

## Originalarbeit

*Aureliano Crameri, Margit Koemeda, Volker Tschuschke, Peter Schulthess, Agnes von Wyl*

# Ergebnisqualität ambulanter Psychotherapie: Ergebnisse aus der Grundversorgung in der Schweiz

**Zusammenfassung:** Die Ergebnisqualität von ambulanten Psychotherapien aus der Grundversorgung in der Schweiz wurde überprüft. Die untersuchte Stichprobe umfasste 300 Patienten, welche mehrheitlich mit einem tiefenpsychologischen oder humanistischen Therapieverfahren behandelt wurden. Die häufigsten Hauptdiagnosen waren affektive Störungen und Angststörungen; zwei Drittel der Patienten hatten eine oder mehrere komorbide Störungen. Fehlende Outcome-Werte wurden mit der multiplen Imputation behandelt. Die Analyse der Prä-post-Veränderungen zeigte grosse signifikante Effektstärken bei den globalen Masssen der psychosozialen Beeinträchtigung und der Symptombelastung. Die Ergebnisse der vorliegenden Stichprobe wurden mit denen aus Effectiveness-Studien verglichen, in welchen das Outcome von kognitiv-behavioralen Behandlungsansätzen evaluiert wurde. Der Vergleich zeigte, dass die Effektivität von humanistischen und tiefenpsychologischen Therapien vergleichbar mit der von evidenzbasierten kognitiv-verhaltenstherapeutischen Verfahren ist.

**Schlüsselwörter:** Ergebnisqualität, Routineversorgung, ambulante Psychotherapie, Prä-post-Vergleich, multiple Imputation

### The outcome quality of ambulant psychotherapy: findings from the primary health care in Switzerland

**Abstract:** This study investigated the outcome quality of outpatient psychotherapies in Switzerland's primary health care. The sample analyzed encompassed 300 patients, the majority of whom had been treated with humanistic-experiential psychotherapy or psychodynamic psychotherapy. The most frequent axis I diagnoses were mood disorders or anxiety disorders; two-thirds of the patients had one or more comorbid disorders. Missing outcome values were handled by multiple imputation. The analysis of the pre-post differences revealed large significant effect sizes in the global measures "psychosocial functioning" and "symptom distress". The results from this sample were compared with the results of published studies that tested the effectiveness of evidence-based cognitive-behavioral psychotherapies. The comparison revealed that the effectiveness of humanistic-experiential and psychodynamic treatment approaches was similar to the effectiveness of cognitive-behavioral approaches.

**Keywords:** outcome quality, primary care, outpatient psychotherapy, pre-post comparison, multiple imputation

### Qualità dei risultati della psicoterapia ambulatoriale: risultati riguardanti l'assistenza di base in Svizzera

**Riassunto:** Valutazione della qualità dei risultati delle psicoterapie ambulatoriali prestate nell'ambito dell'assistenza di base. Il campione esaminato comprendeva 300 pazienti, la maggioranza dei quali è stata trattata con metodi terapeutici della psicologia del profondo o umanistici. Le diagnosi principali più frequenti riguardavano disturbi affettivi e d'ansia; due terzi dei pazienti presentavano uno o più disturbi concomitanti. I valori di outcome assenti sono stati trattati con l'imputazione multipla. L'analisi dei cambiamenti pre-post ha mostrato grandi e significative dimensioni dell'effetto per l'ampiezza globale della limitazione psicosociale e del peso dei sintomi. I risultati del presente campione sono stati confrontati con quelli degli studi di effectiveness che valutavano l'outcome degli approcci di trattamento cognitivo-comportamentali. Il confronto ha mostrato che l'efficacia delle terapie umanistiche e della psicologia del profondo è paragonabile a quella delle procedure della terapia cognitiva e comportamentale basate su prove di efficacia.

**Parole chiave:** qualità dei risultati, assistenza di routine, psicoterapia ambulatoriale, confronti pre-post, imputazione multipla

## Einleitung

Massnahmen zur Sicherstellung und Verbesserung der psychotherapeutischen Grundversorgung gehören zum Standard eines modernen Gesundheitssystems (Laireiter, 1997). Eine der ersten Definitionen von Qualität im Gesundheitswesen stammt von Donabedian (1966), welche auf die drei in Abhängigkeit zueinander stehenden Aspekte der Struktur-, Prozess- und Ergebnisqualität fokussiert. In einem psychotherapeutischen Kontext ist das Konzept der Qualität, insbesondere der Ergebnisqualität, eng mit dem der Wirksamkeit verknüpft. Die allgemeine Wirksamkeit von Psychotherapie ist inzwischen durch eine Vielzahl von Studien abgesichert, wie dies in der aktuellen Auflage des „Handbook of Psychotherapy and Behavior Change“ (Lambert, 2013) ausführlich dokumentiert ist. Offen bleibt noch die Frage, welche der verschiedenen Therapieverfahren das Label „evidenzbasiert“ geltend machen dürfen. Vor allem für die kognitiv-behavioralen Therapieformen existiert eine umfangreiche Anzahl an randomisiert kontrollierten Studien, d. h. Studien mit einer hohen internen Validität, welche ihre Wirksamkeit belegen (Chambless & Ollendick, 2001). Die Effektivität von evidenzbasierten Therapieverfahren kann in der Alltagspraxis jedoch anders ausfallen als in kontrollierten Studien (Schindler et al., 2011). In der Praxis werden Therapien nicht selten in eklektischer Form und somit

abweichend von der getesteten manualisierten Version durchgeführt. Darüber hinaus ist die Effektivität der einzelnen Behandlungen von variierenden Faktoren abhängig, wie der Person des Therapeuten (Saxon & Barkham, 2012) oder der Beziehung zwischen Patient und behandelnder Person (Piechotta, 2008). Die Qualität von Behandlungen in der Grundversorgung sollte deshalb kontinuierlich überprüft werden.

Die Schweizer Charta für Psychotherapie als Dachverband für psychotherapeutische Ausbildungsinstitutionen, Fachverbände und Berufsverbände betreibt seit Jahren die Förderung qualitativer Standards in der Ausbildung zur Psychotherapie, in der Berufs- und Weiterbildungsethik und in der Wissenschaftlichkeit. Die Charta führte bereits 2001 eine repräsentative Erhebung unter Schweizer Psychotherapeuten durch, welche Fragen zur Ausbildung und Arbeitsweise der Therapeuten sowie zu den von ihnen behandelten Patienten umfasste (Schweizer et al., 2002).

Die Praxisstudie Ambulante Psychotherapie – Schweiz (PAP-S) ist ein weiteres von der Charta lanciertes Projekt, um die Prozess- und Ergebnisqualität von verschiedenen Therapiemethoden zu untersuchen. Eine ausführliche Beschreibung des Designs ist bei von Wyl et al. (2013) zu finden. Bezüglich der Prozessqualität wurden erste Ergebnisse, welche auf der qualitativen und quantitativen Auswertung von Tonaufnahmen von Therapiesitzungen basieren, bereits von Tschuschke et al. (2014) publiziert. Der vorliegende Artikel berichtet über das Outcome der in der Studie beforschten Therapieverläufe.

## **Methode**

Im Folgenden werden die Teile des Studiendesigns erläutert, mit welchen die Daten für die Überprüfung der Ergebnisqualität generiert wurden. Der Fokus der präsentierten Analysen liegt auf Prä-post-Vergleichen.

### *Messinstrumente und -zeitpunkte*

Die Patientinnen und Patienten nahmen an mindestens zwei externen Assessments teil, eines am Anfang (Prä-Assessment) und eines am Schluss (Post-Assessment) der Therapie. In diesen Assessments unterzog sich der Patient oder die Patientin einem diagnostischen Interview und füllte Fragebögen aus.

Folgende Indikatoren der Ergebnisqualität wurden herangezogen:

- *Hauptdiagnosen* auf der Achse I (klinische Syndrome) und der Achse II (Persönlichkeitsstörungen) des DSM-IV – für die Abklärung der Diagnosekriterien kamen die Strukturierten Klinischen Interviews des DSM-IV zum Einsatz (SKID-I und SKID-II; Wittchen et al., 1997)
- Aspekte der *Psychodynamik* wie *Konfliktausprägungen* (Operationalisierte Psychodynamische Diagnostik, OPD-2, Achse III; Arbeitskreis OPD, 2006) und *Strukturdefizite* (OPD-2, Achse IV) – die Items der Achse III und IV wurden zu drei Skalen zusammengefasst: (1) Skala „Konflikte“ als Mittelwert der 7 Items zu den repetitiv dysfunktionalen Konflikten (Achse III), (2) Skala „Selbstbezug“ als Mittelwert der Items 1a, 2a, 3a, 4a (Achse IV) und Skala „Objektbezug“ als Mittelwert der Items 1b, 2b, 3b, 4b (Achse IV)
- das allgemeine *Funktionsniveau* (Global Assessment of Functioning, GAF; Endicott et al., 1976; Saß et al., 2003)
- die Selbsteinschätzung der *psychischen Belastung* anhand folgender Fragebögen: Outcome Questionnaire-45 (OQ-45; Lambert et al., 2002), Brief Symptom Inventory (BSI; Franke, 2000), Beck Depression Inventory (BDI; Hautzinger et al., 1994), Sense of Coherence (SOC-13; Antonovsky, 1987) und Inkongruenzfragebogen (K-INK; Grosse-Holtforth & Grawe, 2003)

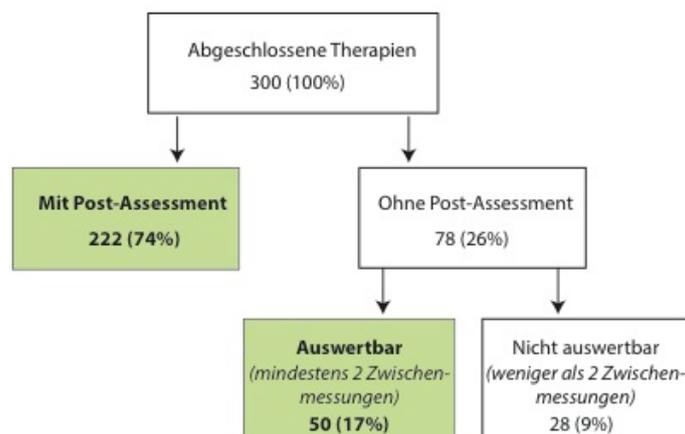
Zusätzlich wurden nach jeder 5. Sitzung im Rahmen der Prozessmessungen folgende Indikatoren erhoben:

- Selbsteinschätzung der *psychischen Belastung* anhand des OQ-45
- Selbsteinschätzung der *therapeutischen Beziehung* mit der Patientenversion des Helping Alliance Questionnaire (HAQ-P; Bassler et al., 1995) – dieser Fragebogen erfasst Aspekte der Qualität der therapeutischen Beziehung, für die vorliegenden Auswertungen haben wir die Zwei-Faktorenstruktur, welche Bassler et al. (1995) und De Weert-Van Oene et al. (1999) ermittelt haben, verwendet: die erste Skala misst die Qualität der Zusammenarbeit, die zweite das Ausmass an empfundener Hilfe; Bassler et al. (1995) betrachten die zweite Skala auch als Mass der Erfolgsszufriedenheit
- Einschätzung der therapeutischen Beziehung durch den Therapeuten mit der Therapeutenversion des HAQ-P (HAQ-T; Groß & Riedel, 1995).

## Stichprobe und Datenanalyse

Ein grosses Problem bei der Analyse des Behandlungsergebnisses (Outcome) unter naturalistischen Bedingungen stellen die fehlenden Werte dar, welche entstehen, wenn Patienten aus der Studie aussteigen und an der Schlussmessung nicht teilnehmen (diese Patienten werden im Folgenden als Drop-outs bezeichnet). Die Analyse von ausschliesslich vollständigen Fällen führt zu verzerrten Schätzungen des Outcomes, da unter ihnen der Anteil an Misserfolgen in der Regel kleiner ist als unter den Drop-outs. Um eine unverzerrte Behandlungsevaluation zu erhalten, sollten die Daten von allen behandelten Patienten in die Analyse einbezogen werden (Lachin, 2000). Eine inferenzstatistische Methode, welche zum State of the Art in der Behandlung von fehlenden Werten gehört, ist die multiple Imputation (Schafer & Graham, 2002; Jekauc et al., 2012). Mit dieser Methode werden anhand statistischer Modelle für jeden fehlenden Wert mehrere plausible Werte generiert. Die Imputation des Outcomes setzt neben der Prä-Messung regelmässige Zwischenmessungen im Laufe der Behandlung voraus. In der PAP-S-Studie wurden die Fortschritte regelmässig nach jeder 5. Sitzung mit Hilfe der Fragebögen OQ-45 und des HAQ gemessen. Cramer (2013) konnte anhand ausführlicher Simulationen zeigen, dass für eine robuste Schätzung des unvollständigen Outcomes die zwei letzten Zwischenmessungen ausreichen. Die multiplen Imputationen wurden mit dem R-Paket *mi* (Su et al., 2011) generiert.

Für die vorliegende Evaluation der Ergebnisqualität wurden die Therapien berücksichtigt, welche bis Ende Juni 2012 abgeschlossen wurden (Abb. 1). Es handelte sich insgesamt um 300 Patienten, 78 davon nahmen am Post-Assessment nicht teil. Für 50 dieser Drop-outs, welche mindestens 10 Therapiesitzungen in Anspruch nahmen, war es möglich, das Outcome zu imputieren. Insgesamt konnten deshalb 272 Behandlungen, d. h. mehr als 90 % der Stichprobe, in die Evaluation einbezogen werden.



**Abb. 1:** Auswertungsstichprobe

Um den Vergleich der Ergebnisse in den verschiedenen Skalen zu erleichtern, wurde die Effektstärke  $g$  herangezogen, welche aus der Division der Prä-post-Differenz durch die Standardabweichung der Prä-Werte berechnet wird (Becker, 1988; Glass, 1976; Grawe & Braun, 1994). Die Effektstärke wurde jeweils so umgepolt, dass ein positiver Wert einen positiven Therapieeffekt bedeutet. Für die Beurteilung der Signifikanz wurde für jede Effektstärke das 95 %-Konfidenzintervall berechnet. Umschliesst dieses Intervall den Wert null, so gilt die Effektstärke als nicht signifikant von null verschieden (Signifikanzniveau 5 %).

Tabelle 1 und 2 zeigen eine Beschreibung der analysierten Stichprobe. Der grösste Teil der Patienten wurden entweder mit einem humanistischen oder einem tiefenpsychologischen Ansatz behandelt (Schlegel et al., 2011). Die analysierte Stichprobe wurde von insgesamt 71 Therapeutinnen und Therapeuten rekrutiert. Davon haben 58 % eine Therapieausbildung humanistischer Richtung und 38 % eine Ausbildung tiefenpsychologischer Richtung; die restlichen 4 % sind ausgebildete Kunsttherapeutinnen.

Etwa 10 % der Patienten hatten keine Berufsbildung, weitere 35 % absolvierten eine Berufslehre, die restlichen wiesen eine höhere Schulbildung auf. Affektive Störungen und Angststörungen sind die häufigsten Hauptdiagnosen. Bei etwa zwei Dritteln der Patienten sind eine oder mehrere komorbide Störungen vorhanden. Rund 60 % litten auch in der Vergangenheit an einer psychischen Störung. Eine oder mehrere Persönlichkeitsstörungen wurden bei circa 40 % der Patienten diagnostiziert. Eine derart hohe Prävalenz an Persönlichkeitsstörungen (PS) unter den Patienten, die wegen psychischer Probleme professionelle Hilfe in Anspruch nehmen, wird von epidemiologischen Studien bestätigt (Grant et al., 2005).

Etwa drei Viertel der Therapien wurden regulär abgeschlossen (d. h. im gegenseitigen Einverständnis). In der analysierten Stichprobe dauerten die humanistischen Therapien im Durchschnitt 43 Sitzungen und die tiefenpsychologischen Therapien 36 Sitzungen. Ein solcher Unterschied in der Therapiedauer erwies sich aber als nicht signifikant ( $t[177] = -1,5$ ,  $p = 0,126$ ). Ebenfalls nicht signifikant war der Zusammenhang zwischen der Art des Behandlungsabschlusses und der des angewendeten Therapieverfahrens ( $\chi^2[3] = 2,2$ ,  $p = 0,529$ ).

Hauptansatz / Therapieverfahren	N	%
<b>Humanistische Psychotherapie</b>		
Integrative Körperpsychotherapie	72	26,5
Transaktionsanalyse	51	18,8
Gestalttherapie	28	10,3
Personale Existenzanalyse nach Längle	14	5,1
Logotherapie und Existenzanalyse nach Frankl	9	3,3
<b>Total</b>	<b>174</b>	<b>64,0</b>
<b>Tiefenpsychologie</b>		
Prozessorientierte Psychotherapie	48	17,6
Bioenergetische Analyse	27	9,9
Analytische Psychologie	7	2,6
Psychoanalyse	4	0,4
<b>Total</b>	<b>86</b>	<b>31,6</b>
<b>Integrative Psychotherapie</b>		
Kunst- und ausdrucksorientierte Psychotherapie	12	4,4
<b>Total</b>	<b>12</b>	<b>4,4</b>

**Tabelle 1:** Anzahl Patienten pro Therapieverfahren und Hauptansätze

	N (%) oder Mittelw. (SD)	
<b>Geschlecht</b>		
Frauen	179	(65,8 %)
Männer	93	(34,2 %)
<b>Alter</b>	39,9	(11,7)
<b>Schul- / Berufsbildung</b>		
Keine Berufsbildung	26	(9,6 %)
Berufslehre	94	(34,6 %)
Höhere Schulbildung	152	(55,9 %)
<b>Achse I, Hauptdiagnose</b>		
Affektive Störung	106	(39,0 %)
Angststörung	65	(23,9 %)
Anpassungsstörung	45	(16,5 %)
Andere	22	(8,0 %)
Keine (V-Kodierungen <sup>a</sup> )	34	(12,5 %)
<b>Achse I, Komorbidität</b>	182	(66,9 %)
<b>Achse I, Lifetime</b>	167	(60,0 %)
<b>Achse II, Hauptdiagnose</b>		
Cluster A (schizoid, paranoid u.a.)	7	(2,6 %)
Cluster B (Borderline, narzisstisch u.a.)	35	(12,9 %)
Cluster C (selbstunsicher, zwanghaft u.a.)	68	(25,0 %)
Keine	162	(59,5 %)
<b>Anzahl Therapiesitzungen</b>	40,7	(33,7)
<b>Art des Abschlusses</b>		
Im gegenseitigen Einverständnis	208	(76,5 %)
Auf Initiative des Patienten	52	(19,1 %)
Auf Initiative des Therapeuten	10	(3,7 %)
Abbruch wegen fehlender Kostensprache des Krankenversicherers	2	(0,7 %)

<sup>a</sup> V-Kodierungen, Zustände, die nicht einer psychischen Störung zuzuschreiben sind, aber Anlass zur Beobachtung oder Behandlung geben

**Tabelle 2:** Deskriptive Statistik (N = 272)

## Resultate

### Alle Patienten

In Tabelle 3 sind die Prä-post-Vergleiche der gesamten Stichprobe zusammengefasst. Alle Effektstärken sind statistisch signifikant. Nach Cohen (1992) entspricht eine standardisierte Prä-post-Differenz von 0,5 einem mittleren und eine von 0,8 einem starken Effekt. Betrachtet man globale Masse der psychosozialen Beeinträchtigung und der Symptombelastung, wie GAF, OQ-Total oder GSI, so befinden sich die Effekte in einem hohen Bereich. Die kleinsten Effekte ergeben sich in den Hauptdimensionen der psychodynamischen Struktur, d. h. Selbst- und Objektbezug. Die positiven Entwicklungen dieser zwei Komponenten der Persönlichkeit sind als klein bis mittel einzustufen.

		Prä <sup>a</sup>		Post		Effektstärke <sup>b</sup>		
		M	SE	M	SE	untere	g	obere
OQ-45	Symptombelastung	38,34	0,94	23,31	0,83	0,90	1,04	1,17
	Interpers. Beziehungen	15,88	0,41	10,62	0,43	0,70	0,85	1,00
	Soziale Integration	12,62	0,33	8,65	0,30	0,67	0,80	0,94
	Total	66,84	1,43	42,58	1,40	0,96	1,09	1,22
BSI	Somatisierung	0,67	0,04	0,38	0,03	0,33	0,48	0,64
	Ängstlichkeit	0,97	0,06	0,47	0,03	0,52	0,68	0,84
	Phobische Angst	0,54	0,04	0,22	0,02	0,40	0,53	0,66
	Zwanghaftigkeit	1,19	0,05	0,61	0,04	0,60	0,73	0,86
	Unsicherheit	1,30	0,06	0,69	0,04	0,56	0,70	0,83
	Paranoides Denken	0,93	0,06	0,46	0,03	0,45	0,59	0,74
	Globale Belastung (GSI)	0,96	0,04	0,46	0,03	0,76	0,89	1,01
BDI	Depression	16,84	0,73	6,81	0,46	0,86	1,00	1,15
SOC	Kohärenzgefühl	54,00	0,90	66,76	0,86	0,86	1,04	1,22
INK-K	Annäherung	3,09	0,06	3,80	0,06	0,76	0,94	1,13
	Vermeidung	2,69	0,05	2,01	0,05	0,69	0,87	1,05
OPD	Konflikte	1,72	0,04	1,36	0,03	0,57	0,73	0,88
	Selbstbezug	1,98	0,03	1,75	0,02	0,40	0,52	0,65
	Objektbezug	1,90	0,03	1,72	0,03	0,24	0,37	0,51
GAF	Allg. Funktionsniveau	62,39	1,09	76,30	0,98	0,79	0,99	1,19

<sup>a</sup> M, Mittelwert; SE, Standardfehler

<sup>b</sup> g, Effektstärke mit oberer und unterer Grenze des Konfidenzintervalls.

**Tabelle 3:** Prä-post-Vergleiche bei allen Patienten (N = 272)

Die Einschätzungen bezüglich der therapeutischen Beziehung während der Anfangs- (erste Prozessmessung nach der 5. Sitzung) und Schlussphase (letzte Prozessmessung) sind in Tabelle 4 zusammengefasst. Die Einschätzungen der Patienten sind im Durchschnitt höher als die entsprechenden Werte der Therapeuten. Das Ausmass an empfundener Hilfe nimmt aus der Perspektive beider Beteiligten im Laufe der Behandlung mit einer mittleren bis grossen Effektstärke zu. Die durchschnittliche Einschätzung der Kooperation aus Sicht der Patienten nimmt leicht ab. Diese leichte Abnahme ist durch die Therapieabbrecher bedingt, deren Einschätzungen von einem anfänglichen Mittelwert von 5,31 zu einem abschliessenden Mittelwert von 5,08 sinken. Die Einschätzungen der Patienten mit einem regulären Abschluss hingegen bleiben mit einem Durchschnitt von 5,44 konstant. Die fehlende Zunahme der von dieser Patientengruppe gelieferten Werte ist möglicherweise vom vorhandenen Deckeneffekt in der Skala bedingt. Die entsprechende Skala des Therapeuten, welche weniger stark von einem Deckeneffekt betroffen ist als die des Patienten, registriert hingegen eine Zunahme in der Kooperationsqualität.

		Erste Prozessmessung		Letzte Prozessmessung		Effektstärke		
		M	SE	M	SE	untere	g	obere
HAQ-P	Kooperation	5,41	0,03	5,35	0,04	-0,26	-0,12	0,03
	Empfundene Hilfe	4,41	0,05	4,83	0,06	0,41	0,55	0,69
HAQ-T	Kooperation	4,63	0,04	4,78	0,04	0,10	0,26	0,41
	Empfundene Hilfe	4,07	0,04	4,57	0,05	0,59	0,76	0,93

**Tabelle 4:** Entwicklung der therapeutischen Beziehung in der Gesamtstichprobe (N = 272)

Von den 238 Patienten, welche am Anfang der Therapie die Kriterien einer Achse-I-Störung erfüllten, erreichten 170 (71 %) eine Remission ihrer Symptome beim Abschluss ihrer Behandlung (Ergebnis des SKID-I-Interviews).

#### *Patienten mit affektiven Störungen oder Angststörungen*

Zusammengefasst in den Tabellen 5 und 6 sind die Prä-post-Veränderungen für die zwei wichtigsten Störungsgruppen der Achse I, nämlich Patienten mit affektiven Störungen und mit Angststörungen. Bezüglich der Reduktion der globalen Belastung (OQ-45-Total und GSI) sind bei beiden Gruppen ähnlich starke Effekte der Behandlung zu registrieren. Für Patienten mit einer affektiven Störung fällt die Effektstärke des GAF-Wertes grösser aus als bei Patienten mit einer Angststörung. Dieser Unterschied ist durch die verschiedenen Anfangswerte der zwei Gruppen bedingt. Beide erreichen am Schluss der Behandlung einen Durchschnitt von etwa 75 Punkten auf der GAF-Skala.

		Prä		Post		Effektstärke		
		M	SE	M	SE	untere	g	obere
OQ-45	Total	74,07	2,04	49,38	2,25	0,90	1,11	1,31
BSI	Globale Belastung (GSI)	1,15	0,06	0,58	0,05	0,81	1,02	1,24
GAF	Allg. Funktionsniveau	57,36	1,51	74,78	1,53	0,99	1,24	1,49
BDI	Depression	19,97	1,07	8,47	0,74	0,95	1,16	1,36
BSI	Ängstlichkeit	1,13	0,09	0,59	0,05	0,48	0,73	0,99

**Tabelle 5:** Prä-post-Vergleiche bei Patienten mit einer affektiven Störung (N = 106)

		Prä		Post		Effektstärke		
		M	SE	M	SE	untere	g	obere
OQ-45	Total	72,14	2,77	46,35	2,96	0,90	1,16	1,41
BSI	Globale Belastung (GSI)	1,06	0,07	0,50	0,06	0,70	0,99	1,27
GAF	Allg. Funktionsniveau	60,54	1,74	74,59	1,81	0,69	1,00	1,31
BDI	Depression	17,72	1,38	6,87	1,01	0,81	1,09	1,36
BSI	Ängstlichkeit	1,16	0,10	0,51	0,07	0,59	0,90	1,20

**Tabelle 6:** Prä-post-Vergleiche bei Patienten mit einer Angststörung (N = 65)

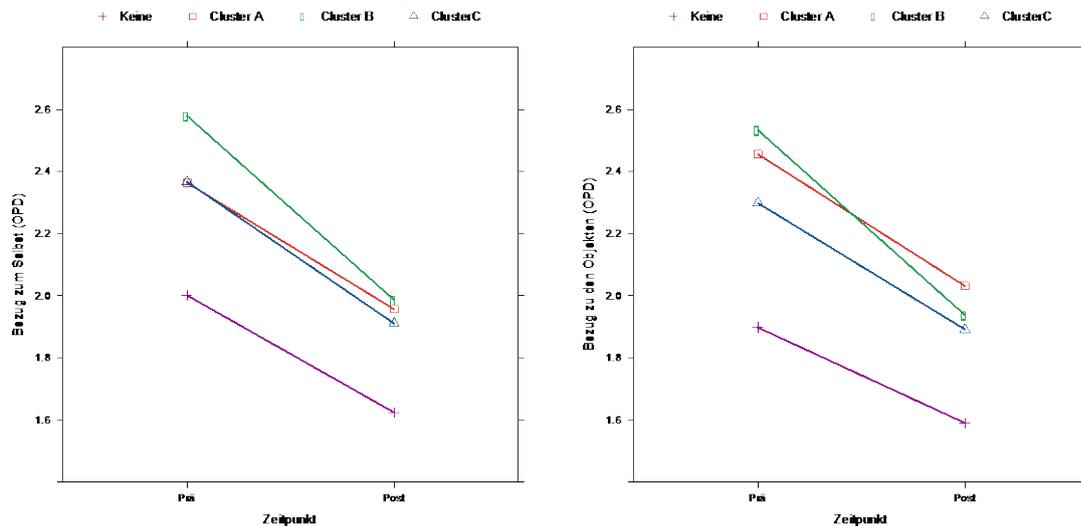
Eine Remission der Symptome wird von 66 % der Patienten mit einer affektiven Störung und von 67 % der Patienten mit einer Angststörung erreicht.

#### *Patienten mit Persönlichkeitsstörungen*

In Tabelle 7 finden sich die Ergebnisse der Patienten mit Persönlichkeitsstörungen (PS). Auch für diese Gruppe bewegen sich die Effektstärken bei den globalen Massen der Beeinträchtigung in einem hohen Bereich und sind nicht signifikant kleiner als die Effektstärken für Patienten ohne PS (z. B. OQ-45-Total:  $F[1, 69,2] = 0,3, p = 0,602$ ). Signifikante Unterschiede lassen sich zwischen Patienten mit einer PS des Clusters B (mehrheitlich PS des Typs Borderline oder des narzisstischen Typs) und solchen ohne PS bei den Hauptdimensionen der psychodynamischen Struktur feststellen (Abb. 2).

		Prä		Post		Effektstärke		
		M	SE	M	SE	untere	g	obere
OQ-45	Total	73,49	2,44	50,39	2,51	0,82	1,04	1,26
BSI	Globale Belastung (GSI)	1,14	0,07	0,60	0,05	0,74	0,97	1,19
GAF	Allg. Funktionsniveau	57,36	1,46	71,40	1,59	0,73	1,00	1,27
BDI	Depression	19,24	1,30	9,06	0,84	0,78	1,02	1,27
BSI	Ängstlichkeit	1,09	0,10	0,56	0,06	0,46	0,73	1,01

**Tabelle 7:** Prä-post-Vergleiche bei Patienten mit einer Persönlichkeitsstörung (N = 110)



**Abb. 2:** Prä-post-Veränderungen in den Hauptdimensionen der psychodynamischen Struktur in Abhängigkeit von Achse-II-Diagnosen.

Die Effektstärken in der Dimension „Selbstbezug“ betragen für Patienten ohne PS 0,38 (SE = 0,04) und für Patienten mit einer PS des Clusters B 0,59 (SE = 0,08); dieser Unterschied ist signifikant ( $t[135,5] = 2,5, p = 0,014$ ). In der Dimension „Objektbezug“ ist die Effektstärke für Patienten vom Cluster B mit 0,60 (SE = 0,08) ebenfalls signifikant grösser als die der Patienten ohne PS, welche nur 0,31 (SE = 0,04) beträgt ( $t[144,8] = 3,8, p < 0,001$ ).

#### Therapieverfahren

Ein Vergleich zwischen humanistischen und tiefenpsychologischen Therapien bezüglich der Verbesserungen auf globaler Ebene ist in Tabelle 8 zu sehen. Auf der Skala OQ-45-Total erreichen humanistische Therapien im Durchschnitt eine höhere Effektstärke als tiefenpsychologische Therapien; dieser Unterschied ist aber nicht signifikant ( $F[1, 4164,3] = 2,2, p = 0,136$ ). Ebenfalls nicht signifikant ist der Unterschied auf der Skala GSI ( $F[1, 131,6] = 0,2, p = 0,674$ ).

In 73,7 % der humanistischen Therapien wird eine Remission der Achse-I-Symptome erzielt, während dies nur in 67,6 % der tiefenpsychologischen Therapien der Fall ist. Dieser Unterschied ist jedoch nicht signifikant ( $OR = 1.34, CI95\% = \{0,71, 2,54\}$ ).

		Prä		Post		Effektstärke		
		M	SE	M	SE	untere	g	obere
OQ Total	Humanistisch	67,83	1,79	41,76	1,76	1,01	1,17	1,33
	Tiefenpsychologisch	65,43	2,51	43,79	2,46	0,75	0,97	1,19
GSI	Humanistisch	0,99	0,05	0,47	0,03	0,77	0,93	1,10
	Tiefenpsychologisch	0,94	0,06	0,44	0,04	0,65	0,87	1,08

**Tabelle 8:** Prä-post-Vergleiche zwischen humanistischen und tiefenpsychologischen Therapien (N = 174 bzw. N = 86)

An den humanistischen und tiefenpsychologischen Therapieformen wird nicht selten ein Mangel an evidenzbasierten Wirksamkeitsstudien kritisiert. Im Gegensatz dazu existiert für die kognitiv-behavioralen Therapieformen eine umfangreiche Anzahl Studien, welche ihre Wirksamkeit belegen (Chambless & Ollendick, 2001). Wir haben deshalb die Ergebnisse der PAP-S-Studie mit denen aus 5 Effectiveness-Studien (d. h. Studien unter Praxisbedingungen) zu kognitiv-behavioralen Therapieansätzen verglichen (Tabelle 9).

Studie	N	Diagnosegruppe	Sitzungen (M)	Prä-Werte		Prä-post- Effektstärke <sup>a</sup>	
				GSI	BDI	GSI	BDI
Nelson & Hiller (2013)	1866	gemischt (40 % Depr., 26 % Angstst.)	33,8	1,09	18,9	0,72	0,54
Puschner et al. (2008)	64	gemischt (48 % F3, 41 % F4)	32,3	0,96		0,68	
Westbrook & Kirk (2005)	151	gemischt (19 % Depr., 56 % Angstst.)	13,2		16,9		0,67
Gibbons et al. (2010)	217	Depressionen	15,9		26,4		1,06
Strauss et al (2006)	40	PS (selbstunsicher, zwanghaft)	43,3		15,9		1,18

<sup>a</sup> Prä-post-Effektstärken, falls im Original nicht vorhanden, wurden nach Becker (1988) berechnet.

**Tabelle 9:** Studien zur Effectiveness von kognitiv-behavioralen Therapieansätzen

Zwei Studien sind aus dem deutschsprachigen Raum (Nelson & Hiller, 2013; Puschner et al., 2008), eine aus Grossbritannien (Westbrook & Kirk, 2007) und zwei aus den USA (Gibbons et al., 2010; Strauss et al., 2006). Unsere Stichprobe weist hinsichtlich der Zusammensetzung der Diagnosen und der initialen Belastung die grösste Ähnlichkeit mit der Stichprobe von Nelson und Hiller (2013) auf. Insgesamt lässt sich erkennen, dass die Effektstärken der Therapien aus der PAP-S-Studie, welche mehrheitlich nach einem humanistischen oder tiefenpsychologischen Konzept durchgeführt wurden, in den meisten Vergleichen nicht geringer sind als die Effektstärken von den sogenannten evidenzbasierten verhaltenstherapeutischen Behandlungsansätzen. Beschränkt man den Vergleich auf Studien mit einer heterogenen Diagnosezusammensetzung (Nelson & Hiller, 2013; Puschner et al., 2008; Westbrook & Kirk, 2007), so stellt man im Gegenteil grössere Effekte bei den Behandlungen aus der PAP-S-Studie fest. Diese Unterschiede müssen aber nicht unbedingt mit der Therapieform zusammenhängen, sie können auch ein Ausdruck der Dosis-Wirkungs-Beziehung sein. Die Dauer der Verhaltenstherapien ist bei den meisten Studien nämlich kürzer als bei den Therapien in unserer Stichprobe. Nichtsdestotrotz unterstützen die Ergebnisse der fünf Vergleichsstudien die Hypothese, dass unter Praxisbedingungen die Wirksamkeit der kognitiv-behavioralen Ansätze im gleichen Spektrum wie die der humanistischen und tiefenpsychologischen Ansätze liegt. Outcome-Unterschiede sind vermutlich eher auf die Forschungsmethodologie als auf die Art der beforschten Interventionen zurückzuführen. Gemäss Schindler et al. (2011) liegt die durchschnittliche Effektstärke bei randomisierten Studien mit einer hohen internen Validität, in welchen die Wirksamkeit von Verhaltenstherapie bei Depressionen überprüft wird, bei 1,39 und somit höher als der Durchschnitt bei Effectiveness-Studien mit Schwerpunkt auf externer Validität mit vergleichbaren Patientenstichproben.

### Zusammenfassung und Diskussion

Das Ziel der evidenzbasierten Medizin ist „die Spreu vom Weizen zu trennen“, d. h. wirksame von unwirksamen oder wenig wirksamen Behandlungsmethoden zu unterscheiden. Dafür werden randomisiert-kontrollierte Studien durchgeführt, in welchen Behandlungen unter möglichst optimalen Bedingungen geprüft werden (z. B. die Patienten haben in der Regel keinen Substanzmissbrauch als Komorbidität oder die Therapeuten erhalten Supervision). Die Effektivität von Behandlungen, die in der Grundversorgung durchgeführt werden, ist aber nicht nur von der zugrunde liegenden „evidenzbasierten“ Technik abhängig, sondern auch von anderen Faktoren, wie dem Schweregrad der Störung, der Qualität der therapeutischen Beziehung oder der Person des Therapeuten. Deshalb sollte die Ergebnisqualität von psychotherapeutischen Routinebehandlungen ebenfalls untersucht werden. In der vorliegenden Arbeit wird die Ergebnisqualität anhand von Prä-post-Vergleichen evaluiert. Die Analysen basieren auf einer Stichprobe von 300 konsekutiv aufgenommenen Patienten, die im Rahmen der PAP-S-Studie rekrutiert wurden. Die häufigsten Hauptdiagnosen in dieser Stichprobe waren affektive Störungen und Angststörungen. Der Anteil an Patienten mit einer oder mehreren komorbiden Störungen betrug zwei Drittel. Durch die Behandlung der fehlenden Werte mit der multiplen Imputation konnten die Daten von 91 % der Therapieverläufe für die statistische Analyse verwertet werden. Mit diesem Vorgehen wurde dafür gesorgt, dass auch Drop-outs mit einem bescheidenen Behandlungsergebnis in der Evaluation berücksichtigt wurden.

Die Analyse der Prä-post-Veränderungen zeigte grosse signifikante Effektstärken bei den globalen Massen der psychosozialen Beeinträchtigung und der Symptombelastung. Die Zustandsverbesserungen von Patienten mit einer affektiven Störung und von denen mit einer Angststörung waren vergleichbar ausgeprägt. Patienten mit einer PS des Clusters B (mehrheitlich Borderline oder narzisstisch) hatten zu Therapiebeginn grössere Defizite der psychischen Struktur als Patienten ohne PS. Die erste Gruppe erreichte auch signifikant mehr positive strukturelle Veränderungen als die zweite Gruppe. Diese Ergebnisse weisen darauf hin, dass Patienten mit einer

Borderline-PS oder mit einer narzisstischen PS von einer humanistischen oder tiefenpsychologischen Therapie gut profitieren können.

Ein Vergleich zwischen humanistischen und tiefenpsychologischen Therapien zeigte höhere Effektstärken beim ersten Behandlungsansatz. Der Unterschied, der nicht signifikant ausfiel, lässt sich aber auch durch die unterschiedliche „Dosis“ erklären. Die humanistischen Behandlungen waren nämlich in unserer Stichprobe tendenziell länger als die tiefenpsychologischen Behandlungen.

Die Ergebnisse unserer Stichproben wurden mit denen aus 5 publizierten Effectiveness-Studien verglichen, in welchen das Outcome von kognitiv-behavioralen Behandlungsansätzen evaluiert wurde. Der Vergleich zeigte, dass die Effektstärken bei unserer Stichprobe gleich oder höher waren als die bei Verhaltenstherapien. Es muss aber erwähnt werden, dass die Verhaltenstherapien kürzer waren als die humanistischen und tiefenpsychologischen Behandlungen in unserer Studie. Der Unterschied könnte deshalb nicht primär durch Unterschiede in den Interventionen bedingt sein, sondern lediglich den Effekt einer unterschiedlichen Dosierung darstellen.

Nichtsdestotrotz unterstützen die Ergebnisse die Hypothese, dass unter Praxisbedingungen die Wirksamkeit von humanistischen und tiefenpsychologischen Ansätzen im gleichen Spektrum wie die der kognitiv-behavioralen Ansätze liegt.

### **Autoren und Autorinnen**

Aureliano Cramer, MSc, Qualitätsmanager, Wissenschaftlicher Mitarbeiter an der Abteilung Forschung Psychotherapie und psychische Gesundheit an der Zürcher Hochschule für Angewandte Wissenschaften, Departement Angewandte Psychologie, Dozent für Forschungsmethoden.

Margit Koemeda, Dr. Dipl.-Psych., Psychotherapeutin ASP, Lehrtherapeutin und Trainerin SGBAT, IIBA. Mitglied der Wissenschaftskommission der Schweizer Charta für Psychotherapie; Redaktionsmitglied in drei Fachzeitschriften; Autorin von Büchern und Fachartikeln zu Themen der Psychotherapieforschung und Körperpsychotherapie.

Volker Tschuschke, Univ.-Prof. Dr. Dipl.-Psych., ist emeritierter ehemaliger Lehrstuhlinhaber des Faches Medizinische Psychologie an der Universitätsklinik Köln, jetzt Leiter des Studiengangs Psychotherapiewissenschaft an der Sigmund-Freud-Privatuniversität Berlin. Er ist Psychoanalytiker und war jahrelang Dozent an verschiedenen Ausbildungsinstituten, ist Autor von zahlreichen Büchern und Beiträgen in internationalen und nationalen Fachzeitschriften und arbeitet außerdem als Psychotherapeut und Supervisor.

Peter Schulthess, MSc, Psychotherapeut ASP, Ausbilder für Gestalttherapie am IGW Schweiz. Internationale Lehrtätigkeit für Gestalttherapie. Präsident der Schweizer Charta für Psychotherapie und Mitglied in deren Wissenschaftskommission. Redaktion „Psychotherapie-Wissenschaft“. Diverse Publikationen zur Psychotherapie und Gestalttherapie, Herausgeber einer Buchreihe zur Gestalttherapie im Verlag EHP.

Agnes von Wyl, Prof. Dr., ist Leiterin der Abteilung Forschung Psychotherapie und psychische Gesundheit an der Zürcher Hochschule für Angewandte Wissenschaften, Departement Angewandte Psychologie. Sie arbeitet ausserdem als Dozentin und psychoanalytische Psychotherapeutin. Ihre Forschungsgebiete umfassen Psychotherapieforschung, Infant Mental Health und psychische Gesundheitsförderung.

### **Korrespondenz**

Aureliano Cramer  
ZHAW Zürcher Hochschule für Angewandte Wissenschaften  
Departement Angewandte Psychologie  
Pfungstweidstrasse 96  
Postfach 707  
8037 Zürich  
E-Mail: aureliano.cramer@zhaw.ch

### **Literatur**

- Antonovsky, A. (1987). *Unraveling the mystery of health: how people manage stress and stay well*. San Francisco: Jossey-Bass.
- Arbeitskreis OPD (2006). *Operationalisierte psychodynamische Diagnostik OPD-2: das Manual für Diagnostik und Therapieplanung*. Bern: Huber.
- Bassler, M., Potratz, B., & Krauthauser, H. (1995). Der „Helping Alliance Questionnaire“ (HAQ) von Luborsky. *Psychotherapeut*, 40, 23–32.
- Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41, 257–278.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: controversies and evidence. *Annual Review of Psychology*, 52, 685–716.

- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Cramer, A. (2013). Die multiple Imputation in der Bewertung der Ergebnisqualität von Psychotherapien unterschiedlicher Dauer. Bericht. Zürich: Departement Angewandte Psychologie, Zürcher Hochschule für Angewandte Wissenschaften.
- De Weert-Van Oene, G. H., De Jong, C. A., Jörg, F., & Schrijvers, G. J. (1999). The Helping Alliance Questionnaire: psychometric properties in patients with substance dependence. *Substance Use and Misuse*, 34, 1549–1569.
- Donabedian, A. (1966). Evaluating the quality of medical care. *Milbank Memorial Fund Quarterly*, 44(3, Suppl.), 166–203.
- 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, 776–771.
- Franke, G. H. (2000). BSI: brief symptom inventory von L. R. Derogatis (Kurzform der SCL-90 R) – deutsche Version. Göttingen: Beltz Test.
- Gibbons, C. J., Fournier, J. C., Stirman, S. W., DeRubeis, R. J., Crits-Christoph, P., & Beck, A. T. (2010). The clinical effectiveness of cognitive therapy for depression in an outpatient clinic. *Journal of Affective Disorders*, 125, 169–176.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3–8.
- Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Chou, S. P., Ruan, W. J., & Huang, B. (2005). Co-occurrence of 12-month mood and anxiety disorders and personality disorders in the US: results from the national epidemiologic survey on alcohol and related conditions. *Journal of Psychiatric Research*, 39, 1–9.
- Grawe, K., & Braun, U. (1994). Qualitätskontrolle in der Psychotherapie. *Zeitschrift für Klinische Psychologie*, 23, 242–267.
- Groß, A., & Riedel, W.P. (1995). Therapieergebnis und Komplementarität in der Therapeut-Patient-Beziehung: eine Analyse mit Hilfe von SASB. Regensburg: Roderer.
- Grosse-Holtforth, M., & Grawe, K. (2003). Der Inkongruenzfragebogen (INK): ein Messinstrument zur Analyse motivationaler Inkongruenz. *Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis*, 32, 315–323.
- Hautzinger, M., Bailer, M., Worall, H., & Keller, F. (Bearb.) (1994). Beck-Depressions-Inventar (BDI). Bern: Huber.
- Jekauc, D., Völkle, M., Lämmle, L., & Woll, A. (2012) Fehlende Werte in sportwissenschaftlichen Untersuchungen: eine anwendungsorientierte Einführung in die multiple Imputation mit SPSS. *Sportwissenschaft*, 42, 126–136.
- Lachin, J. M. (2000). Statistical considerations in the intent-to-treat principle. *Controlled Clinical Trials*, 21, 167–189; Erratum, *Controlled Clinical Trials*, 21, 526.
- Laireiter, A.-R. (1997). Qualitätssicherung von Psychotherapie: Struktur-, Prozess- und Ergebnisqualität in der ambulanten Praxis. *Psychotherapie Forum*, 5, 203–218.
- Lambert, J. M. (Hrsg.) (2013). *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*, 6. Aufl. Hoboken, N.J.: Wiley.
- Lambert, M. J., Hannöver, W., Nisslmüller, K., Richard, M., & Kordy, H. (2002). Fragebogen zum Ergebnis von Psychotherapie: zur Reliabilität und Validität der deutschen Übersetzung des Outcome Questionnaire 45.2 (OQ-45.2). *Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis*, 31, 40–46.
- Nelson, N. E., & Hiller, W. (2013). Therapeutischer Misserfolg in der ambulanten Psychotherapie: Daten aus einer verhaltenstherapeutischen Hochschulambulanz. *Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis*, 42, 217–229.
- Piechotta, B. (2008). *PsyQM: Qualitätsmanagement für psychotherapeutische Praxen*. Heidelberg: Springer.
- Puschner, B., Wolf, M., & Kraft, S. (2008). Helping alliance and outcome in psychotherapy: what predicts what in routine outpatient treatment? *Psychotherapy Research*, 18, 167–178.
- Saß, H., Wittchen, H.-U., Zaudig, M., & Houben, I. (Bearb.) (2003). *Diagnostisches und statistisches Manual psychischer Störungen: Textrevision – DSM-IV-TR*. Göttingen: Hogrefe.
- Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology*, 80, 535–546; Erratum, *Journal of Consulting and Clinical Psychology*, 80, 546.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Schindler, A. C., Hiller, W., & Withöft, M. (2011). Benchmarking of cognitive-behavioral therapy for depression in efficacy and effectiveness studies: how do exclusion criteria affect treatment outcome? *Psychotherapy Research*, 21, 644–657.
- Schlegel, M., Meier, I., & Schulthess, P. (2011). *Psychotherapien: ein Führer der Schweizer Charta für Psychotherapie für die in ihr vertretenen tiefenpsychologischen, humanistischen und integrativen Psychotherapieverfahren*. Zürich: Schweizer Charta für Psychotherapie.
- Schweizer, M., Buchmann, R., Schlegel, M., & Schulthess, P. (2002). Struktur und Leistung der Psychotherapieversorgung in der Schweiz: Erhebung der Schweizer Charta für Psychotherapie. *Psychotherapie Forum*, 10, 127–146.
- Strauss, J. L., Hayes, A. M., Johnson, S. L., Newman, C. F., Brown, G. K., Barber, J. P., Laurenceau, J.-P., & Beck, A. T. (2006). Early alliance, alliance ruptures, and symptom change in a nonrandomized trial of cognitive therapy for avoidant and obsessive-compulsive personality disorders. *Journal of Consulting and Clinical Psychology*, 74, 337–345.
- Su, Y.-S., Gelman, A., Hill, J., & Yajima, M. (2011). Multiple imputation with diagnostics (mi) in R: opening windows into the black box. *Journal of Statistical Software*, 45, issue 2.
- Tschuschke, V., Cramer, A., Koehler, M., Berglar, J., Muth, K., Staczan, P., von Wyl, A., Schulthess, P., & Koemeda-Lutz, M. (2014). The role of therapists' treatment adherence, professional experience, therapeutic alliance, and clients' severity of psychological problems: prediction of treatment outcome in eight different psychotherapy approaches; preliminary results of a naturalistic study. *Psychotherapy Research*, DOI 10.1080/10503307.2014.896055.

- Von Wyl, A., Cramer, A., Koemeda, M., Tschuschke, V., & Schulthess, P. (2013). Praxisstudie ambulante Psychotherapie Schweiz (PAP-S): Studiendesign und Machbarkeit. *Psychotherapie-Wissenschaft*, 3, 6–22.
- Westbrook, D., & Kirk, J. (2005). The clinical effectiveness of cognitive behaviour therapy: outcome for a large sample of adults treated in routine practice. *Behaviour Research and Therapy*, 43, 1243–1261; Erratum, *Behaviour Research and Therapy*, 45, 1703–1704.
- Wittchen H.-U., Zaudig M., & Fydrich T (1997). *Strukturiertes klinisches Interview für DSM-IV: Achse I und II*. Göttingen: Hogrefe.

## **La qualité des résultats de la psychothérapie ambulatoire: Résultats des soins de base en Suisse**

La médecine fondée sur les faits a pour objet de «séparer le bon grain de l'ivraie», c'est-à-dire de faire la distinction entre méthodes de traitement efficaces et méthodes de traitement inefficaces. A cet effet, on effectue des essais randomisés contrôlés dans lesquels les traitements sont testés dans les meilleures conditions possibles (par exemple, les patients ne présentent en général ni abus de substances ni comorbidités ou bien les thérapeutes sont suivis par supervision). Toutefois, l'efficacité des traitements qui sont administrés dans le cadre des soins de base, ne dépend pas uniquement de la technique sous-jacente «fondée sur les faits», mais aussi d'autres facteurs comme le degré de gravité du trouble, la qualité de la relation thérapeutique ou la personne du ou de la thérapeute. C'est pourquoi, il convient également d'examiner la qualité des résultats des traitements psychothérapeutiques de routine. Dans le présent papier, la qualité des résultats est évaluée en fonction de comparaisons pré- et post-traitement. Les analyses se fondent sur un échantillon de 300 patients consécutifs, recrutés dans le cadre de l'étude sur la pratique de la psychothérapie ambulatoire (PAP-S). Dans cet échantillon, les diagnostics les plus fréquents étaient des troubles affectifs et anxieux. La part des patients présentant un ou plusieurs troubles comorbides était de deux tiers. Le traitement par imputation multiple des valeurs manquantes nous a permis d'exploiter les données de 91% des processus thérapeutiques pour l'analyse statistique. Cette démarche nous a ainsi permis d'intégrer dans l'évaluation les abandons présentant de modestes résultats de traitement.

L'analyse des modifications pré- et post-traitement a révélé d'importantes tailles d'effets significatives dans les masses globales de troubles psychosociaux et de symptômes. Le niveau d'amélioration de l'état des patients souffrant de troubles affectifs et des patients souffrant de troubles anxieux était comparable. Les patients du cluster B souffrant de troubles de la personnalité (pour la plupart borderline ou narcissique) présentaient, au début du traitement, des déficits de la structure psychique plus importants que les patients sans troubles de la personnalité. Ils ont également accompli des transformations structurelles positives significativement plus importantes. Ces résultats démontrent que les patients souffrant de troubles de la personnalité borderline ou narcissique peuvent tirer profit d'une thérapie humaniste ou de psychologie des profondeurs.

Une comparaison entre thérapies humanistes et thérapies de psychologie des profondeurs a révélé une taille d'effet plus importante lors du premier traitement. La différence, statistiquement insignifiante, peut aussi s'expliquer par la différence de «dosage». Dans notre échantillon en effet, les traitements humanistes avaient tendance à être plus longs que les traitements de psychologie des profondeurs.

Les résultats de notre échantillon ont été comparés à ceux de 5 études d'efficacité dans lesquelles étaient évalués les résultats de démarches thérapeutiques cognitivo-comportementales. La comparaison a montré que les tailles d'effet dans notre échantillon étaient égales ou supérieures à celles des thérapies comportementales. Il faut encore mentionner que les thérapies comportementales étaient plus courtes que les thérapies humanistes et de psychologie des profondeurs retenues pour notre étude. C'est pourquoi, les divergences pourraient ne pas découler de différences dans les interventions, mais être seulement l'effet d'un dosage différent.

Néanmoins les résultats confortent l'hypothèse que, dans la pratique professionnelle, l'efficacité des approches humanistes et de psychologie des profondeurs est équivalente à celle des approches cognitivo-comportementales.

**Mots clés:** qualité des résultats, soins de routine, psychothérapie ambulatoire, comparaisons pré- et post-opératoire, imputation multiple