% im Jahr 2007). Die Response- und Remissionsquoten bzgl. des BDI lagen bei uns im aktuellen Evaluationszeitraum bei 61,5% und 57,7%, in der Mainzer Ambulanz bei 74,6% und 55,9%.

Für zukünftige Vergleiche wäre es nun wichtig, interne Standards für Qualitätssicherung zu entwickeln, aufrechtzuerhalten und kontinuierlich zu verbessern. Der Vergleich der Kohorten aus den Jahren 2004-2009 im Verhältnis zur Kohorte 2009-2011 hat dafür eine gute Grundlage geliefert, indem am ZPP jetzt z.B. die Höhe von erwarteten Effektstärken oder Rücklaufquoten realitätsnah festgelegt werden kann, die es aus Qualitätssicht mindestens zu erreichen gilt. Zum anderen wäre wünschenswert, auch über die Einrichtungen hinweg kritisch darüber zu diskutieren und sich anzunähern, um letztlich realistische Aussagen über die Wirksamkeit ambulanter Psychotherapien im klinischen Alltag einer Hochschulambulanz treffen zu können.

LITERATUR

- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Berlin: Springer.
- Buchkremer, G. & Klingberg, S. (2001). Was ist wissenschaftlich fundierte Psychotherapie? *Nervenarzt*, 72, 20-30.
- Dilling, H., Mombour, W. & Schmidt, M. H. (Hrsg.) (2011). *Internationale Klassifikation psychischer Störungen. ICD-10 Kapitel V (F)*. Bern: Huber.
- Ehlers, A., Margraf, J. & Chambless, D. (2001). *Fragebogen zu körperbezogenen Ängsten, Kognitionen und Vermeidung (AKV)*. Göttingen: Hogrefe.
- Franke, G. (1995). *Die Symptom Checkliste von Derogatis – Deutsche Version (SCL-90-R)*. Göttingen: Beltz.
- Hautzinger, M., Bailer, M., Worall, H. & Keller, F. (1994). *Das Beck Depressionsinventar (BDI)*. Bern: Huber.
- Hautzinger, M. (2010). *Akute Depression. Fortschritte der Psychotherapie*. (Band 40). Göttingen: Hogrefe
- Hiller, W., Bleichhardt, G., Haaf, B., Legenbauer, T., Mauer-Matzen, K. & Rübler, D. (2006). Zertifizierung einer Psychotherapeutischen Hochschulambulanz nach DIN EN ISO 9001. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 35, 225-233.
- Hiller, W., Bleichhardt, G. & Schindler, A. (2009). Evaluation von Psychotherapien aus der Perspektive von Qualitätssicherung und Qualitätsmanagement. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 57, 7-22.
- Hiller, W., Schindler, A., Andor, T., Rist, F. (2011). Vorschläge zur Evaluation regulärer Psychotherapien an Hochschulambulanzen im Sinne der Phase-IV-Therapieforschung. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 40, 22-32.
- 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, 12-19.
- Kawski, S., Koch, U., Lubecki, P. & Schulz, H. (2005). Qualitätssicherung der Psychotherapie. In: W. Senf & M. Broda (Hrsg.), *Praxis der Psychotherapie. Ein integratives Lehrbuch* (S. 379-386). Stuttgart: Thieme.
- Lang, T., Helbig-Lang, S., Westphal, D., Glosster, A. T. & Wittchen, H.-U. (2012). *Expositionsbasierte Therapie der Panikstörung mit Agoraphobie. Ein Behandlungsmanual*. Göttingen: Hogrefe.
- Lutz, W. & Böhnke, J. R. (2010). Psychotherapieforschung: Verläufe, Prozesse, Ergebnisse und Qualitätssicherung. In: Lutz, W. (Hrsg.). *Lehrbuch Psychotherapie*. Bern: Huber.
- Maier-Riehle, B. & Zwingmann, C. (2000). Effektivitätsvarianten beim Eingruppen-Prä-Post-Design: Eine kritische Betrachtung. *Rehabilitation*, 39, 189-199.
- Richter, S. (2010). *Qualitätssicherung in der ambulanten Psychotherapie. Evaluation der ambulanten psychotherapeutischen Versorgung in der Poliklinischen Institutsambulanz der Ernst-Moritz-Arndt-Universität Greifswald*. Saarbrücken: SVH-Verlag.
- Schmidt-Bodenstein, S. (1998). Versorgungsstrukturelle Aspekte der Qualitätssicherung. In: A.-R. Laireiter & H. Vogel (Hrsg.), *Qualitätssicherung in der Psychotherapie und psychosozialen Versorgung* (S. 493-509). Tübingen: dgvt-Verlag.
- Schulz, W., Hoyer, H. & Hahlweg, K. (1998). Qualitätssicherung ambulanter Psychotherapie. Erste Ergebnisse einer Evaluierung der Verhaltenstherapie-Ambulanz der Technischen Universität Braunschweig. In: A.-R. Laireiter & Vogel, H. (Hrsg.), *Qualitätssicherung in der Psychotherapie und psychosozialen Versorgung* (S. 135-156). Tübingen: dgvt-Verlag.
- Seggar, L. B., Lambert, M. J. & Hansen, N. B. (2002). Assessing clinical significance: Application to the Beck depression inventory. *Behavior Therapy*, 33, 253-269.

Sulz, S. K. D. (2007). Richtliniengetreues Qualitätsmanagement für Psychotherapeuten – Kann psychotherapeutische Qualität zeitsparend gesteigert werden? *Psychotherapie in Psychiatrie, Psychotherapeutischer Medizin und Klinischer Psychologie*, 12, 216-227.

Wittchen, H.-U. & Pfister, H. (1997). *Diagnostisches Expertensystem DIA-X*. Frankfurt: Swets Test Services.



DR. SUSAN RICHTER

ERNST-MORITZ-ARNDT-UNIVERSITÄT
GREIFSWALD

INSTITUT FÜR PSYCHOLOGIE
ZENTRUM FÜR PSYCHOLOGISCHE
PSYCHOTHERAPIE (ZPP)

RUDOLF-BREITSCHIED-STR. 27
D-17487 GREIFSWALD

E-MAIL: susan.richter@uni-greifswald.de

FELICITAS GUTT

ERNST-MORITZ-ARNDT-UNIVERSITÄT
GREIFSWALD

INSTITUT FÜR PSYCHOLOGIE
ZENTRUM FÜR PSYCHOLOGISCHE
PSYCHOTHERAPIE (ZPP)

RUDOLF-BREITSCHIED-STR. 27
D-17487 GREIFSWALD

E-MAIL: gutt.felicitas@googlemail.com

PROF. DR. ALFONS HAMM

ERNST-MORITZ-ARNDT-UNIVERSITÄT
GREIFSWALD

INSTITUT FÜR PSYCHOLOGIE
LEHRSTUHLINHABER FÜR KLINISCHE
PSYCHOLOGIE/PSYCHOTHERAPIE

FRANZ-MEHRING-STR. 47
D-17487 GREIFSWALD

E-MAIL: hamm@uni-greifswald.de