

Frank Jacobi

**Kosten-Effektivitäts- und Kosten-Nutzen-Analyse
psychologischer Angstbehandlung**

**Dissertationsschrift zur Erlangung des Doktorgrades Dr. rer. nat.
an der Technischen Universität Dresden**

Juli 2001

Betreuer der Arbeit:

Prof. Dr. Jürgen Margraf (Universität Basel)

Prof. Dr. Hans-Ulrich Wittchen (TU Dresden)

Disputation der Arbeit und Rigorosum: 3.6.2002

Die Arbeit wurde gefördert von der Christoph-Dornier-Stiftung für Klinische Psychologie (CDS) und von der Deutschen Forschungsgemeinschaft (Aktenzeichen Ma-1116-8)

Danksagungen

Wie es der Brauch will, möchte ich zu Anfang denen danken, die mich bei dieser Dissertationsarbeit unterstützt haben.

Geistig-moralische und materielle Hilfe erhielt ich vor allem von meiner Familie, meinen engen Kolleginnen und Kollegen und weiteren Freunden, die wissen, was „Dizz“ bedeutet. Und von Prof. Dr. Jürgen Margraf (u.a. Realisierung des DFG-Projektes).

Besonders hervorheben möchte ich außerdem meine sehr guten DiplomandInnen Dipl.-Psych. Uta Gebel, Dipl.-Psych. Katja Pieper, Dipl.-Psych. Gisela Runde, Dipl.-Psych. Christian Sekot und Dipl.-Psych. Steffi Sommer. Der technische Abschluss der Arbeit wurde gefördert von cand.-psych. Christine Fröhlich und von cand.-psych. Johannes Hacker.

Außerdem bin ich dem "Mondfisch" zu Dank verpflichtet; dort hat die Arbeit wichtige Momente erlebt.

Herzlich möchte ich mich bei der Christoph-Dornier-Stiftung für Klinische Psychologie bedanken, also bei Prof. Dr. Wolfgang Fiegenbaum (Vorstand), Prof. Dr. Kurt Hahlweg (stellvertretend für das Kuratorium) und den Institutsleiterinnen Dr. Silvia Schneider, Dr. Monika Frank, Dr. Brigitte Schröder, Dipl.-Psych. Gabi Fünfstück und Dipl.-Psych. Ines v. Witzleben. Ohne die durch diese Personen geschaffenen Rahmenbedingungen wäre diese Arbeit nicht durchführbar gewesen.

Diese Arbeit hatte mehrere Verzögerungsmomente. Als MotivatorInnen für den Abschluss der Arbeit halfen mir letztendlich besonders Prof. Dr. Uli Wittchen, PD Dr. Jürgen Hoyer, Dr. Anne Ullrich, sowie meine Mutter, Dipl.-Psych. Brigitte Jacobi, und, wenn man so will, mein Vater Prof. Dr. Peter Jacobi († 1989).

Dresden im Juli 2001, Dipl.-Psych. Frank Jacobi

Vorbemerkung zur Einführung in das Thema

Gesundheit gilt als höchstes Gut – insbesondere wenn Krankheit Leiden und Beeinträchtigung bei den Betroffenen und deren Angehörigen verursacht. Auch über diese subjektive Perspektive hinaus stellt Gesundheit einen hohen Wert dar, z.B. wenn die Kosten hoch sind, die durch eine Krankheit oder eine Störung verursacht werden. Dies gilt gleichermaßen für somatische wie für psychische Erkrankungen.

„*Gesund ist nur, wer noch nicht eingehend untersucht wurde.*“ Dieser Aphorismus des ehemaligen Gesundheitsministers Seehofer betrifft Diagnostik und Behandlungsbedürftigkeit von „Krankheit“. Psychische Störungen sind, wenn man genauer hinschaut, häufiger, als man gemeinhin annimmt: die 12-Monats-Prävalenz liegt bei Erwachsenen zwischen 18 und 65 Jahren etwa bei 33%. Auch wenn die *awareness* gegenüber psychischen Störungen in den vergangenen Jahrzehnten gestiegen ist, werden immer noch viele ernstzunehmende psychische Störungen häufig nicht als solche erkannt. Andererseits wird man bei jedem, wenn man nur lange genug sucht, im Laufe seines Lebens Zeiten finden, in der er oder sie z.B. durch Ängste oder depressive Stimmung belastet war – aber nicht in allen Fällen (auch nicht immer bei Vorliegen einer Diagnose) ist deswegen eine Behandlung indiziert.

Schwere psychische Störungen, wie z.B. viele Angststörungen, sind behandlungsbedürftig. Die Patienten, um die es in dieser Arbeit geht, erfüllten allerdings alle die internationalen Kriterien für eine oder mehrere Diagnosen, waren in der Regel schwer belastet und beeinträchtigt und auch behandlungsbereit. Es darf angenommen werden, dass von den geschätzt 12 Millionen Betroffenen unter den Erwachsenen, die zwar die Diagnose einer psychischen Störung aufweisen, aber aktuell *keinerlei* Behandlung deswegen erhalten, zumindest mehr in diesem Sinne behandlungsbedürftig sind, als überhaupt Behandlungsmöglichkeiten zur Verfügung stehen.

Klinisch-psychologische Behandlungsmethoden sind bei behandlungsbedürftigen Störungen wirksam und kosten-effektiv. Wie in dieser Arbeit anhand der kognitiven Verhaltenstherapie für Angststörungen gezeigt wird, kann moderne Psychotherapie bei vertretbarem Aufwand beachtliche Erfolge verbuchen.

Klinisch-psychologische Behandlungsmethoden sind darüber hinaus auch wirtschaftlich, derart, dass sie über den reinen Behandlungserfolg hinaus zu Kosteneinsparungen beitragen könnten. Es ist erstaunlich, dass diese Perspektive der wissenschaftlichen Bewertung von Psychotherapie noch recht neu ist. Die Klinische Psychologie kann davon profitieren, diese Dimension (neben den beachtlichen Erkenntnisfortschritten in bezug auf Klassifikation, Ätiologie, Verlaufsforschung und Behandlungsverfahren) mehr herauszustellen.

Inhaltsverzeichnis

1	Einführung	1
1.1	Anlass der Arbeit: Warum Kostenstudien in der Klinischen Psychologie?	1
1.2	Angststörungen und klinisch-psychologische Behandlungsmethoden: Überblick	5
1.2.1	Panikstörungen mit Agoraphobie sind schwere Angsterkrankungen	5
1.2.2	Panikstörungen mit Agoraphobie sind häufig (und meist komorbid)	6
1.2.3	Panikstörungen mit Agoraphobie haben einen ungünstigen Spontanverlauf	8
1.2.4	Panikstörungen mit Agoraphobie sind behandlungsfähig	10
1.2.5	Klinisch-Psychologische Behandlungsverfahren: Begriffsbestimmung	12
1.3	Kostenanalysen: Begriffsklärung	16
1.3.1	Definition von „Kosten“ und „Nutzen“	16
1.3.2	Kosten-Berechnung: Betriebskosten- vs. umfassender Ansatz	18
1.3.3	Nutzen-Berechnung	20
1.3.4	Analyse-Typen	20
1.4	Ethik von Kostenanalysen im Gesundheitswesen	24
1.5	Fallbeispiele zur Illustration unterschiedlicher Kosten-Nutzen-Indices bei gleichem Outcome	25
1.5.1	Exzellentes Kosten-Nutzen-Verhältnis	25
1.5.2	Durchschnittliches Kosten-Nutzen-Verhältnis	26
1.5.3	Schlechtes Kosten-Nutzen-Verhältnis	26
2	Theorie und Forschungsstand	28
2.1	Outcome-Forschung im Bereich Psychotherapie	28
2.1.1	Welche Therapie ist erwiesenermaßen wirksam?	28
2.1.2	Von der Wirksamkeits- zur Effektivitätsforschung	30
2.2	Kosten-Effektivitäts-Analysen	36
2.2.1	Allgemeines Vorgehen bei einer Kosten-Effektivitäts-Analyse	36
2.2.2	Therapieerfolgsmaße und Therapieerfolgskriterien	38
2.3	Kosten-Nutzen-Analysen	49
2.3.1	Richtlinien zur Durchführung von Kosten-Nutzen-Analysen	49
2.3.2	Forschungsstand zu Kosten-Analysen im Psychotherapiebereich (insbes. Angststörungen)	52

2.4	Einbezug des subjektiven/intangiblen Nutzens in eine Kosten-Nutzen-Analyse	65
2.4.1	Die Zahlungsbereitschaft (willingness to pay, WTP)	65
2.4.2	Die Ermittlung der Zahlungsbereitschaft	67
2.4.3	Einbezug der WTP in Kosten-Nutzen-Analysen	73
2.4.4	Forschungsstand zur Methode der Zahlungsbereitschaft	73
3	Kosten-Effektivitäts-Analyse	77
3.1	Fragestellungen	77
3.2	Methoden	82
3.2.1	Behandlungskosten	82
3.2.2	Verwendete Effektivitätsmaße	82
3.2.3	Effektivitätsberechnung und Kriterien klinischer Signifikanz	85
3.2.4	Darstellung der Kosten-Effektivität	92
3.3	Rahmenbedingungen und Stichprobe	93
3.3.1	Therapieform, Therapeuten, klinische Diagnostik	93
3.3.2	Stichprobe	94
3.4	Kosteneffektivitätsstudie: Ergebnisse	98
3.4.1	Kosten	98
3.4.2	Effektivität (1) – globale Therapieerfolgseinschätzung durch Patienten und Therapeuten	99
3.4.3	Effektivität (2) – Störungsfolgen: Beeinträchtigungsebene	103
3.4.4	Effektivität (3) – Fragebogenmaße	112
3.5	Variabilität des Kosten-Effektivitäts-Index in Abhängigkeit von Berechnungsmethoden und (Sub-) Stichproben	124
3.5.1	Messebene	124
3.5.2	Therapie-Preis	125
3.5.3	Störungsspezifische Variablen	126
3.5.4	Soziodemographische Variablen	128
3.5.5	Variabilität der Kosten-Effektivitäts-Indices: Zusammenfassung	129
3.6	Überprüfung von Selektionseffekten durch Dropouts	131
3.6.1	Unterscheiden sich nachuntersuchte und nicht nachuntersuchte Patienten?	131
3.6.2	Hinweise auf Unsystematik von Drop-Outs aufgrund des Settings	135
3.7	Einordnung der Effekte in den Forschungsstand	137
4	Kosten-Nutzen-Analyse	139
4.1	Fragestellungen	139

4.2	Methoden	141
4.2.1	Berechnung der Behandlungskosten	141
4.2.2	Nutzenberechnung I: Bestimmung der Kostenfaktoren vor und nach der Behandlung	141
4.2.3	Nutzenberechnung II: Umrechnung der Kostenfaktoren in monetäre Einheiten	143
4.2.4	Sensitivitätsanalyse	146
4.2.5	Diskontierung	146
4.2.6	Marginalanalyse	147
4.3	Rahmenbedingungen und Stichprobe	148
4.3.1	Soziodemografische Variablen	148
4.3.2	Störungsspezifische Variablen	149
4.4	Ergebnisse	151
4.4.1	Der Nutzen aus der Perspektive volkswirtschaftlicher Einsparungen	151
4.4.2	Weitere Posten	155
4.4.3	Kosten-Nutzen-Index	159
4.4.4	Sensitivitätsanalyse	160
4.5	Die Perspektive des Patienten: Einbezug der Zahlungsbereitschaft (WTP)	163
4.5.1	Fragen zur Validität der Methode	163
4.5.2	Methoden der Datenerhebung: Auswahl der hypothetischen Situation und Umsetzung der NOAA-Richtlinien	164
4.5.3	Stichprobe	167
4.5.4	Ergebnisse	168
4.6	Vereinfachung in Durchführung und Auswertung von Kosten-Nutzen- Analysen im Psychotherapiebereich	174
4.6.1	Querschnitt statt Längsschnitt?	174
4.6.2	Ein Programm zur Vereinfachung von Kosten-Nutzen-Berechnungen: Der „Cost-Benefit-Calculator“	175
5	Zusammenfassungen und Diskussion	179
5.1	Kosten-Effektivitäts-Analyse: Zusammenfassung	179
5.2	Sind klinisch-psychologische Behandlungen modernen Zuschnitts effektiv?	181
5.2.1	Allgemeine Bewertung der Effektgrößen und der Besserungsraten	181
5.2.2	Berücksichtigung einer „subklinischen“ Kategorie	183
5.2.3	Anregungen für die Therapieerfolgsmessung	184
5.2.4	Validität und Design:	186
5.2.5	Forderungen an eine verbesserte standardmäßige zukünftige Forschung	189
5.2.6	Fazit Therapieerfolg: Klinisch-psychologische Behandlungsmethoden sind wirksam und effektiv, kommen aber in der Versorgungsrealität häufig nicht zum Zuge	190

5.3	Spezielle Aspekte der Kosten-Effektivitäts-Analyse	192
5.3.1	Therapiekosten und aufgewendete Ressourcen	192
5.3.2	Die Kosten-Effektivität als Erweiterung klassischer Outcome-Forschung	
5.3.3	Optimierung der Kosten-Effektivität einer Behandlung	193
5.3.4	Variabilität des Kosten-Effektivitäts-Index in Abhängigkeit von Auswertungsmethode und (Sub-) Stichproben	195
5.4	Kosten-Nutzen-Analyse: Zusammenfassung	198
5.5	Spezielle Aspekte der Kosten-Nutzen-Analyse	200
5.5.1	Validität bezüglich Design und Datenerhebung	200
5.5.2	Vereinfachung von Datenerhebung und Berechnung	203
5.5.3	Der Einbezug der Zahlungsbereitschaft (WTP) zur Ermittlung intangiblen Nutzens	204
5.6	Einordnung der Ergebnisse in den Forschungsstand zu Kosten–Nutzen–Analysen in der Klinischen Psychologie	207
6	Allgemeine Diskussion zu Kosten-Studien und Ausblick	210
	Literatur	217
Anhang:	Erklärungen und eingesetzte Fragebogeninstrumente	

1 Einführung

In diesem ersten Teil wird zunächst einführend auf die Rolle von Kostenstudien für die Klinische Psychologie und Psychotherapie eingegangen.

Die Ziele dieser Arbeit sind, einen Überblick über bislang nur defizitär untersuchter Aspekte der Kostenseite psychischer Störungen und deren Behandlung mit klinisch-psychologischen Methoden zu geben, sowie eigene empirischer Kosten-Studien zu präsentieren; deshalb sind die Ausführungen zu den Konzepten psychischer Störungen in der Klinischen Psychologie, zu moderner klassifikatorischen Diagnostik und zur Charakteristik von Angststörungen und der wichtigsten psychotherapeutischen Ansätze knapp gehalten.

Ausführlicher ist die Begriffsklärung aus dem Bereich von Kostenstudien im allgemeinen, da hier ein entsprechendes Vorwissen bei klinischen Psychologen nicht immer vorausgesetzt werden kann.

Nach einem Exkurs zu ethischen Aspekten von Kostenstudien im Gesundheitswesen sollen drei Fallbeispiele das Zusammenspiel von Therapieerfolg aus Patienten- und Kostenperspektive und der dafür aufgewendeten Ressourcen in einer Kosten-Nutzen-Analyse illustrieren.

1.1 Anlass der Arbeit: Warum Kostenstudien in der Klinischen Psychologie?

Psychische Störungen werden leider vielerorts immer noch tabuisiert und von vielen Menschen eher als persönlicher Makel oder als Ausdruck von Schwäche und Versagen denn als behandlungsbedürftige und behandlungsfähige Erkrankungen angesehen. Dabei wissen wir seit vielen Jahren, dass fast jeder zweite Mensch im Verlauf seines Lebens zumindest einmal an einer psychischen Störung so schwer erkrankt, dass es zu schwerwiegenden Folgen für seine Arbeitsfähigkeit, die Familie, den Partner und für die Lebensqualität kommt. Psychische Störungen gehören zu den häufigsten Erkrankungsformen in unserer Gesellschaft und sind mindestens ebenso belastend wie andere Formen körperlicher Krankheiten (vgl. z.B. Jenkins, 1997, zur großen „burden of disease“-Studie der Harvard Medical School, der Weltgesundheitsorganisation und der Weltbank).

Moderne klinisch-psychologische Behandlungsmethoden basieren auf Erkenntnissen der empirischen wissenschaftlichen Psychologie und zielen neben der Behandlung psychischer Störungen auch auf die günstige Beeinflussung psychischer Aspekte somatischer Erkrankungen ab (Baumann & Perrez,

1998). Patienten mit psychischen Störungen profitieren von klinisch-psychologischen Behandlungsmethoden empirisch abgesichert in weit über der Hälfte der Fälle ähnlich gut und oft nachhaltiger als von medikamentöser Behandlung, und die positiven Wirkungen psychotherapeutischer Behandlung geht in den meisten Fällen über die bloße Beseitigung oder Besserung der Störung hinaus (Grawe, 1992; Grawe, Donati & Bernauer, 1994).

Angesichts dieser Befundlage kann mit Grawe et al. (1994) insbesondere für die mit Abstand am besten untersuchten Behandlungen aus dem kognitiv-behavioralen Spektrum gefordert werden, in der Gesundheitsversorgung eine prominente Rolle zu spielen. Dass dem nicht so ist, liegt zu einem erheblichen Teil an Kostengründen: die – sei es aus „echter“ gesundheitspolitischer Besorgnis oder aus berufspolitischen Gründen (Nuber, 1992) geäußerte – Befürchtung, dass durch ein flächendeckendes Angebot von psychologischer Psychotherapie als Regelleistung die Kosten für das Gesundheitswesen dramatisch in die Höhe getrieben würden, führte in Deutschland zu einer strikten Deckelung der Ausgaben in diesem Bereich. Die Gesamtausgaben für Gesundheitsleistungen betragen 1998 etwa 250 Milliarden DM (Statistisches Bundesamt, 1998). Das Budget für die ambulante Psychotherapie betrug im ersten Jahr nach Inkrafttreten des neuen Psychotherapeutengesetzes etwa 1,5 Milliarden DM. Damit ist klar, dass Psychotherapie nur einen winzigen Bruchteil (0,6%) des Gesamtbudgets ausmacht; und die kognitiven Verhaltenstherapie macht in diesem Teilbudget wiederum lediglich nur einen kleinen Teil aus. Angesichts dieser Zahlen können bei der Frage „Kostenlawine oder Kostenentlastung durch Psychotherapie?“ in keiner Richtung dramatische Effekte erwartet werden.

Welche Bedeutung hat also die Kostenfrage für die Klinische Psychologie und Psychotherapie, nachdem einerseits die positiven Auswirkungen klinisch-psychologischer Behandlung bereits eindrucksvoll dokumentiert werden konnten, andererseits aber die gesundheitspolitische Dimension angesichts des geringen Budgets so zurechtgerückt wurde?

Die Kostenaspekte einer Behandlung sind grundsätzlich wichtige Beurteilungskategorien, und dies nicht nur in Zeiten des knappen Geldes:

- In der derzeitigen Situation, in der bemerkenswert wenige Befunde zu den tatsächlichen finanziellen Auswirkungen psychischer Störungen und ihrer Behandlung vorliegen, ist es besonders wichtig, exakte Daten über die Kosten zu erheben, die Patienten mit psychischen Störungen dem Gesundheitswesen verursachen. Auch die finanziellen Effekte einer psychologischen Behandlung solcher Störungen sollten einer kritischen Analyse unterzogen werden. Für die Wissenschaft bzw. für die Psychotherapieforschung beleuchten Kostenstudien den bislang vernachlässigten Aspekt der Ressourcen, die für psychotherapeutische Behandlungen aufgewendet wurden und schaffen somit mehr Transparenz. Auch der systematische

Einbezug von Kostenaspekten bei der Evaluation psychotherapeutischer Maßnahmen liefert per se einen Erkenntnisgewinn und ist deshalb zu begrüßen (Bühringer & Hahlweg, 1986), zumal der Legitimationsdruck für die Klinische Psychologie, die bislang Therapieverfahren fast ausschließlich unter der Perspektive ihrer Wirksamkeit untersucht hat, mit den berufs- und sozialrechtlichen Neuerungen des Psychotherapeutengesetzes erheblich gestiegen ist.

- Psychotherapie ist – wie jede andere Krankenbehandlung auch – eine Dienstleistung, und wer Dienstleistungen „kauft“, „verkauft“ oder in anderer Weise mit ihrer Verbreitung zu tun hat, sollte über den Preis informiert sein. Informationen zu Kosten, Effektivität und Nutzen helfen bei der Entscheidung über die Einleitung und die Dauer einer Behandlung, über evtl. anschließende Therapien sowie über die Auswahl zwischen konkurrierenden Verfahren.
- Gesamtgesellschaftlich sind Kosten-Nutzen-Fragen wesentliche Hilfen bei Entscheidungen über die Verteilung von Ressourcen; sie sind somit in politische und ethische Entscheidungen mit einzubeziehen.

Selbst bei einem so hohen Gut wie der Gesundheit und bei selbst angesichts der eindeutig positiven Befundlage zu Wirksamkeit und Effektivität klinisch-psychologischer Behandlungsmethoden kann also auf die „Niederungen“ der Kostenanalyse nicht verzichtet werden. Insbesondere für die Verhaltenstherapie mit ihrem Anspruch auf Transparenz und Wissenschaftlichkeit (Margraf & Lieb, 1995) sollte das Kriterium der Kosten-Nutzen-Relation, das in anderen Lebensbereichen grundlegend ist, als selbstverständlich gelten.

Es gibt bereits eine Anzahl von derartigen Kostenstudien, allerdings noch kaum empirische Ergebnisse für den ambulanten Bereich. Darüber einen Überblick zu geben, war ein erstes Anliegen dieser Arbeit; in den Theorieteilen finden sich entsprechende Verweise auf die relevante Literatur.

Der Hauptteil dieser Dissertation beschäftigt sich mit Erhebung und Auswertung eigener empirischer Kosten- und Outcome-Daten für die ambulante psychologische Behandlung von Angststörungen. Besonders hervorzuheben ist dabei das klinisch relevante Setting, da eindringlich gefordert wird, die klassischen kontrollierten und standardisierten Wirksamkeitsstudien (hohe interne Validität) um klinisch relevante, versorgungsnaher Effektivitätsstudien (hohe externe Validität) zu ergänzen (z.B. Seligman, 1995; Shadish et al., 1998).

Ein weiterer Aspekt der vorliegenden Arbeit ist der Vergleich von monetären und psychologischen bzw. subjektiven Kosten und Nutzen. Dies ist ein mit verschiedenen wissenschaftstheoretischen, weltanschaulichen und methodischen Problemen behaftetes Unterfangen. Betrachten wir eine Aussage wie:

„Die Verbesserung der Lebenszufriedenheit bei gleichzeitiger Abnahme störungsbedingter Symptome und Beeinträchtigungen durch eine erfolgreiche Therapie kostet 6000.- DM.“

Einen Preis für eine bestimmte Lebenszufriedenheit anzugeben, könnte manchem etwas befremdlich anmuten. Es entspricht aber dem Denken vieler Menschen, die die Gesundheit als höchstes Gut betrachten, für das auch große Ressourcen bzw. Geldbeträge aufgewendet werden sollten. Um einer Umrechnung von Geldeinheiten und psychischem Wohlbefinden zumindest nahezukommen, werden in dieser Arbeit also neben den „objektiven“ wirtschaftlichen Aspekten der Psychotherapie (Einsparung für Kostenträger oder Volkswirtschaft) auch subjektiver Gesundheit und individuell erlebten Kosten und Nutzen von Störung und Therapie untersucht.

Damit unsere Analysen möglichst Beispielcharakter gewinnen, sollen sie anhand der häufigsten psychischen Störungen, der Angststörungen, erfolgen. Hier wiederum wird ein Störungsbild, die Panikstörung mit Agoraphobie gesondert herausgegriffen; ein Überblick zu Störungsbild und Behandlung findet sich im folgenden Abschnitt.

1.2 Angststörungen und klinisch-psychologische Behandlungsmethoden: Überblick

Niemand ist gegen seelisches Leiden gefeit. Es gibt nur wenige Familien ohne Familienmitglieder, die mit Depressionen oder Angst zu kämpfen haben, emotionale Krisen durchmachen, Drogen- oder Verhaltensprobleme haben, unter sexuellen Funktionsstörungen, Schlaflosigkeit, psychotischen Symptomen, Demenz oder irgendeiner anderen der zahlreichen Varianten von psychischen Störungen leiden (s. z.B. Robins & Regier, 1991; Wittchen & Perkonig, 1996). Die vorliegende Arbeit befasst sich mit der Untersuchung der Behandlung von Angststörungen im Sinne von Störungen mit Krankheitswert, deren internationale diagnostische Kriterien in den modernen klassifikatorischen Diagnosesystemen der vierten Version des *Diagnostic and Statistical Manual of Mental Disorders* der American Psychiatric Association (DSM-IV; APA, 1993) bzw. der zehnten Revision der *International Classification of Diseases* der Weltgesundheitsorganisation (ICD-10; WHO, 1994) niedergelegt sind.

Da Angststörungen eine außerordentlich heterogene Gruppe psychischer Störungen mit sehr unterschiedlichen Vulnerabilitäten, Risiko- und Verlaufstendenzen umfassen, wird im folgenden die Panikstörung mit Agoraphobie herausgegriffen – nicht zuletzt weil im empirischen Teil dieser Arbeit der Großteil der untersuchten Patienten diese Diagnose aufweist. Der folgende Überblick bezieht sich auf Beeinträchtigung, Epidemiologie, Spontanverlauf und Behandelbarkeit dieser Störung. Außerdem soll begriffsklärend auf den Bereich klinisch-psychologischer Behandlungsmethoden, Psychotherapie, Verhaltenstherapie etc. im allgemeinen eingegangen werden.

1.2.1 Panikstörungen mit Agoraphobie sind schwere Angsterkrankungen

Die Panikstörung mit Agoraphobie (DSM-IV: 300.21) ist einerseits gekennzeichnet durch wiederkehrende unerwartete Panikattacken, bei denen mindestens vier abrupt auftretende Angstsymptome innerhalb von zehn Minuten ihren Höhepunkt erreichen, die nicht auf die direkte körperliche Wirkung einer Substanz oder eines medizinischen Krankheitsfaktors zurückgeführt werden können. Diese Panikattacken führen zu anhaltender Besorgnis über das Auftreten weiterer Panikattacken oder mögliche katastrophale Konsequenzen, oder zu deutlichen Verhaltensänderungen infolge der Attacken. Außerdem werden zumindest einige Angstanfälle durch bestimmte Situationen ausgelöst und diese infolgedessen vermieden oder nur unter starker Angst ertragen. Agoraphobische Vermeidung bezieht sich in der Regel auf Situationen, in denen es be-

sonders unangenehm oder gefährlich sein könnte, einen Angstanfall zu haben (z.B. Autofahren, Kaufhäuser, Supermärkte, Fahrstühle, Menschenmengen, allein das Haus verlassen, Schlange stehen, Reisen mit öffentlichen Verkehrsmitteln, Flugzeuge, Kinos, Theater). Das gemeinsame dieser Situationen ist, dass die betreffende Person fürchtet, im Falle eines Angstanfalls nicht schnell genug aus der Situation flüchten zu können bzw. dass Hilfe nicht schnell genug erreichbar ist.

Im Laufe der Zeit können „bei perfekter Vermeidung“ die Angstanfälle völlig verschwinden. Manchmal können die gefürchteten Situationen unter extremer Angst ertragen werden. Bei der Agoraphobie ohne Paniksyndrom werden dieselben Situationen aus anderen Gründen vermieden (z.B. Angst vor plötzlichem Durchfall oder exzessive Angst im Zusammenhang mit körperlichen Krankheiten). Einzelne Situationen können auch von spezifischen oder Sozialphobikern vermieden bzw. gefürchtet werden. Agoraphobiker vermeiden jedoch mehr Situationen und befürchten v.a. Angstanfälle bzw. deren katastrophale Folgen, spezifische Phobiker dagegen in der Regel unmittelbar von einem einzelnen speziellen phobischen Objekt ausgehende Gefahren (z.B. Flugzeugabsturz, Hundebiss) und Sozialphobiker eine Blamage bzw. negative Bewertung durch andere.

Die Panikanfälle werden von den Betroffenen als außerordentlich belastend erlebt (bis hin zur Todesangst), und das Vermeidungsverhalten führt oft zu erheblichen Beeinträchtigungen im familiären, sonstigem sozialen und beruflichen Bereich (Binder und Angst, 1981; Leon et al., 1995). Darüber hinaus zeigen Menschen mit Angsterkrankungen einen deutlich erhöhten Alkoholkonsum und deutlich erhöhte Suizidalität. Die eingeschränkte Mobilität kann als Behinderung so weit gehen, dass die eigene Wohnung nicht mehr (alleine) verlassen werden kann. Auch der Befund, dass sich bei dieser Störung häufig sekundäre Depressionen entwickeln, mag für den hohen Schweregrad von Panikstörung und Agoraphobie sprechen.

1.2.2 Panikstörungen mit Agoraphobie sind häufig (und meist komorbid)

Relevant im Zusammenhang mit der vorliegenden Arbeit sind vor allem Häufigkeit und Verlauf von Angststörungen, da dadurch die Störungskosten erheblich moderiert werden.

Bei Frauen stehen Angsterkrankungen bezüglich ihrer Häufigkeit an erster Stelle der psychischen Erkrankungen, bei Männern sind sie nach Abhängigkeitssyndromen die zweithäufigste Erkrankung. Tabelle 1.1 zeigt aktuelle für die erwachsene Bevölkerung (18-65 Jahre) repräsentative Prävalenzdaten aus dem Zusatzmodul „Psychische Störungen“ des Bundesgesundheits surveys 1998/99 (Bellach, 1998; Jacobi, Jahn & Wittchen, 2001; Wittchen et al., 1998,

1999). Die Diagnosen wurden dabei mit dem für Angststörungen sehr reliablen standardisierten CIDI-Interview ermittelt (Wittchen & Pfister, 1997).

Die Diagnose für Panikstörung mit Agoraphobie erfüllten 1.2% der untersuchten Personen, wobei die Störung bei Frauen nahezu doppelt so häufig auftritt wie bei Männern. Hochgerechnet entspräche dies knapp 600000 Betroffenen in der Bundesrepublik im 1-Jahres-Zeitraum.

Komorbidität lässt sich definieren als das Auftreten von mehr als einer spezifisch diagnostizierbaren psychischen Störung bei einer Person in einem definierten Zeitintervall. Wittchen und Vossen (2000) geben einen Überblick über die komplexen Komorbiditätsstrukturen bei Angststörungen. Der überwiegende Teil der Panikpatienten ist auch bei der Wahl eines kleinen Zeitfensters komorbid (Wittchen, 1991). Insbesondere aber bei großem Zeitfenster (Lifetime-Komorbidität) ist die Komorbiditätsrate bei Angststörungen außerordentlich hoch, insbesondere mit anderen Angststörungen (Magee et al., 1996; s. auch Tabelle 1.2) und depressiven Störungen (Merikangas et al., 1996).

Tabelle 1.1: Prävalenzen von Angststörungen in Deutschland laut Bundesgesundheits-survey 1998/99 (Zusatzmodul „Psychische Störungen“); N=4181 (erwachsene Wohnbevölkerung 18-65 Jahre)

Störung bzw. Störungsgruppe	12-Monats- Prävalenz: % (95%-KI)	12-Monats- Prävalenz: % (95%-KI)	12-Monats- Prävalenz: % (95%-KI)
	Männer	Frauen	gesamt
Angststörungen:			
irgendeine Angststörung (einschl. Zwangsstörung, ohne PTSD) (N=604)	9.20 (7.98-10.58)	19.77 (18.12-21.54)	14.45 (13.42-15.55)
Agoraphobie (ohne Panikstörung) (N=84)	0.98 (0.64-1.49)	3.07 (2.44-3.86)	2.02 (1.65-2.47)
Panikstörung (ohne Agoraphobie) (N=48)	0.83 (0.54-1.29)	1.45 (1.04-2.04)	1.14 (0.88-1.49)
Panikstörung mit Agoraphobie (N=50)	0.83 (0.48-1.45)	1.57 (1.14-2.17)	1.20 (0.90-1.60)
Soziale Phobie (N=83)	1.32 (0.88-1.96)	2.65 (2.05-3.43)	1.98 (1.59-2.46)
Generalisierte Angststörung (N=63)	0.99 (0.64-1.53)	2.05 (1.52-2.77)	1.52 (1.18-1.94)
Irgendeine spezifische Phobie (N=319)	4.51 (3.65-5.55)	10.78 (9.52-12.18)	7.63 (6.85-8.48)
Zwangsstörung (N=30)	0.55 (0.32-0.96)	0.89 (0.58-1.36)	0.72 (0.51-1.01)
Angststörung NNB (N=143)	1.93 (1.42-2.61)	4.90 (4.05-5.92)	3.41 (2.90-4.00)

Tabelle 1.2: Lifetime-Komorbidität von Angststörungen im U.S.-amerikanischen *National Comorbidity Survey* (Wittchen et al., 1994; Magee et al., 1996): Prozentsatz der Personen mit einer Störung, die auch eine der in der linken Spalte aufgeführten Störungen aufwies. Die Angaben in Klammern geben die Odds Ratios für das entsprechend erhöhte Risiko an.

Diagnose nach CIDI/DSM-III-R	Agoraphobie	Spezifische Phobie	Soziale Phobie	Generalisierte Angststörung
Agoraphobie	-	27,0% (8,9)	23,3% (7,1)	25,7% (5,8)
Spezifische Phobie	45,6% (8,7)	-	37,6% (7,8)	35,1% (4,9)
Soziale Phobie	46,5% (7,1)	44,5% (7,8)	-	34,4% (3,4)
Gen. Angststörung	19,8% (5,8)	16,0% (4,9)	13,3% (3,8)	-
Panikstörung	21,6% (11,9)	14,8% (7,8)	10,9% (4,8)	23,5% (12,3)
Panikattacken	35,8% (10,0)	27,0% (7,3)	20,7% (4,7)	41,5% (8,9)

Auch wenn es aus der Perspektive der Versorgungsforschung nicht adäquat ist, das Vorliegen einer Diagnose automatisch mit „Behandlungsbedarf“ gleichzusetzen (Andrews & Henderson, 2000; Wittchen, 2000), so ist angesichts der Häufigkeit der Diagnose und der erwähnten starken Belastungen und Beeinträchtigungen, die im allgemeinen mit dieser Diagnose einhergehen von einem enormen Behandlungsbedarf auszugehen, der angesichts der derzeitigen Versorgungslage im Bereich der Psychotherapie kaum jemals abgedeckt werden kann.

1.2.3 Panikstörungen mit Agoraphobie haben einen ungünstigen Spontanverlauf

Angaben über den „natürlichen Verlauf“ von unbehandelten Angsterkrankungen sind z.T. älteren Datums und damit insofern nicht valide, da sich seitdem die Diagnosekriterien weiterentwickelt haben (z.B. Paul, 1967, Remissionsrate um 18%; Noyes & Clancy, 1976, ca. 67%). Durch Unterschiede der Patientenselektion, der diagnostischen Kategorisierung, der Besserungskriterien, der Berücksichtigung verschiedener Einflussgrößen (z.B. Symptombelastung vs. soziales Funktionsniveau) und unterschiedlicher Nachuntersuchungsdesigns ist eine Interpretation solch widersprüchlicher Befunde erheblich erschwert.

In der *Münchener Follow-up-Studie* (MFS; u.a. Wittchen & v. Zerssen, 1987), ein zehnjähriges Forschungsprogramm des Max-Planck-Instituts für Psychiatrie, bei dem der Langzeitverlauf ehemals stationär behandelter psychiatrischer Patienten und der „Spontanverlauf“ weitgehend unbehandelt gebliebener Personen aus der Durchschnittsbevölkerung der BRD über einen Zeitraum von durchschnittlich sieben Jahren (1974 bzw. 1981) vergleichend untersucht wurde, ergab sich für die Störung ein ungünstiger Spontanverlauf. Nur etwa 14% der Betroffenen remittierten ohne Behandlung komplett (Witt-

chen, 1988; 1991). Auch die Befunde der von der Erhebungsmethodik her vergleichbaren *Epidemiologic Catchment Area Studie* (ECA; u.a. Robins & Regier, 1991) des U.S.-amerikanischen National Institute of Mental Health (NIMH), unterstützen deutlich die Hypothese der niedrigen Spontanremission: die Differenz zwischen Lebenszeit- und Punktprävalenz ist deutlich niedriger als für andere Störungsgruppen (Robins et al., 1984; Myers et al., 1984).

Auch die prospektive Münchner *Early Developmental Stages of Psychopathology*-Studie (EDSP; Wittchen et al., 1998), in der Jugendliche und junge Erwachsene über einen Verlauf von fünf Jahren umfassend psychopathologisch untersucht wurden, liefert Befunde zum Verlauf von Angststörungen. Angststörungen insbesondere phobischer Prägung manifestieren sich früh, haben einen nachgewiesenen Einfluss auf Sozialisation und Entwicklung, sowie einen persistierenden, wenn auch fluktuierenden Verlauf mit seltenen kompletten Spontanremissionen. Sie stellen markante Risikofaktoren für die Entwicklung weiterer Störungen, insbesondere sekundärer Depression dar (Wittchen et al., 1999).

Vor dem Hintergrund dieser neueren Ergebnisse kann die bekannte, von Eysenck (1952) erstmals vorgebrachte und z.B. auch von Jablensky und Hugler (1982) unterstützte These, dass mehr als ein Drittel aller Neurosen von begrenzter Dauer (bis zu vier Monaten) sind und weniger als ein Drittel der sogenannten neurotischen Erkrankungen länger als drei Jahre persistieren, mittlerweile insbesondere für Angststörungen als völlig überholt gelten.

Auch die langen Krankheitsdauern der im empirischen Teil dieser Arbeit untersuchten Angstpatienten sprechen für einen chronischen Verlauf: lediglich ein Viertel der Patienten hatte die Störung seit weniger als 3 Jahren; bei einem weiteren Viertel betrug die Krankheitsdauer 3 bis 6, bei einem weiteren 7 bis 13 und beim letzten Viertel über 14 Jahre. Im Rahmen dieses Dissertationsprojekts wurde in einer Erhebung an einer Unterstichprobe, für die sehr genaue Angaben zur Häufigkeit und zum Zeitpunkt der Inanspruchnahme von Gesundheitsleistungen und zur Arbeitsunfähigkeit vorlagen, der Verlauf untersucht, wie diese Kostenfaktoren im Rahmen der bisherigen Behandlungsgeschichte anfielen (Sommer, 1999). Dabei wurde der Zeitraum zwischen Störungsbeginn und Diagnostik für jeden Patienten gedrittelt und die durchschnittliche Häufigkeit der Kostenfaktoren für jeden der drei Teile ermittelt und auf signifikante Differenzen untersucht.

In Tabelle 1.3 ist das Ergebnis dargestellt. In keiner der untersuchten Variablen der Behandlungsgeschichte und der Arbeitsfähigkeit konnte mit dem ein signifikanter Unterschied zwischen den drei Zeitabschnitten gefunden werden. Eine Resignation der Patienten in dem Sinne, dass sie im Zeitverlauf immer weniger Behandlungen in Anspruch nehmen, ist also nicht nachweisbar. Im Gegenteil: tendenziell scheint im Verlauf der Erkrankung die Zahl der Sitzungen in Psychotherapien (von durchschnittlich 7,55 auf 8,60), der Praxiskontakte (von durchschnittlich 19,65 über 21,71 auf 26,75) sowie die Aufenthalts-

dauer in Kliniken (von durchschnittlich 8,00 über 9,15 auf 8,90 Tage) zu steigen, was daran liegen könnte, dass sich die verschiedenen Behandlungsmöglichkeiten den Patienten erst im Zeitverlauf eröffnen oder dass sich bestimmte Behandlungsmethoden erst in letzter Zeit etabliert haben.

Tabelle 1.3: Zeitverlauf von Behandlungsumfang und Arbeitsunfähigkeit bei N=20 Angstpatienten (Panikstörung mit Agoraphobie als Hauptdiagnose); die drei Abschnitte entsprechen jeweils einem Drittel der individuellen Krankheitsdauer

	1. Abschnitt \bar{x}	2. Abschnitt \bar{x}	3. Abschnitt \bar{x}	Signifikanz (Friedmann- Test)
Sitzungen in ambulanten Psychotherapien	7.55 (s= 14.63)	7.55 (s= 13.10)	8.60 (s= 11.14)	0.71
Anzahl weiterer Praxiskontakte	19.65 (s= 51.81)	21.75 (s= 48.91)	26.75 (s= 58.51)	0.26
Aufenthaltsdauer in Kliniken in Tagen	8.00 (s=24.83)	9.150 (s= 28.03)	8.90 (s= 26.37)	0.96
Dauer der Krankschreibung in Tagen	36.15 (s=112.45)	25.25 (s= 61.51)	13.25 (s= 41.49)	0.85

Die Dauer der Krankschreibungen sinkt dagegen von 36 über bzw. 25 auf 13 Tage. Diese Tendenz wurde wahrscheinlich aufgrund hoher Varianz nur wegen der geringen Stichprobengröße nicht signifikant. Eine mögliche zu diskutierende Ursache dieser Entwicklung könnte im Erwerb von Bewältigungsstrategien durch die Patienten liegen. Die entspricht auch dem Befund aus der Münchner Follow-up-Studie, dass die psychosoziale Remissionsrate höher liegt als die der Symptombelastung.

In jedem Fall ist das Ergebnis dahingehend zu bewerten, dass zumindest bei *Angstpatienten*, die sich nach z.T. langer Krankheitszeit zu einer Therapie vorstellen, kein günstiger Verlauf in der Krankheitsgeschichte (im Sinne einer Abnahme von unspezifischen Behandlungen) verzeichnet werden kann.

1.2.4 Panikstörungen mit Agoraphobie sind behandlungsfähig

Das Grundprinzip der heute üblichen Behandlung von agoraphobischem Vermeidungsverhalten, die Konfrontation mit angstausslösenden Situationen, war schon vor der Beschäftigung der Fachwissenschaften mit dem Thema bekannt. Prominente Aussagen zur Wirksamkeit dieser Methoden stammen z.B. von Johann Wolfgang v. Goethe, der verschiedene seiner Ängste selbst damit „be-

handelte“, oder von Sigmund Freud, der feststellte, dass eine psychoanalytische Behandlung bei (Agora-) Phobikern nur dann erfolgreich verlaufen kann, wenn der Betroffene zuvor seine Ängste in der konkret gefürchteten Situation bekämpft hat (vgl. Margraf, 2000).

Die Entstehung der störungsspezifischen klinisch-psychologischen Angstbehandlung ist von der wechselseitigen Entwicklung ätiologischer Modelle und einzelnen Behandlungselementen gekennzeichnet (vgl. 1.2.5). Wenn im folgenden, sowie im Abschnitt 2.1.1 zum Forschungsstand bei der Outcome-Forschung in der Psychotherapie von „empirischer Überprüfung“ die Rede ist, so ist zunächst zu fordern, dass ein methodologisches Regelwerk vorhanden ist, in das man das zu überprüfende Verfahren einzuordnen kann. Bei der Suche nach Gesetzmäßigkeiten müssen Beobachtbarkeit und Operationalisierbarkeit, sowie empirische Testbarkeit und tatsächliche empirische Prüfung gewährleistet sein. Das Vorgehen bei der empirischen Prüfung und die „Zertifizierung“ psychotherapeutischer Verfahren als „empirisch validiert“ ist dabei durchaus umstritten (Kasten 1.1).

In den letzten drei Jahrzehnten wurde die Konfrontationsbehandlung systematisiert und empirisch überprüft (Übersicht bei Marks, 1987; Schneider & Margraf, 1998). Zudem wurden störungsspezifische kognitive Verfahren entwickelt und wissenschaftlich erprobt (Barlow, 1988; Barlow et al., 1989; Clark, 1994; Margraf & Schneider, 1990). Wie auch die monumentale Literaturauswertung von Grawe et al. (1994) eindrucksvoll zeigt, wurden diese kognitiv-behavioralen Verfahren in methodisch überdurchschnittlicher Qualität besonders häufig untersucht. Die Ergebnisse zur massierten Konfrontation zeigen konsistent sehr starke Wirkungen auf die Hauptsymptomatik (Ängste und Vermeidungsverhalten), aber auch auf individuell definierte andere Ziele, allgemeines Wohlbefinden sowie Arbeit und Freizeit, seltener dagegen positive Effekte auf Persönlichkeitsmaße oder Symptome anderer Störungen (sofern überhaupt erhoben).

Die Erfolgsraten in Bezug auf die Hauptsymptomatik schwanken je nach Untersuchungsdesign und Operationalisierung des Therapieerfolgs zwischen 50% und 90%, was bei der Behandlung psychischer Störungen als ausgesprochen günstig zu bewerten ist. Interessanterweise zeigen sich auch keine nennenswerten Risiken und Nebenwirkungen dieser Behandlungsform – ein Thema, das bislang in der bisherigen Psychotherapieforschung im Gegensatz zur medikamentösen Behandlung sehr vernachlässigt wurde (Jacobi, 2001).

Kasten 1.1: Empirisch überprüfte Therapien?

Angesichts der stark ungleichen Forschungsbudgets von Psychotherapien auf der einen Seite und Pharmaindustrie auf der anderen, beschloss die American Psychological Association zu Beginn der 90er Jahre eine „Gegenoffensive“. Die klinisch überprüfte Leistungsfähigkeit psychologischer Verfahren sollte analog der *evidence based medicine* dokumentiert und herausgestellt werden. So wurde eine erste Liste von empirisch validierten Treatments aufgestellt (EVTs; Task Force, 1995; deutsch: Hahlweg, 1995), die später ergänzt und in Empirically Supported Psychological Therapies umgetauft wurden (ESPTs; Kendall, 1998, im zugehörigen Themenheft der Ausgabe 66 (1) der Zeitschrift *Journal of Consulting and Clinical Psychology*).

Gegner dieser Strömung befürchten einen ungerechtfertigten Ausschluß psychotherapeutischer Verfahren aus Praxis und Forschung und durch die Manualisierung eine Entwicklung in Richtung „Kochbuch-“ oder „Malen-nach-Zahlen-Psychotherapie“ (Silverman, 1996).

Die Zeitschrift *Psychotherapy Research* widmet sich in der Ausgabe 8 (2) (1998) ebenfalls eingehend diesem Thema. Der Herausgeber Robert Elliot stellt dabei zu Beginn einen „reader’s guide“ zur Kontroverse um die EVT’s bereit.

Die Behandlungsfähigkeit von Angststörungen im allgemeinen und von Panikstörung mit Agoraphobie im besonderen ist also seit der Entwicklung moderner kognitiv-behavioraler störungsspezifischer Therapien als ausgesprochen gut zu bewerten.

Allerdings mangelt es an Studien, die diese störungsspezifischen klinisch-psychologischen Behandlungsansätze auch unter routinemäßigen ambulanten Praxisbedingungen überprüft haben, so dass es auch ein Anliegen der vorliegenden Arbeit ist, die Befunde der kontrollierten Studien in diesem Setting zu replizieren.

1.2.5 Klinisch-Psychologische Behandlungsverfahren: Begriffsbestimmung

Wenn in dieser Arbeit im empirischen Teil von „Psychotherapie“, „psychologischer Angstbehandlung“ oder „kognitiver Verhaltenstherapie“ die Rede ist, so sind damit moderne störungsspezifische klinisch-psychologische Behandlungsverfahren gemeint, die auf Grundlage der empirischen klinischen Psychologie entwickelt wurden. Dieser Psychotherapiebegriff ist also zunächst von den herkömmlichen „Therapieschulen“ mit ihren unzähligen Unterformen (s. Tabelle 1.4) abzugrenzen.

Tabelle 1.4: Übersicht über die wichtigsten Therapieschulen (modif. nach nach Linster & Wetzel, 1980)

Tiefenpsychologische Therapien		Verhaltensorientierte Therapien		Erlebnisorientierte Therapien		Systemische Therapien		Körperorientierte Therapien	
Stichwort	Vertreter	Stichwort	Vertreter	Stichwort	Vertreter	Stichwort	Vertreter	Stichwort	Vertreter
Psychoanalyse	Freud	Verhaltensanalyse	Skinner	Existenzanalyse	Binswanger	Kommunikationsth.	Watzlawick/Jackson	Charakteranalyse	Reich
Analytische Therapie	Jung	Verhaltenstherapie	Eysenck	Daseinsanalyse	Boss	Strategische Th.	Haley	Bioenergetik	Lowen
Individualtherapie	Adler	Systemat. Desensibilisierung	Wolpe	Logotherapie	Frankl	Conjoint Family Therapy	Satir	Biodynamik Strukturanalyse	Boyesen Rolf
Interpersonale Psychiatrie	Sullivan	Sozialkognitive Lerntheorie	Bandura, Kanfer	Klientenzentrierte Gesprächstherapie	Rogers	System. Familientherapie (Ital. Schule)	Selvini/Palozzi/Andoli	Autogenes Training	I.H. Schultz
Ich-Analyse	A.Freud	Fixed-Role Therapy	Kelly	psychoth. Gestaltth.	Tausch Perls	Strukturelle Familientherapie	Minuchin	Progressive Entspannung	Jakobson
Neoanalyse	Schulz-Hencke	Rational-emotive Therapie	Ellis	Experimental Th./Focusing	Gendlin	Familienth (Heidelb. Schule)	Stierlin	Eutonie	G. Alexander
Fokaltherapie	Malan, Balint	Struktur. Lerntherapie	Goldstein	Primärtherapie	Janov	Mehrgenerationen-Familienth.	Boszormenyi-Nagi	Alexander-Technik	M. Alexander
Psychodrama	Moreno	Multimodale VT	A. Lazarus	Feelingtherapie	Hart	Beziehungsanalyse	Bauriedl	Bewusstheit d. Bewegung	Feldenkrais
Katathymes Bilderleben	Leuner	Kognitive VT	Mahoney, Beck, Meichenbaum	Initialische Therapie	Dürckheim	Kollusion/Koevolution	Willi	Atemtherapie	Middendorf
Transaktionsanalyse	Berne			Integrative Therapie	Petzold			Rebirthing	Orr
Dynamische Psychiatrie	Ammon			Neuroling. Programmieren (NLP)	Bandler Grin-der				

Dennoch kann eine Zuordnung getroffen werden: die untersuchte Therapie ist am ehesten im Bereich der Verhaltenstherapie zu verorten. Es stellt sich nun die Frage, was Verhaltenstherapie eigentlich ist. Grawe et al. (1994) weisen darauf hin, dass es sich bei der außerordentlich dynamischen Entwicklung von Therapien aus dem kognitiv-behavioralen Spektrum in den letzten drei Jahrzehnten nicht um die Entwicklung innerhalb einer Therapieschule handelt, sondern um den Beginn einer bedeutsamen Veränderung im Bereich der Psychotherapie, die deren Erscheinungsbild grundlegend gewandelt hat und weiter wandeln wird. Viele dieser Verfahren wurden erst seit den 70er Jahren entwickelt und haben fast nichts mehr mit den Konzepten zu tun, die Wolpe (1958) und Eysenck (1959) einst dem zugrunde legten, was sie „Verhaltenstherapie“ nannten. Es hat also etwas begonnen, was es vorher in der Psychotherapie nicht gab: eine intensive Forschung zu spezifischen Störungsbildern, die das Verständnis dieser Störungen völlig verändert und damit ganz neue Veränderungsmöglichkeiten aufgezeigt hat. Diese Entwicklung ist eng gekoppelt mit der sonstigen Entwicklung der empirischen (klinischen) Psychologie, die in dieser Zeit einen Grundsätzlichen Wandel in ihrer Sicht des Menschen vollzogen hat (Grawe et al., 1994, S. 744 ff.)

Viele Autoren haben versucht, die Verhaltenstherapie zu definieren, und die einflussreiche Definition von Yates (1970, zit. nach Margraf, 1996) soll hier stellvertretend für viele andere mögliche Definitionsversuche zitiert werden:

„Verhaltenstherapie ist der Versuch, den gesamten empirischen und theoretischen Wissensbestand, wie er durch den Einsatz experimenteller Methoden in der Psychologie und ihren Nachbardisziplinen (Physiologie und Neurophysiologie) angesammelt werden konnte, in systematischer Weise zu benutzen, um Entstehung und Beibehaltung abweichender Verhaltensmuster zu erklären, und weiterhin der Versuch, dieses Wissen bei der Behandlung oder Prävention solcher Fehlverhaltensweisen einzusetzen, und zwar mit Hilfe kontrollierter experimenteller Untersuchungen am einzelnen Patienten.“

Allerdings reicht eine reine Definition nicht aus, um eine aussagekräftige und zugleich zukunfts offene Standortbestimmung der modernen Verhaltenstherapie zu geben. Es müssen darüber hinaus die Grundprinzipien des verhaltenstherapeutischen Vorgehens und die zugrundeliegende Methodologie erläutert werden (Margraf & Lieb, 1995). Verhaltenstherapie wird als ein der klinischen Psychologie entsprungener Heilkundeansatz aufgefasst, der eine Vielzahl unterschiedlicher spezifischer Techniken und Behandlungsmaßnahmen beinhaltet. In Abhängigkeit der jeweiligen Problematik der Patienten werden diese Techniken unterschiedlich kombiniert und eingesetzt.

Margraf (1996) definiert die moderne Verhaltenstherapie schließlich als

”... eine auf der empirischen Psychologie basierenden psychotherapeutischen Grundorientierung, die störungsspezifische und -unspezifische Therapieverfahren umfaßt, die aufgrund von möglichst hinreichend überprüftem Störungswissen und psychologischem Änderungswissen eine systematische Besserung der zu behandelnden Problematik anstreben. Die Maßnahmen verfolgen konkrete und operationalisierte Ziele auf den verschiedenen Ebenen des Verhaltens und Erlebens, leiten sich aus einer Störungsdiagnostik und individuellen Problemanalyse ab und setzen an prädisponierenden, auslösenden und/oder aufrechterhaltenden Problemänderungen an. Die in ständiger Entwicklung befindliche Verhaltenstherapie hat den Anspruch, ihre Effektivität empirisch abzusichern” (Margraf, 1996, S. 3).

Innerhalb der therapeutischen Maßnahmen, die diese klinisch-psychologische Behandlungsweise kennzeichnen, können drei Gruppen unterschieden werden:

- Basisfertigkeiten wie Gesprächsführung, Beziehungsgestaltung und Motivationsarbeit,
- störungsübergreifende verhaltenstherapeutische Maßnahmen, die jeder Verhaltenstherapeut flexibel in den jeweiligen Behandlungsplan einfügen können muss: z.B. Konfrontationsverfahren in sensu und in vivo (Reizüberflutung, Habituationstraining, Reaktionsverhinderung, systematische Desensibilisierung etc.), Entspannungsverfahren (z.B. progressive Muskelrelaxation), kognitive Methoden (Selbstinstruktionstraining, Problemlösetraining, Modifikation dysfunktionaler Kognitionen, Reattribution, Analyse fehlerhafter Logik, Entkatastrophisieren etc.), Training sozialer Kompetenz und andere, sowie
- störungsspezifische Therapieprogramme, die möglichst genau auf die speziellen Gegebenheiten der verschiedenen Störungsbilder zugeschnitten sind.

In diesem Sinne ist die hier untersuchte klinisch-psychologische Behandlung als Verhaltenstherapie zu bezeichnen (mit Zuschnitt auf störungsspezifische Angstbehandlung), und als Verhaltenstherapie wurde sie auch abgerechnet.

1.3 Kostenanalysen: Begriffsklärung

Im Gesundheitswesen kommt Kostenanalysen eine besondere Bedeutung zu, da hier weitgehend nicht durch marktbezogene Prozesse, sondern durch staatliche Eingriffe reguliert wird. So soll Entscheidungsträgern eine Orientierung über eine ökonomisch sinnvolle Verwendung der verfügbaren Mittel (Ressourcenallokation) ermöglicht werden, zumal derzeit im gesamten Gesundheitswesen enorm gespart werden muss.

Noch bis in die siebziger Jahre hinein wurden die ansteigenden Ausgaben im Gesundheitswesen als positives Zeichen für den steigenden medizinischen Fortschritt bewertet. In den letzten Jahrzehnten entwickelten sie sich jedoch zunehmend zum gesellschaftlichen Problem, wie die Diskussion um die sogenannte Kostenexplosion im Gesundheitswesen widerspiegelt. Angesichts der hohen Ausgaben im Gesundheitswesen (1996 lag der Anteil der Gesamtkosten für das Gesundheitswesen am Bruttoinlandsprodukt bei über 14%) wird die Identifikation wirksamer, effizienter und kostengünstiger medizinischer Leistungen unabdingbar.

Hier bietet sich als Entscheidungshilfe eine Kosten-Nutzen-Analyse an. Die Begrenzung von unwirksamen und/oder unwirtschaftlichen Leistungen dürfte sich vielfach schwierig gestalten; sie stellt aber angesichts der Knappheit der Ressourcen eine unverzichtbare Aufgabe dar (Sondergutachten des Sachverständigenrates für die konzertierte Aktion im Gesundheitswesen, 1995). Obwohl die Forderung der ökonomischen Evaluation öffentlicher Maßnahmen im Haushaltsrecht von Bund und Ländern verankert ist, werden Kosten-Nutzen-Analysen bisher nur selten durchgeführt.

Im folgenden werden die für Kosten-Studien relevanten Begriffe überblicksartig erläutert. Das aus der Literatur abgeleitete exakte Vorgehen in dieser Studie wird in den späteren Abschnitten detaillierter ausgeführt.

1.3.1 Definition von „Kosten“ und „Nutzen“

Kosten-Nutzen-Analysen basieren auf dem sogenannten *Opportunitätskostenprinzip*, wobei die Opportunitätskosten als entgangener Nutzen für eine alternative Verwendung der Ressourcen definiert sind (Wittmann, 1985). Der Wert dieser Opportunitäten wird in Geldeinheiten erfaßt. Im Bereich der Psychotherapie kann dies z.B. bedeuten, dass die Kosten der untersuchten Behandlung zunächst einmal ausgegeben werden müssen, d.h. an anderer Stelle im Gesundheitssystem (oder im Budget eines selbst zahlenden Patienten) fehlen. Übersteigt aber der Nutzen der Therapie (z.B. Einsparungen weiterer Gesundheitsleistungen, Verminderung von Arbeitsunfähigkeit) den Nutzen einer al-

alternativen Verwendung des Geldes (z.B. für eine andere Therapie oder bei Nichtbehandlung für etwas, das mit der Störung gar nichts zu tun hat), dann kann man von einer positiven Kosten-Nutzen-Relation sprechen.

Es gibt verschiedene Arten von Kosten und Nutzen, die in einer Analyse betrachtet werden können. In Tabelle 1.5 werden die verschiedenen Arten von Kosten und Nutzen an Beispielen erläutert (vgl. für den medizinischen Bereich Galler, 1990). Man unterscheidet zwischen *tangiblen* und *intangiblen* und *direkten* und *indirekten* Kosten und Nutzen. Unter *tangiblen* Nutzen versteht man den Nutzen, dem man direkt einen Marktpreis zuordnen kann (z.B. höherer Gewinn eines Unternehmens bei besserer Motivation der Arbeitnehmer nach einer arbeitspsychologischen Intervention). Für den *intangiblen* Nutzen hingegen gibt es keine klar bezifferbaren Preise. Das trifft beispielweise auf bessere Lebensqualität zu. Direkte Kosten sind diejenigen Werte, die man bewusst in das Projekt investiert, d.h. die nicht unbeabsichtigt an anderer Stelle entstehen. Direkter Nutzen ist der durch das Programm erstrebte Nutzen, der indirekte entsteht an anderer Stelle ungeplant.

Tabelle 1.5: Übersicht über die verschiedenen Arten von Kosten und Nutzen von Psychotherapie

	Kosten	Nutzen
direkt tangibel	z.B. in Rechnung gestellte Kosten der Behandlung pro Patient	Reduktion der Inanspruchnahme anderer medizinischer Leistungen
direkt intangibel	Unannehmlichkeiten durch die Behandlung für den Patienten (z.B. Zeitaufwand für Therapie)	Bessere Lebensqualität nach Ende der Behandlung
indirekt tangibel	Nebenkosten der Therapie, die der Patient selbst übernehmen muss, z.B. Fahrtkosten zum Behandlungsort; ggf. Übernachtungskosten; Kosten, die Rahmen von Konfrontationsübungen entstehen (z.B. Zugfahrten)	Reduktion der Arbeitsunfähigkeit, weitere Aspekte wie z.B. höheres zukünftiges Gehalt für einen Patienten, der Dank der Therapie seine Ausbildung abschließen bzw. seine Karriere weiterverfolgen kann
indirekt intangibel	Unannehmlichkeiten für Verwandte durch die Therapie (z.B. Mehraufwand bei der Haushaltsführung, wenn ein Familienmitglied sich für Wochen in einer stationären Maßnahme befindet)	Bessere Lebensqualität der Verwandten in Folge der Therapie (z.B. wenn geheilter Patient wieder mehr familiäre Aufgaben übernehmen kann)

Kosten- und Nutzen-Perspektiven

Die Schwerpunkte unterschiedlicher Analysen werden zudem hinsichtlich der *Perspektive* unterschiedlich gesetzt werden. Vor allem der Nutzen kann aus unterschiedlichen Perspektiven betrachtet werden. Die folgende Darstellung lehnt sich an die Beschreibung in Warner und Luce (1982):

- *Persönlicher Nutzen*: bessere Lebensqualität des einzelnen behandelten Patienten, bessere Lebensqualität seiner Verwandten,
- *Nutzen für das Gesundheitswesen*: Einsparungen durch geringere Behandlungskosten an anderer Stelle infolge der Intervention,
- *Nutzen für die Volkswirtschaft*: Reduktion des Produktivitätsausfalls infolge von Arbeitsunfähigkeit, sowie
- *Sozialer Nutzen*: beispielsweise die allgemein bessere Zugänglichkeit psychologischer Leistungen für Menschen niedrigerer Schichten und damit eine Besserung der gesamtgesellschaftlichen Gerechtigkeit.

Man sollte in einer Kostenstudie immer definieren, welche Kosten und Nutzen als direkt oder indirekt bezeichnet werden, da dies durchaus unklar sein kann. So wird z.B. die Wiedererlangung von Arbeitsfähigkeit im medizinischen Bereich in der Regel als indirekter Nutzen gewertet. Im Zusammenhang mit der hier untersuchten psychologischen Angstbehandlung könnte sie durchaus auch unter den direkten Nutzen gefasst werden, da die Verbesserung der Arbeitsfähigkeit bei Problemen in diesem Bereich in der Regel ein explizites psychotherapeutisches Therapieziel darstellt und nicht unbeabsichtigt „nebenbei“ anfällt.

Die Ergebnisse einer Kostenstudie sollten bei der Darstellung für die verschiedenen Perspektiven getrennt ausgewiesen werden.

1.3.2 Kosten-Berechnung: Betriebskosten- vs. umfassender Ansatz

Es gibt mehrere Möglichkeiten, die Kosten einer Behandlung zu berechnen. Meistens wird in Kosten-Studien der *Betriebskostenansatz* gewählt (direkter tangibler Nutzen, vgl. Bühringer & Hahlweg, 1986). In die Betriebskosten gehen Personalkosten, Einrichtungskosten und Materialkosten für den Behandler bzw. für die behandelnde Institution ein. Die Bezifferung der Betriebskosten erfolgt über die den Patienten oder den Kostenträgern in Rechnung gestellten Therapiestunden, stationäre Pflage tage etc., da davon ausgegangen wird, dass darin der Wert der für die direkt in der Therapie aufgewen-

deten Ressourcen enthalten ist. Dieser Weg wird auch in der vorliegenden Studie beschränkt.

Neben den Betriebskosten gibt es aber noch weitere Berechnungsgrundlagen, die zur Quantifizierung der in der Therapie aufgewendeten Ressourcen dienen können (vgl. Yates & Newman, 1980). Dabei können auch die oben erwähnten indirekten und intangiblen Kosten mit einbezogen werden. Aber auch nicht in Rechnung gestellte Leistungen, wie ehrenamtliche Arbeit oder unterbezahlte therapeutische Arbeit einzelner Mitarbeiter einer Therapieeinrichtung (z.B. im Rahmen von Ausbildungstherapien oder beim obligatorischen „Psychiatrie-Jahr“) sind für eine Gesellschaft nicht kostenlos. Solange diese Zuwendungen theoretisch auch anderen Zwecken hätten dienen können, zählen sie zu den Opportunitätskosten der Behandlung.

Weiterhin können in einen Ansatz, der umfassend alle Gesamtkosten berücksichtigt, auch noch Folgekosten eingehen, die durch die jetzige Therapie verursacht werden (wobei dieser Aspekt bei der Psychotherapie wohl in geringerem Ausmaß eine Rolle spielt als in der Medizin, wo z.B. in bestimmten lebensrettenden Bereichen deutliche schädigende Nebenwirkungen einer Therapie in Kauf genommen werden müssen).

Ein Beispiel von Siegert & Yates (1980) zeigt, dass es einen Unterschied machen kann, ob man den einfachen Betriebskostenansatz (*operations perspective*) oder den skizzierten umfassenden Ansatz (*comprehensive approach* unter Einbezug indirekter und intangibler Kosten) zur Berechnung der Kosten einer Maßnahme heranzieht. Wurden lediglich die Betriebskosten berücksichtigt, lagen in dieser Studie die Kosten dreier Trainingsalternativen für Eltern verhaltensgestörter Kinder eng beieinander. Bezog man aber im umfassenden Ansatz auch indirekte Kosten mit ein (z.B. Zeitaufwand für die Eltern), so erwies sich eine der drei Alternativen als deutlich teurer als die anderen beiden.

Warum aber sollte man dann einen Betriebskostenansatz der umfassenden Methode zur Berechnung der Behandlungskosten vorziehen? Dies hat pragmatische bzw. Machbarkeitsgründe. Es wäre im Rahmen unserer (multizentrischen) Studie nicht möglich gewesen, rückwirkend bei allen Patienten, deren Therapien z.T. Jahre zurücklagen, die zuvor nicht erfaßten indirekten Therapiekosten nachzuerheben. Und im Rahmen der später ausgeführten Kosten-Nutzen-Analyse hätten gar die gesamten Kosten vor und nach der Therapie, die durch weitere Behandlungen aufgrund der Störung angefallen sind, dann ebenfalls umfassend berechnet werden müssen. Diese immensen Untersuchungskosten hätten den zu erwartenden Nutzen solcher Analysen nicht gerechtfertigt. Dass eine Behandlung einer Störung eben „das kostet, was sie kostet“, ist zwar eine Vereinfachung, aber durchaus eine Basis, auf der sich Kosten-Effektivitäts- und Kosten-Nutzen-Analysen anstellen lassen. Es wird an entsprechenden Stellen allerdings diskutiert, inwieweit von einer umfassenden Berechnungsgrundlage für die Therapiekosten wichtige andere oder zusätzliche Ergebnisse zu erwarten wären.

1.3.3 Nutzen-Berechnung

Der Nutzen der untersuchten Behandlung wird in dieser Arbeit über Einsparungen bei Gesundheitsleistungen und Arbeitsunfähigkeit beziffert (direkter tangibler Nutzen). Dabei werden die störungsbezogenen Kosten vor und nach der Psychotherapie ebenfalls nach dem Betriebskostenansatz bestimmt.

Diese Perspektive wird in einem eigenen Abschnitt durch die subjektive, d.h. direkte und indirekte intangible Perspektive ergänzt. Es gibt Methoden, auch den subjektiven Nutzen monetär zu bewerten: in dieser Arbeit wird erstmals im Psychotherapiebereich das Maß der sogenannten *Zahlungsbereitschaft* verwendet.

1.3.4 Analyse-Typen

Kosten-Effektivitäts-Analyse

Die Kosten-Effektivitäts- oder Kosten-Wirksamkeits-Analyse (*cost effectiveness analysis*) geht der Frage nach, was die Erreichung eines bestimmten therapeutischen Ergebnisses kostet. Dabei werden die Kosten (in monetären Einheiten) der Effektivität (beliebige eindimensionale Einheit, z.B. reduziertes Gewicht in kg, klinisch relevante Verbesserung in einem störungsspezifischen Fragebogen) gegenübergestellt und ein Quotient gebildet.

Ein häufiges Problem ist aber, dass in den einzelnen Studien zur Kosten-Effektivität eines Programmes unterschiedliche Effektivitätseinheiten gewählt wurden und die Ergebnisse dadurch nicht vergleichbar sind. Meistens wird in der Forschung ein Quotient von Gesamtkosten in Relation zur Anzahl „erfolgreich behandelte“ Patienten verwendet (z.B.: kostet eine Tinnitus-therapie von 10 Patienten 10000,- DM und sind danach 5 Patienten klinisch bedeutsam gebessert, dann kostet ein erfolgreich therapierter Patient 2000,- DM). Ist ein Vergleich zwischen den Alternativen aufgrund gleicher Effektivitätsdimensionen möglich, so ist das Verfahren mit dem günstigsten Kosten-Effektivitäts-Verhältnis zu bevorzugen.

Weiterhin müssen bei einer Entscheidung für oder gegen eine Alternative Überlegungen zu Kostenober- und Effektivitätsuntergrenzen mit einbezogen werden, da u.U. ein sehr effektives Programm unbezahlbar oder ein sehr billiges Programm zu ineffektiv sein kann.

Abschnitt 2.2 informiert ausführlich über Theorie und Forschungsstand von Kosten-Effektivitäts-Analysen.

Kosten-Nutzwert-Analyse

Wie eben geschildert ist die Anwendung der Kosten-Effektivitäts-Analyse auf den Vergleich von Programmstudien beschränkt, welche die Effektivität in ein und derselben Einheit gemessen haben. Häufig werden aber sehr unterschiedliche Maße verwendet, so dass die Studien untereinander nicht vergleichbar sind. Dieser Nachteil führte in der Medizin zur Entwicklung der Kosten-Nutzwert-Analyse (*cost utility analysis*). Sie verwendet ein einheitliches, aber mehrdimensionales Maß als Indikator für den Erfolg einer Therapie. Dabei werden quantitative und qualitative Dimensionen der Gesundheit (d.h. die Wirkung der Therapie auf die Länge und die Qualität des Lebens) berücksichtigt. Solche Maße sind z.B. Quality Adjusted Life Years (QALYs) oder Health Years Equivalents (HYEs). Als Ergebnis der Analyse erhält man das Kosten-Nutzwert-Verhältnis (z.B. Kosten pro QALY).

Vorwiegend wird der Nutzwert über die Methoden „Time trade off“ oder „Standard Gamble“ bestimmt. Beim Time Trade Off wird den Probanden ein schlechter Gesundheitszustand beschrieben. Sie sollen sich vorstellen, sich kostenlos behandeln lassen zu können, woraufhin ihre Gesundheit um ein definiertes Maß steigen würde, allerdings nur in Verbindung mit einer Verkürzung der Lebensdauer. Diese zu akzeptierende Lebensdauerverkürzung wird so lange gesteigert, bis der Befragte die Alternative Behandlung der Alternative Nichtbehandlung nicht mehr vorzieht. Beispielsweise wird ein an rheumatischer Erkrankung leidender 20-jähriger Patient zu einer Behandlung befragt, die eine völlige Schmerzfreiheit von 10 Stunden am Tag verspricht und leichte Schmerzen in Form eines Ziehens in Armen und Beinen in der restlichen Zeit des Tages. Für diese Behandlung würde er eine Lebensdauerverkürzung von 10 Jahren, 20 Jahren, 25 Jahren, aber nicht mehr von 26 Jahren in Kauf nehmen. Der dem geschilderten Gesundheitszustand entsprechende Nutzwert würde also 25 Jahre betragen. Selbstverständlich muß hier das Lebensalter des Patienten als Korrekturgröße berücksichtigt werden.

Bei der Methode Standard Gamble erfolgt dieselbe Prozedur, nur geht es hier nicht um eine Verkürzung der Lebensdauer, sondern um die Wahrscheinlichkeit, bei der Intervention zu sterben. Hier bildet die höchste akzeptierte Wahrscheinlichkeit zu sterben das Maß für den Nutzen der gewonnenen Lebensqualität.

Weil die Kosten-Nutzwert-Analyse und die Kosten-Effektivitäts-Analyse sehr eng verwandt sind, unterscheiden viele Autoren nicht zwischen ihnen. Einzelne Studien zu diesem Ansatz zu finden, ist deshalb sehr schwierig; zudem ist diese Analysemethode bislang auf den medizinischen Bereich beschränkt. In der klinisch-psychologischen Forschung gibt es eigene etablierte Maße zur Bestimmung der Effektivität einer Maßnahme, die geeigneter erscheinen als etwa QALYs oder HYEes. Deshalb sei diese Methode in der vorliegenden Arbeit ausgeklammert.

Kosten-Nutzen-Analyse

Bei der Kosten-Nutzen-Analyse (engl. *cost benefit analysis*) werden die Kosten der Therapie in Relation zum Nutzen gesetzt. Im Gegensatz zur Kosten-Effektivitäts- und zur Kosten-Nutzwert-Analyse werden aber für beide Dimensionen die gleichen Einheiten verwendet, meistens monetäre Skalen.

Damit ist auch eine absolute Beurteilung der Therapie möglich. Übersteigt nämlich der erhaltene Quotient den Wert 1, so handelt es sich um eine wirtschaftlich nützliche Maßnahme per se, da die Ausgaben dafür aufgewogen werden.

Darüber hinaus ist natürlich noch eine vergleichende Betrachtung mit anderen Alternativen möglich. Die alleinige Darstellung des Quotienten ist allerdings von Nachteil, da dieser dann keine Aussage mehr über die absolute Höhe der zugrunde liegenden Kosten und Nutzen erlaubt. So ist z.B. der *Nettonutzen* (Nutzen nach Abzug der Kosten) bei einer Maßnahme, die bei Kosten von DM 2000.- eine Einsparung von DM 4000.- erbringt (Nettonutzen = DM 2000.-) geringer als der einer Maßnahme, die bei Kosten von DM 4000.- DM 8000.- einspart (Nettonutzen = DM 4000.-), obwohl beide Maßnahmen das selbe Kosten-Nutzen-Verhältnis von 1:2 aufweisen.

Besondere Erwähnung verdient der Einbezug des intangiblen Nutzens aus Patientensicht in eine Kosten-Nutzen-Analyse, z.B. mit der bereits erwähnten Methode der Zahlungsbereitschaft. Auf diese Weise ermittelte Nutzens-Werte sollten aber in Kosten-Nutzen-Analysen getrennt vom tangiblen Nutzen (für den geeignete Preise leichter zu ermitteln sind) dargestellt werden. Abschnitt 2.3 informiert ausführlich über Theorie und Forschungsstand von Kosten-Nutzen-Analysen.

Vergleich von Kosten-Effektivität und Kosten-Nutzen

Die Kosten-Effektivität besitzt als Erfolgsmaß besondere Vorteile gegenüber Kosten-Nutzen-Quotienten:

- *Plausibilität und Flexibilität.* Den Outcome bzw. den Nutzen einer Therapie in monetären Einheiten auszudrücken, wirft bei Kosten-Nutzen-Quotienten grundsätzliche Fragen auf: Kann man die Steigerung an subjektivem Wohlbefinden direkt in Geld umrechnen? Ist es nicht eine unzulässige Reduktion, „Therapieerfolg“ als Einsparung für das Gesundheitssystem oder als Wiedererlangung volkswirtschaftlich produktiver Arbeitskraft zu definieren? Die Kosten-Effektivität hingegen ist unmittelbar plausibel: Je nach Perspektive kann auf unterschiedliche Outcome-Maße zurückgegriffen werden. Es herrscht zwar auch hier bei weitem noch keine Einigkeit, wie eine „erfolgreiche Behandlung“ zu definieren ist, aber die inhaltlich bisweilen störende Übersetzung von Therapieerfolg in monetären Nutzen entfällt. Bei einer Kosten-Effektivitäts-Analyse ist man in der Wahl sinnvoller Effektivitäts-

kriterien grundsätzlich frei – wichtig ist allerdings die Transparenz des gewählten Vorgehens und, beim Vergleich verschiedener Behandlungen, die Vergleichbarkeit der gewählten Ansätze und Methoden zur Kosten- und Effektivitätsbestimmung.

- *Praktikabilität.* Die Kosten einer therapeutischen Maßnahme sind bereits nicht immer einfach zu bestimmen. Bei weitem aufwendiger ist es im Rahmen einer Kosten-Nutzen-Analyse, den Nutzen zu berechnen. Dies erfordert eine Vielzahl an Erhebungen, die mit der eigentlich zu untersuchenden Behandlungsform nichts zu tun haben (z.B. detaillierte Erhebung der durch eine Störung verursachten Kostenfaktoren für angemessene Zeiträume vor und nach der Behandlung). Die Effektivitätsmessung hat dagegen einen unmittelbaren Bezug zu diagnostischen Routineerhebungen und normalen bewährten Katamnesen, wie sie im Rahmen von Qualitätssicherung und in der Forschung eingesetzt werden.

In der Öffentlichkeit haben Kosten-Nutzen-Quotienten allerdings auch einen entscheidenden Vorteil gegenüber der Kosten-Effektivität, denn sie stehen für sich:

„Jede Mark, die in die Psychotherapie gesteckt wird, kann später drei Mark sparen!“

Solche Werte müssen nicht hinsichtlich eines bestimmten Effektivitätskriteriums interpretiert werden, wie z.B. die kompliziertere Aussage:

„Psychotherapie führt, bei Kosten um DM 5000.-, zu klinisch signifikanten Befindlichkeitsverbesserungen bei 80% der behandelten Patienten“, oder
„Behandlungsmethode A hat mit DM 12000.- eine günstigere Kosten-Effektivität als Behandlungsmethode B mit DM 16500.-“.

Inhaltlich sind die letzten beiden Aussagen aussagekräftiger als die erste, zumal es angesichts der Strukturen in unserem Gesundheitssystem sowieso fraglich ist, ob eine Reduktion in Anspruch genommener Behandlungen nach einer Psychotherapie wirklich irgendwo Geld einspart. Der Preis für eine höhere Aussagekraft und größere Differenzierung psychologisch sinnvoller Therapienutzenbestimmung liegt in der niedrigeren Prägnanz – vielleicht führt die Kosten-Effektivität deshalb noch ein eher stiefmütterliches Dasein.

Gemeinsam ist beiden sogenannten zweidimensionalen Ansätzen, dass sie mehr relevante Informationen berücksichtigen als eindimensionale Ansätze, die die Seite der für die Behandlung aufgewendeten Ressourcen ausblenden. Und es handelt sich um umfassendere Beurteilungen, bei der stets auch mehrere Ergebnisaspekte integriert und in klinisch und praktisch bedeutsamer Weise aggregiert werden (vgl. auch 1.5: Fallbeispiele).

1.4 Ethik von Kostenanalysen im Gesundheitswesen

Kostenanalysen im Gesundheitswesen sind nicht unumstritten. Gesundheit und Wirtschaftlichkeit stellen konkurrierende Werte dar, und die Gegner derartiger Untersuchungen argumentieren, dass jeder Mensch ein Recht auf Gesundheit besitze, das ihm nicht aus ökonomischen Gründen verweigert werden könne. Dieser Satz wird so sicher überall auf Zustimmung stoßen. Was geschieht jedoch, wenn eine effektive Behandlung (z.B. gegen AIDS) entwickelt würde, die pro Patient 10 Mio. DM kostet? Wo ist die Grenze des Bezahlbaren? Diese Frage muß letztendlich gesellschaftlich bzw. politisch ausgehandelt werden.

Der monetäre gesellschaftliche Nutzen allein darf niemals das einzige Entscheidungskriterium über die Durchführung einer Behandlung sein. Wichtig ist immer auch die Betrachtung des individuellen, subjektiven Nutzens. Kosten-Nutzen-Überlegungen dürfen nur im Sinne einer summativen Evaluation herangezogen werden (Neumer & Margraf, 1996). Es gilt daher: Keine Kostenanalyse ohne Berücksichtigung der Effektivität bzw. des intangiblen Nutzens der Therapie!

Ein weiterer Punkt ist ebenfalls zu bedenken. Berechnet man die durchschnittliche Einsparung einer Therapie, so kann es problematisch sein, diese Ergebnisse auf den einzelnen Patienten zu übertragen. Der positive monetäre Nutzen bei einem Spitzenverdiener ist ein viel größerer als derjenige bei einem Erwerbslosen. Solche Überlegungen dürfen aber niemals dazu führen, dass die Therapie eher Personen zugute kommt, bei denen es sich mehr „lohnt“. Kosten-Nutzen-Analysen können also nur eine Grundlage für Entscheidungen über Ressourcenallokationen darstellen und sind nicht die Entscheidung selbst. Dieser Punkt wird auch im folgenden Abschnitt (Fallbeispiele aus der Kosten-Nutzen-Analyse) angesprochen.

Ebenso verhält es sich mit der Therapie verschiedener Störungsbilder: Die ambulante Therapie von Angststörungen weist, wie im Hauptteil ausgeführt, ein gutes Kosten-Nutzen-Verhältnis auf. Dieses darf nicht ohne weiteres zum Vergleich mit anderen Störungsbildern herangezogen werden, denn es wäre ethisch unhaltbar, aufgrund eines solchen Vergleiches die Behandlung etwa von sexuellen Funktionsstörungen (geringerer direkter tangibler Nutzen) oder von Borderline-Störungen (meist deutlich höhere Behandlungskosten) zu verweigern.

Der Grundsatz der Gleichbehandlung muß also gewahrt werden, und ökonomische Evaluatoren bzw. Entscheidungsträger müssen sich derartige ethische Prinzipien stets besonders vor Augen halten.

1.5 Fallbeispiele zur Illustration unterschiedlicher Kosten-Nutzen-Indices bei gleichem Outcome

Die folgenden Beispiele dreier erfolgreicher Therapien aus der eigenen Praxis des Autors sollen veranschaulichen, dass Therapieerfolg und die Güte des Kosten-Nutzen-Verhältnisses stark voneinander abweichen können. Dies liegt vor allem an der Schwierigkeit, bestimmte Kostenfaktoren monetär zu bewerten.

1.5.1 Exzellentes Kosten-Nutzen-Verhältnis

Die 29jährige Frau M. suchte die Behandlung nach einer außergewöhnlich umfangreichen Krankheits- und Behandlungskarriere auf. Primärdiagnose war eine zu diesem Zeitpunkt schwere Zwangsstörung mit umfangreichen Wasch- und Reinigungsritualen, die einen Großteil des Tages einnahmen. Die Patientin war ausgesprochen komorbid und erfüllte bei Therapiebeginn u.a. die Kriterien für Panikstörung ohne Agoraphobie, depressive Episode, Somatisierungsstörung und einer weiteren spezifischen Phobie.

Die letzten fünf Jahre hatten sie ihre psychischen Probleme völlig in Anspruch genommen, was sich unter anderem im Abbruch des Studiums und völliger Unselbständigkeit niedergeschlagen hatte. Sie hatte unter anderem insgesamt 20 Wochen Psychiatrie bzw. Psychosomatik, etwa 350 Psychoanalysesitzungen, unzählige Medikamente und unzählige Haus- und Facharztbesuche hinter sich. Die dem Forschungsstand entsprechende Methode der Wahl (Reizkonfrontation mit Reaktionsverhinderung und kognitive Therapie) waren ihr bei all diesen Behandlungen nicht – bzw. nur halbherzig¹ – angeboten worden.

Nun hatte sie sich über die störungsspezifische Therapie informiert und hoffte, in intensiver ambulanter Betreuung mit dieser Methode endlich aus ihrem durch die Störung völlig reglementierten Leben auszubrechen. Ihre Krankenkasse (DAK), die in den vergangenen Jahren etwa DM 110000.- für die Behandlungen aufwenden mußte, verweigerte bezeichnenderweise die Kostenerstattung für eine solche Psychotherapie und empfahl ihr einen weiteren stationären Aufenthalt. So ließ sich die Patientin DM 5000.- von ihrer Schwester und bezahlte die Therapie privat.

Die Therapie verlief sehr erfolgreich. Bereits zwei Monate nach Erstkontakt konnte die Patientin wieder selbständig in eigener Wohnung leben und sich einen Job suchen. Heute, drei Jahre später, ist die Patientin gut integriert, weist kaum noch Symptome auf und geht einer geregelten Arbeit nach. Nach Therapieende hatte sie zwar noch viel zu kämpfen, konnte Behandlungen aber auf etwa 20 unregelmäßige telefonische Kontakte zu ihrem Therapeuten begrenzen. Für sie würde das individuelle Kosten-Nutzen-Verhältnis allein für die Behandlungen über 1:10 betragen. Dazu käme ein schwer zu beziffernder Nutzen für den Übergang von der untätigen Studentin auf Sozialhilfeniveau zur dauerhaf-

¹ Dies betrifft übrigens ein in der Praxis regelhaft anzutreffendes Problem: fast jeder Psychotherapeut kennt theoretisch die Vorzüge von Konfrontationsverfahren (die Methode wird auch in jedem Klinik-Konzept herausgestellt) – aber die Durchführung *lege artis* ist eklatant seltener anzutreffen (vgl. Wittchen, in Druck)

ten Berufstätigkeit. Auch Marginalkosten, die bei der Berechnung in dieser Studie nicht berücksichtigt werden konnten, fielen nach der Therapie weg (z.B. gab die Patientin etwa ein Drittel ihres verfügbaren Einkommens vor Therapie für Putz- und Hygieneartikel aus). Abgesehen vom hohen intangiblen Nutzen (der sich auch aus der hohen Zahlungsbereitschaft erschließen läßt) haben wir es also hier mit einem extrem guten direktem tangiblen Kosten-Nutzen-Verhältnis einer störungsspezifischen Psychotherapie zu tun.

1.5.2 Durchschnittliches Kosten-Nutzen-Verhältnis

Der 38jährige Herr A. litt an einer klassischen Panikstörung mit Agoraphobie, als er sich zur Behandlung vorstellte. Aufgrund seiner eingeschränkten Mobilität war er beruflich und privat stark belastet.

Seine Vorbehandlungsgeschichte, bevor er in unserer Behandlungseinrichtung eine störungsspezifische Therapie (insbesondere Konfrontationstherapie) erhielt, entsprach ziemlich genau dem Durchschnitt: in den zwei Jahren vor Therapie durchlief er eine Gruppentherapie in einer Tagesklinik, die er aber bald abbrach, weil sie nichts mit seinem Problem zu tun gehabt habe; außerdem suchte er aufgrund seiner Beschwerden vier Fachärzte je einmal auf und rief zwei mal (medizinisch unbegründet) den Notarzt. Arbeitsunfähig war er bislang nicht, wollte aber die Therapie u.a. deswegen absolvieren, weil er bei einer absehbaren Beförderung Aufgaben übernehmen müßte, die er mit der Störung nicht hätte durchführen können.

Auch hier verlief die Therapie erfolgreich, und auch zwei Jahre nach Therapieende waren keinerlei weitere Behandlungen oder Arbeitsunfähigkeitszeiten aufgrund psychischer Probleme zu verzeichnen. In diesem Fall ist das Kosten-Nutzen-Verhältnis durchschnittlich, d.h. dass die durch die Therapie verursachten Kosten durch Einsparungen etwa nach zwei Jahren amortisiert waren.

1.5.3 Schlechtes Kosten-Nutzen-Verhältnis

Eine 19jährige Abiturientin wollte studieren, fühlte sich aber aufgrund ihrer seit drei Jahren ausgeprägten phobischen Angst vor Dunkelheit und vorm Alleinsein nicht in der Lage, aus dem Elternhaus auszuziehen und sich im gewünschten Studienort ein Zimmer zu nehmen.

Abgesehen von fünf Besuchen bei ihrem früheren Kinderarzt wegen dieses Problems hatte sie vor der Therapie keine weitere Behandlung erhalten, und bisherige Beeinträchtigungen ihrer Arbeitsfähigkeit hätten sich noch nicht finanziell zu Buche schlagen können.

Die Therapie (kognitive Bearbeitung der Angst inhalte, wiederholte Konfrontation mit gefürchteten Reizen) verlief erfolgreich (30 Behandlungseinheiten, Kosten etwa DM 3500.-), so dass die Patientin ihre Ängste dauerhaft bewältigte und mittlerweile den ersten Studienabschnitt in eigener Wohnung erfolgreich absolviert hat. Weitere Behandlungen wurden nicht in Anspruch genommen. Hier handelt es sich unter Kosten-Nutzen-Aspekten, so wie sie in dieser Studie analysiert werden können, um ein ausgesprochen schlechtes Kosten-Nutzen-Verhältnis, da den Therapieausgaben praktisch keine Einsparungen gegenüberstehen. Doch wie teuer wäre die Störung geworden, wenn sie fortbestanden hätte und Studium oder Berufstätigkeit vielleicht auf Jahre verhindert hätte?

Viele Patienten berichten von Brüchen in ihrer Biografie durch Angststörungen, die im frühen Erwachsenenalter manifest wurden und zu ständigen Folgeproblemen führten. Die „Investition“ der DM 3500.- für die genannte junge Patientin scheint unter diesem Gesichtspunkt unbedingt gerechtfertigt, auch wenn der Nutzen erst Jahre später (hypothetisch) meßbar sein wird.

Wie im Abschnitt zur Ethik von Kostenanalysen angesprochen sollten alle drei „Prototypen“ von Betroffenen wegen starker Belastung und Beeinträchtigung, sowie aufgrund ihrer Behandlungsmotivation ein Anrecht auf Behandlung bekommen – auch wenn sich die Behandlung aus volkswirtschaftlicher Sicht nicht in allen Fällen „lohnt“.

2 Theorie und Forschungsstand

2.1 Outcome-Forschung im Bereich Psychotherapie

Zunächst soll die vorliegende Studie in die Psychotherapieforschung eingeordnet werden. Insbesondere soll dabei auf die Validitätsproblematik eingegangen werden: Ist das in dieser Studie genutzte Prä-Post-Design aussagekräftig?

Die in dieser Arbeit aufgeworfenen Fragen „Wie viele psychotherapeutische Ressourcen werden verbraucht, um bestimmte Effekte zu erzielen?“ oder „Wie ist das Verhältnis von Kosten und Nutzen psychotherapeutischer Behandlung?“ beziehen sich auf die Ergebnisse (Outcome) psychotherapeutischer Interventionen. Die Erforschung solcher Ergebnisse hat viele Facetten (Lambert & Hill, 1994): Wie soll Therapieerfolg definiert und gemessen werden? Unterscheiden sich die Erfolge einer bestimmten Interventionsform bei verschiedenen Gruppen (z.B. bei verschiedenen Störungsbildern)? Unter welchen Bedingungen wirkt Psychotherapie, und wie stark?

Und auf einer molaren Ebene:

2.1.1 Welche Therapie ist erwiesenermaßen wirksam?

An der Frage der nachgewiesenen Wirksamkeit und der jeweiligen Überlegenheit verschiedener Therapieverfahren scheiden sich die Geister. Eysenck (1952) provozierte zunächst mit der These, dass Psychotherapie überhaupt nicht helfe, bzw. die Erfolgsraten nicht besser seien als Spontanremissionen². Wie bereits unter 1.2.5 angeführt, ist diese Einschätzung aufgrund neuerer Erkenntnisse der Verlaufsforschung gerade für den Bereich der Angststörungen nicht mehr haltbar.

Weiterhin wurde die Auffassung vertreten, dass Psychotherapie zwar schon hilft (d.h. besser ist als Nichtbehandlung) – aber auf ganz unspezifische Weise: alle therapeutischen Verfahren seien gleich wirksam (vgl. Kiesler, 1966; Luborsky, Singer & Luborsky, 1975).

² In einem Kommentar in der Psychologischen Rundschau stellt Eysenck (1993) allerdings heraus, dass er 1952 einen anderen logischen Schluß gezogen habe: wegen der schlechten Forschungslage könne (noch!) nicht behauptet werden, dass (die damals gängigen) Psychotherapien positive Effekte hätten. Auch seine Behauptung hoher Spontanremissionsraten sei nach wie vor noch nicht überzeugend widerlegt worden (Rachman & Wilson, 1980).

In der Tat erweist sich aus verschiedenen Gründen ein direkter Vergleich beispielsweise zwischen Verhaltenstherapie und Psychoanalyse in ein und derselben Studie als äußerst schwer durchzuführen. Unter anderem deswegen ist die Frage noch unzureichend geklärt, warum die unterschiedlichen Ansätze und therapeutischen Herangehensweisen in unterschiedlichsten Settings doch oft zu ähnlichen Ergebnissen kommen. Eine mögliche Erklärung für nicht-signifikante Unterschiede verschiedener Verfahren könnte sein, dass es sich bei diesen Ergebnissen um methodische Artefakte handelt (Kazdin & Bass, 1989).

Die Entwicklungslinie der Betrachtung von Psychotherapie als generalisierte bzw. allgemeine traditionelle Orientierung wurde später durch strengere Kriterien der Wirksamkeits- und Effektivitätsforschung vor allem in verhaltenstherapeutischen Studien abgelöst. Diese stringenter Art von Forschung verwischt sich allerdings in den seit den 80er Jahren vermehrt durchgeführten Meta-Analysen, die die unterschiedliche Anlage und durchführungstechnische Implikationen der jeweiligen Prüfstrategien nicht mehr adäquat abbilden können. So verwundert es nicht, dass solche Meta-Analysen bei unterschiedlicher Konzeptualisierung zu unterschiedlichen Ergebnissen kommen (vgl. z.B. Crits-Christoph, 1992; Smith, Glas & Miller, 1980; Svartberg & Stiles, 1991).

Viel Diskussionsmaterial stellte die in ihrem Umfang einzigartige Meta-Analyse von Klaus Grawe, Ruth Donati und Friederike Bernauer (1994) bereit. Hier wurde der Versuch unternommen, nicht wie sonst in Meta-Analysen nur eine Stichprobe an Psychotherapiestudien zu untersuchen, sondern die ganze Grundgesamtheit psychotherapeutischer Forschungsergebnisse einzubeziehen. Die sehr hitzig und unsachlich geführte Kontroverse dazu soll hier nicht nachgezeichnet werden. In jedem Fall kann nicht zuletzt aufgrund der Ergebnisse dieser Studie gesagt werden, dass die Auswirkungen von Psychotherapie im Allgemeinen positiv sind, dass die Effektstärken von Psychotherapie deutlich höher als die spontaner Remission reichen, und dass die Effektstärken im Vergleich mit medizinischen Interventionen als durchaus hoch angesehen werden können (s. auch Lipsey & Wilson, 1983). Dass aber die verschiedensten Verfahren gleich, weil unspezifisch wirken, muss, wie bereits angeführt, aus heutiger Sicht bestritten werden („Uniformitätsmythos“).

Insgesamt haben nach Grawe et al. (1994) Verfahren aus dem kognitiv-behavioralen Spektrum die breiteste empirische Basis vorzuweisen. Auch tiefenpsychologische (Kurzzeit-) Therapie und Gesprächstherapie können wissenschaftliche Belege ihrer Wirksamkeit anführen, wenngleich das Ausmaß der Erfolgsraten z.T. wesentlich bescheidener ausfällt. Daraus jedoch allgemeine Schlüsse zu ziehen, etwa dass Psychodrama, systemische und Gestalttherapie unwirksam seien, oder dass tiefenpsychologische Verfahren als Richtlinienverfahren für alle Störungsbereiche indiziert seien, ist dem Forschungsstand nicht angemessen.

In den vergangenen Jahren lassen sich die Prädikate „Verhaltenstherapie“, „Gesprächstherapie“ und anderer in ständiger Entwicklung begriffener Verfahren nur noch begrenzt zur Beurteilung ihrer Güte heranziehen. Denn sowohl die einzelnen Forschergruppen als auch Therapeuten lassen sich zunehmend schlechter solchen Therapieschulen zuordnen. Der Vorschub störungsspezifischer bzw. schulenunabhängiger und „kognitiver“ Methoden hat die Effektivität der Psychotherapien von Therapeutinnen und Therapeuten unterschiedlichster Couleur vermutlich gegenüber früher gesteigert und Unterschiede verwischt – ein Effekt, der in großem Maße durch empirische Therapieforschung produziert wurde (vgl. 1.2.5).

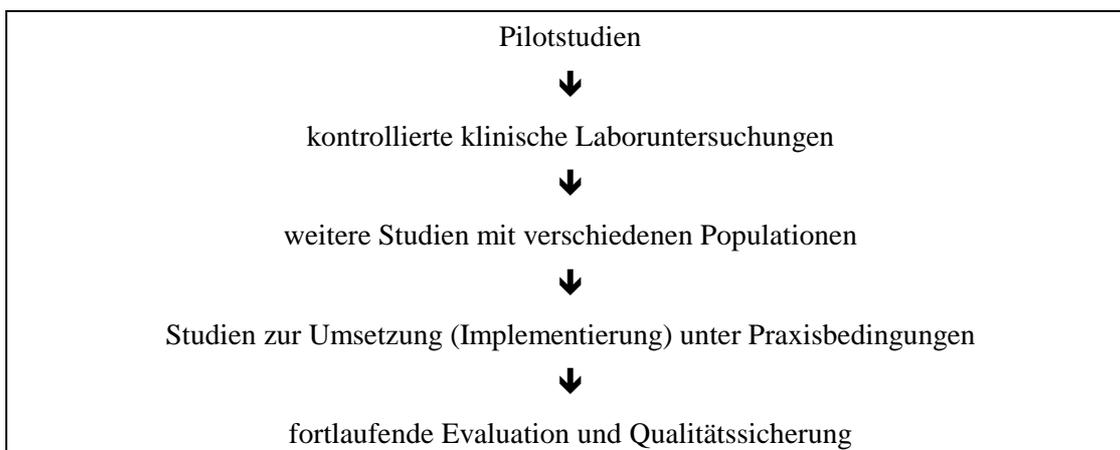
Angemessener als Therapie-Vergleichsstudien ist in der Outcome-Forschung heute die empirische Validierung einzelner Verfahren in klar beschriebenen Settings (z.B. Behandlung bestimmter Störungsbilder; Kontrolle verschiedener Prozessmaße). Kontrollgruppen sind günstigstenfalls Gruppen, die mit einem herkömmlichen bereits validierten Treatment behandelt werden; ebenso möglich ist der Vergleich zu Wartelisten- oder Placebogruppen. Hierbei müssen bei der Interpretation der Ergebnisse mögliche Einschränkungen, etwa durch eine unklare Definition des Placebos oder durch Selektionseffekte berücksichtigt werden (u.U. unterscheiden sich die Patienten, die in der Studie auf einer Warteliste stehen, von der Gesamtpopulation unbehandelter Patienten). Es wurden so in den vergangenen Jahren viele neue und relevante therapeutische Methoden entwickelt, deren spezifische Wirksamkeit klar belegt wurde (vgl. z.B. Fiedler, 1996; s. auch 1.2.5).

Entscheidend für die theoretische Einordnung der vorliegenden Studie ist, dass die untersuchte Therapieform (kognitive Verhaltenstherapie, störungsspezifisch) bei der untersuchten Patientengruppe (Angststörungen) als wirksam erachtet werden darf (kontrollierte Prüfung mit vielen Replikationen). In einem nächsten Schritt soll die Frage aufgegriffen werden, inwieweit die experimentell nachgewiesene Wirksamkeit einer Intervention im Versorgungsalltag repliziert werden kann.

2.1.2 Von der Wirksamkeits- zur Effektivitätsforschung

Allgemeine Anmerkungen zum Forschungsprozess

Neben dem Nachweis der Wirksamkeit geht es auch um die *Evaluation der Umsetzung* psychotherapeutischer Verfahren. In der Psychotherapie-Outcome-Forschung kann man Forschungsfragestellungen folgendermaßen einordnen (Abb. 2.1):

Abbildung 2.1: Von der Wirksamkeits- zur Effektivitätsforschung

Die Abbildung zeigt den Weg von der Erschließung der Erforschung eines Therapieprogramms in *Pilotstudien* (z.B. Ausprobieren neuer Therapiebausteine an Einzelfällen) über die strenge Wirksamkeitsprüfung therapeutischer Interventionen im *kontrollierten Laborexperiment* (z.B. Zeitreihenanalyse oder Kontrollgruppenvergleiche) bis hin zur *Umsetzung* der empirischen Erkenntnisse im *psychotherapeutischen Alltag*.

Die gewählte Forschungsmethodik hängt vom Ort in diesem Prozeß ab. Die Prä-Post-Forschung etwa der Verhaltenstherapie in einer Ambulanz im Rahmen normaler Gesundheitsversorgung stellt lediglich ein *vor-experimentelles Design* dar und weist also nur eine niedrige interne Validität auf (vgl. Cook & Campbell, 1978). Dies wäre in der Tat problematisch, wenn es sich bei der dort durchgeführten Psychotherapie um ein noch ungeprüftes Verfahren handeln würde, d.h. wenn wir uns in der Abbildung im oberen Bereich der Wirksamkeitsstudien bewegen würden. Die Wirksamkeit beispielsweise der Verhaltenstherapie ist für verschiedene Störungsbilder allerdings hinreichend belegt (vgl. Grawe et al., 1994), und nun geht es darum, zusätzlich auch die Effektivität in der klinischen Praxis, also die *klinische Nützlichkeit* zu erforschen. In den USA hat die Forschung bei den als wirksam ausgewiesenen Therapieverfahren bereits den Weg der *Überprüfung der wirksamen Umsetzbarkeit* dieser Verfahren eingeschlagen (Shadish, Navarro & Crits-Christoph, 1997; Wade et al., 1998).

Die Klinische Relevanz von Therapiestudien

Martin Seligman stellt in seinem Artikel „The Effectiveness of Psychotherapy” (1995) die Unterschiede von Wirksamkeits- und Effektivitätsstudie heraus und macht Vorschläge, wie diese beiden Forschungsperspektiven ergänzend eingesetzt werden könnten. Dazu gehört es, in naturalistischen Versorgungs-

studien, wie sie z.B. die amerikanische Verbraucherzeitschrift *Consumer Reports* (vergleichbar mit dem deutschen Organ der *Stiftung Warentest*) in einer großen Umfrage zur Inanspruchnahme und zu den Effekten psychosozialer Dienste durchgeführt hat, auf einige der strengen Kriterien experimenteller und quasi-experimenteller Studien zu verzichten. Denn die externe bzw. ökologische Validität könnte z.B. durch die Auslese der Patienten ohne Komorbidität, Manualisierung, standardisierte Therapiedauer, Kontrollgruppen mit Zufallszuweisung etc. gefährdet werden. Außerdem sollten die Ergebnisse aus Therapiestudien von Forschergruppen repliziert werden, die nicht an der Entwicklung der zu überprüfenden Intervention beteiligt waren, denn besonders starke Identifikation mit und eine außergewöhnliche Expertise in dem beforschten Ansatz können zu unrealistisch hohen Effekten führen.

Shadish, Matt, Navarro und Siegle (1997) stellen *Kriterien für die klinische Relevanz von Psychotherapiestudien* auf und kommen nach der Analyse mehrerer Meta-Analysen zur Wirksamkeitsforschung zu dem ernüchternden Schluß, dass nur ein Bruchteil der Studien unter klinisch relevanten Bedingungen durchgeführt wurde (s. auch Goldfried & Wolfe, 1998; Matt & Cook, 1994). Auch im deutschsprachigen Raum gibt es seit einiger Zeit vermehrte Forderungen nach Therapieforschung in „naturalistischen“ Settings. In diesem Zusammenhang sprechen Linden (1989) und Tschuschke & Kaechele (1996) von der „Phase IV der Therapie-Evaluation“. Auch Rief (1994) plädiert für eine praxisorientierte Psychotherapieforschung.

Andererseits dürfen natürlich die Grenzen der Interpretierbarkeit – vor allem die Selbstselektion der Stichprobe und die subjektive post-hoc Befragung als einzige Datenquelle – von naturalistischen Studien wie von *Consumer Report* nicht übersehen werden. Außerdem sind im Zusammenhang mit den verschiedenen Validitätskriterien auch statistische Fragen von Bedeutung. Die zunehmende klinische Relevanz einer Therapiestudie ist direkt mit einem Anstieg der Heterogenität der untersuchten Patienten und Behandlungen verbunden (z.B. aufgrund Komorbidität oder unterschiedlicher Therapiedauer) und damit auch mit einer Erhöhung der Varianz der Ausgangsdaten. Damit sinkt jedoch die Wahrscheinlichkeit, statistisch signifikante Ergebnisse zu finden. Es können aber weder die Effekte einer Intervention noch die Stichprobe beliebig vergrößert werden. Eine versuchsplanerische Homogenisierung ist also bisweilen notwendig, um eine realistische Chance zu behalten, *reell vorhandene* Effekte aufzudecken.

Die unterschiedlichen Herangehensweisen im Spannungsfeld von interner und externer Validität hängen somit von der jeweiligen Fragestellung ab und sollten sich komplementär ergänzen. Seligman (1995) erklärt, dass Wirksamkeitsstudien in erster Linie sehr geeignet sind, um zu entscheiden, ob eine Behandlung für eine bestimmte Störung besser geeignet ist als eine andere. Seligman führt dabei folgende Vorzüge der Wirksamkeitsstudie an:

- *Patienten werden zufällig der Therapie- und Kontrollgruppe zugeordnet:* Es gibt also auch Gruppen ohne Behandlung und Gruppen mit herkömmlicher oder Placebobehandlung, an deren Wirksamkeit sowohl Patienten wie Therapeuten glauben. Patienten- und Therapeutenvariablen, die einen spezifischen Einfluss auf Therapieverlauf und –Ergebnis ausüben, sollten sich dabei ausmitteln bzw. lediglich als „*unsystematische Störung*“ gelten.
- Behandlungen sind oft manualisiert. Die Einhaltung des Manual-Standards wird durch Videoaufzeichnungen überprüft und bewertet; somit kann die sog. *Behandlungsintegrität* gesichert werden.
- Die *Standardisierung* der Anzahl von Behandlungssitzungen, Operationalisierung von Zielen, Diagnosen und Katamnesen verringert die Varianz der Ausgangsdaten und steigert damit die interne Validität zusätzlich.

Gleichzeitig zeigt Seligman aber auf, dass die Frage, ob eine Behandlung unter sehr gut kontrollierten Bedingungen besser ist als die andere, und die Frage, was in der Praxis tatsächlich funktioniert, völlig verschieden sind. Daher propagiert er die Effektivitätsstudie, die untersucht, wie sich Patienten unter den aktuellen Bedingungen der klinischen Praxis verändern, als zusätzlichen Weg zur empirischen Validierung von Psychotherapie. Psychotherapie – so wie sie im klinischen Alltag praktiziert wird – weist bestimmte Merkmale auf, die in der klassischen Wirksamkeitsstudie nicht enthalten sind:

- Patienten kommen normalerweise aufgrund eines aktiven Suchprozesses zu ihrem Therapeuten. Sie wählen sich ihre Therapieform (je nach Wissensstand) und ihren Therapeuten nach Kriterien wie Sympathie, wahrgenommene Kompetenz und Vertrauenswürdigkeit aus. Dies gilt vor allem für Patienten, die unabhängige Praktiker besuchen, weniger für Patienten von Kliniken und ähnlichen Einrichtungen. Die Patienten in Wirksamkeitsstudien hingegen werden in einem passiven Prozeß per Zufall den Therapiebedingungen zugeordnet (Howard, Orlinsky & Lueger, 1994).
- Patienten in der Therapiepraxis haben oft mehrere Probleme. Psychotherapie soll parallele und interagierende Schwierigkeiten lösen. In Wirksamkeitsstudien werden Personen mit klar definierten Störungen ausgewählt.
- Psychotherapie hat gewöhnlich keine vorher für alle gleich festgesetzte Sitzungsdauer. Sie geht normalerweise so lange, bis der Patient davon merklich profitiert hat oder er die Therapie abbricht. Wirksamkeitsstudien sind von vorneherein auf einen festen Umfang hin angelegt (meist ca. 15 Sitzungen in Studien zu kognitiv-behavioralen Therapieverfahren).

- Psychotherapie ist selbstkorrigierend. Wenn eine Technik keinen Erfolg bringt, wird oft eine andere Technik oder Modalität eingesetzt. Die Wirksamkeitsstudie definiert den Einsatz der Technik vorher, sie wird innerhalb einer Modalität nach einem Manual durchgeführt.
- In der psychotherapeutischen Praxis wird neben der Reduzierung eines speziellen Symptoms und der Verminderung einer bestimmten Störung fast immer zusätzlich besonderes Augenmerk auf die Verbesserung der generellen Funktionsfähigkeit des Patienten gelegt.

Implikationen für die vorliegende Kosten-Effektivitäts-Studie

Die vorliegende Arbeit versucht, die Vorzüge beider Methoden zu kombinieren. Einerseits wird wie in klassischen Wirksamkeitsstudien zu mehreren Meßzeitpunkten vielfältige diagnostische Information erhoben. Das therapeutische Vorgehen war zwar nicht manualisiert, aber die Therapeutinnen und Therapeuten waren angehalten, störungsspezifische kognitiv verhaltenstherapeutische Behandlungen durchzuführen (z.B. sensu Schneider & Margraf, 1998); sie unterlagen dabei regelmäßiger Supervision.

Andererseits waren die Therapien individuell auf die einzelnen Patienten zugeschnitten und damit auch selbstkorrigierend und flexibel bezüglich der Behandlungsdauern. Das Setting entsprach der ambulanten Normalversorgung: die Patienten wurden der Therapie nicht zufällig zugeordnet, sondern haben sie selbst ausgewählt; die Patienten haben teilweise multiple Diagnosen; Schwierigkeiten wegen Finanzierungsfragen (Kostenerstattung) konnten zu Therapieabbrüchen bzw. zur Nicht-Aufnahme einer Therapie nach den probatorischen Sitzungen führen. Die Zahl der untersuchten Patienten ist darüberhinaus beeindruckend groß (N = 493 bei der Kosten-Effektivitäts-Analyse, für die die hier gemachten Ausführungen die größte Bedeutung haben). Die Behandlungseinrichtungen waren zwar an universitäre Lehrstühle angegliedert, jedoch scheint durch die Vielzahl der Behandler in den vier verschiedenen Ambulanzen eine Heterogenität gegeben, die eher dem klinischen Alltag entspricht als die bewußte Standardisierung bei universitären kontrollierten Wirksamkeitsstudien.

Diese hohe Realitätsnähe entspricht den Kriterien, die Seligman (1995) für die Bewertung von Therapie „unter Alltagsbedingungen“ oder Shadish et al. für die „klinische Relevanz“ einer Studie aufstellen; allerdings mit der Einschränkung, dass bezüglich der konkreten, an modernen Forschungsergebnissen optimierten Behandlungstechniken nicht davon ausgegangen werden kann, dass diese in der „Durchschnittspraxis“ regelmäßig ebenso durchgeführt werden.

Die Studie kann natürlich nicht den Anspruch erheben, eine Wirksamkeitsstudie zu sein; es war aber auch gar nicht das Ziel, Wirkfaktoren zu finden

oder die Wirksamkeit bestimmter Interventionen experimentell nachzuweisen. Vielmehr handelt es sich um eine Arbeit aus dem Bereich der Effektivitätsforschung, deren Ziel es in erster Linie ist, die allgemeine Qualität der Therapie einer real existierenden Institution unter Kostengesichtspunkten zu prüfen. In der abschließenden Diskussion wird vor diesem Hintergrund auf die zentrale Frage eingegangen, ob die hier untersuchten Effekte tatsächlich auf die Therapie zurückgeführt werden können. Dabei wird der Standpunkt vertreten, dass die Interpretation als echte Therapieeffekte hoch plausibel ist, insbesondere da alternative Erklärungen durch die Daten kaum gestützt werden.

Zusammenfassend kann also festgehalten werden, dass sich kontrollierte klinische Studien nicht in idealer Weise mit Erhebungen unter psychotherapeutischen Praxisbedingungen kombinieren lassen. Die Operationalisierung, aber auch die Terminologie strengerer wissenschaftlicher Studien sind mit den häufig verwaschenen und unklaren Vorgehensweisen im psychotherapeutischen Sektor nicht in vollständige Deckung zu bringen. Diese Divergenz zwischen beiden Ansätzen wächst erfahrungsgemäß gerade in einem Gebiet wie der klinischen Psychologie weiter an, da die zunehmende Differenzierung in der wissenschaftlichen Forschung nur noch ausschnittsweise von der Praxis rezipiert und umgesetzt wird (zu dieser Thematik des Transports empirisch fundierter oder „evidenzbasierter“ Behandlungen in den klinischen Alltag s. Jacobi & Kosfelder, 2000; Kosfelder, Poldrack & Jacobi, 2000).

Die vorliegende Arbeit untersucht die klinisch-psychologische Behandlung von Angststörungen, wie sie mit einem „scientist-practitioner“-Ansatz in psychologischen Ambulanzen moderner Prägung – allerdings unter den Randbedingungen eines normalen Versorgungsalltags – durchgeführt wurde. Somit können die Ergebnisse nicht ohne weiteres auf „die“ Psychotherapie übertragen werden, so wie sie von der Gesamtheit der geschätzt 20000 ambulanten Psychotherapeutinnen und Psychotherapeuten in Deutschland durchgeführt wird.

2.2 Kosten-Effektivitäts-Analysen

Die Kosten-Effektivitäts-Analyse einer Behandlung berücksichtigt neben dem Therapieerfolg bzw. dem Outcome auch die für diesen Outcome aufgewendeten Ressourcen. Diese aufgewendeten Ressourcen werden in den Kosten ausgedrückt, die durch die Behandlung entstanden sind. Man möchte damit herausfinden, was die Erreichung eines bestimmten Therapieergebnisses (durchschnittlich) kostet.

Es müssen also sowohl Effektivitäts- als auch Kostendaten vorliegen, um die Kosten-Effektivität zu bestimmen. Relevante Darstellungen über die Durchführung von Kosten-Effektivitäts-Analysen, auf die sich auch die folgenden Ausführungen stützen, sind Bühringer und Hahlweg (1986), Drummond, Stoddart und Torrance (1987), Warner und Luce (1982), sowie Yates und Newman (1980) zu entnehmen.

Der Forschungsstand zu empirischen Kosten-Effektivitäts-Studien im Bereich der Psychotherapie wird gemeinsam mit dem Forschungsstand zu Kosten-Nutzen-Analysen referiert (Abschnitt 2.3).

Unerlässlich ist, die Randbedingungen einer Kosten-Effektivitäts-Analyse (Störung, Therapiesetting, Art der Effektivitätsmessung) genau zu spezifizieren, insbesondere wenn man die Kosten-Effektivität mehrerer Behandlungen vergleichen möchte. So wäre es z.B. unfair, die Kosten-Effektivität in einer Behandlungseinrichtung, in der üblicherweise sehr teure bzw. prognostisch vergleichsweise ungünstige Alkoholismusbehandlungen durchgeführt werden, mit derjenigen der ambulanten Angstbehandlung aus dieser Studie zu vergleichen.

2.2.1 Allgemeines Vorgehen bei einer Kosten-Effektivitäts-Analyse

In einer Kosten-Effektivitäts-Analyse bietet es sich zunächst an, die behandelten Patienten hinsichtlich ihres Outcomes in verschiedene Klassen einzuteilen und die jeweiligen Therapiekosten für jede Outcome-Modalität und jede einzelne Klasse zu ermitteln. Yates (1985) empfiehlt in diesem Zusammenhang Kosten-Effektivitäts-Matrizen, in denen differenziert dargestellt werden kann, welchen Veränderungen welche Therapiekosten entsprechen. Entsprechende Darstellungen finden sich in Abschnitt 3.3.

In einem weiteren Schritt kann man diese eher „impressionistischen“ Ergebnisse standardisieren, um damit die Möglichkeit zu haben, die Therapie auch mit anderen Behandlungsmethoden unter Einbezug von Effektivität *und* Kosten zu vergleichen. Dabei werden die Kosten einer Behandlung in einem Quotienten oder über eine Differenzbildung in Relation zur Effektivität gesetzt, also

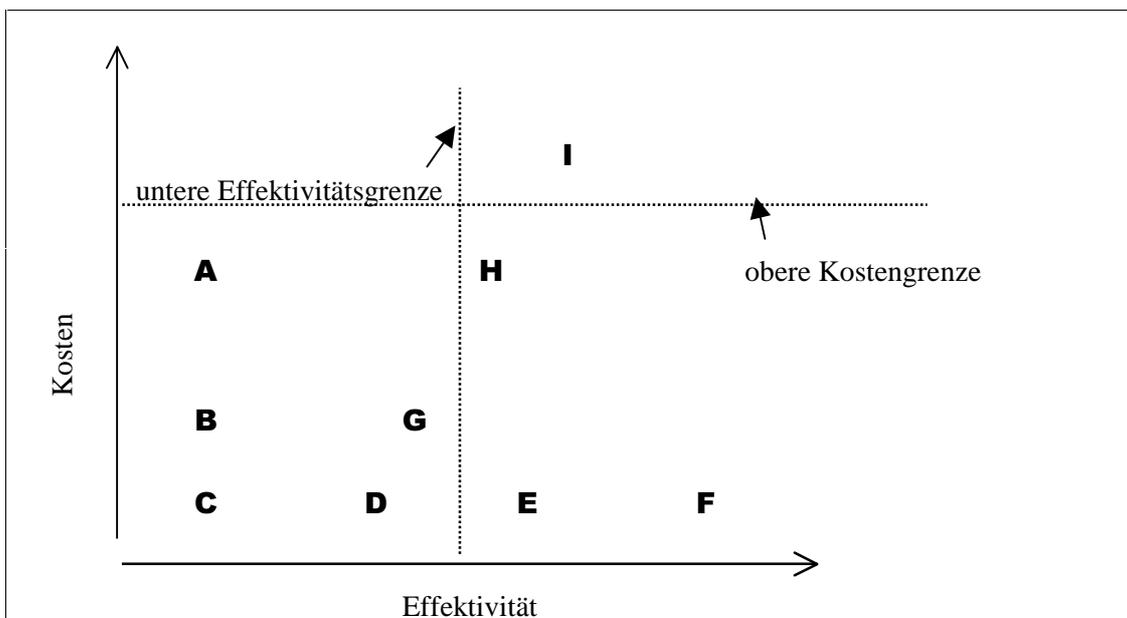
$$\text{Kosteneffektivität} = \frac{\text{Durchschnittliche Therapiekosten (DM)}}{\text{Rate „erfolgreich behandelte“ Patienten}}$$

Hierfür muß für die untersuchte Behandlung ein – möglichst konsensfähiges – Kriterium gefunden werden, mit dem man die Patienten in „erfolgreich behandelt“ vs. „nicht erfolgreich behandelt“ einteilen kann. Ein Beispiel für die Verwendung dieses Index findet sich bei Koester et al. (1982).

Bei der Interpretation der Ergebnisse ist stets zu berücksichtigen, dass sich die ermittelte Kosten-Effektivität auf eben dieses gewählte Erfolgskriterium bezieht – es ist also nicht möglich, „die“ Kosten-Effektivität einer Therapiemethode zu bestimmen. In der vorliegenden Kosten-Effektivitäts-Analyse wurden angesichts der Vielzahl der zur Verfügung stehenden Maße mehrere Kriterien herangezogen und jeweils eigene Kosten-Effektivitäten berechnet.

Ebenso wie sinnvolle Kriterien für einen Therapieerfolg bzw. für eine von der Behandlung geforderte Mindest-Effektivität gefunden werden müssen, so muß auch an eine Kosten-Obergrenze gedacht werden. Denn wie soll verfahren werden, wenn eine Behandlung, die zu sehr guten Ergebnissen kommt, die gleiche Kosten-Effektivität hat wie eine andere Behandlung, die deutlich billiger ist und bereits ausreichend gute Ergebnisse liefert? Yates & Newman (1980) veranschaulichen dieses Problem an einem Beispiel (Abb. 2.2).

Abbildung 2.2: Entscheidung für eine Behandlungsalternative bei gleicher Kosten-Effektivität (nach Yates & Newman, 1980, S.138); Erläuterung im nachfolgenden Text.



Die Behandlungsalternativen A, B und C weisen alle nur eine geringe Effektivität auf. Wenn dennoch eine Behandlung daraus gewählt werden *muß* (z.B. weil keine andere Alternative verfügbar ist und man die niedrige Effektivität immer noch einer Nicht-Behandlung vorzieht), so würde man sich aufgrund der besseren Kosten-Effektivität für C entscheiden. Bei den Alternativen D, E, und F fällt die Entscheidung leicht: F hat die mit Abstand beste Kosten-Effektivität aufzuweisen.

Schwieriger wird es, wenn man eine Wahl zwischen G, H und I treffen müsste: die Quotienten aus Kosten und Effektivität zeigen nur minimale Unterschiede. Nun können zusätzliche Kriterien weiterhelfen. Entscheidet man sich aber für eine Kosten-Obergrenze, so scheiden zu teure Maßnahmen aus (hier: I). Legt man eine Effektivitäts-Untergrenze fest, so scheiden zu schwache Maßnahmen aus (hier: G). H bliebe dann übrig, weil sowohl Effektivität als auch Kosten noch im akzeptablen Bereich liegen.

2.2.2 Therapieerfolgsmaße und Therapieerfolgskriterien

Die Bestimmung der Behandlungskosten wurde bereits eingangs erläutert; in der Regel werden die direkten tangiblen Kosten im Betriebskostenansatz ermittelt. Weitaus schwieriger ist die Definition dessen, was man in der Kosten-Effektivitäts-Analyse diesen Kosten gegenüberstellt: Wie kann Therapieerfolg gemessen werden? Welche Kriterien müssen erfüllt sein, damit ein Patient als „erfolgreich behandelt“ gilt?

Froyd und Lambert (1989, zit. nach Schulte, 1993) fanden in 348 Psychotherapiestudien, die zwischen 1982 und 1988 veröffentlicht worden waren, 1430 verschiedene Maße für den Therapieerfolg. Die Messung von Therapieerfolg ist also äußerst vielfältig – jedoch nicht beliebig (Schulte, 1993). Bei der Definition der Therapieerfolgskriterien muß zwischen den inhaltlichen und methodischen Aspekten unterschieden werden, d.h., wie bei der Messung jedes anderen psychologischen Konstruktes müssen zum einen Festlegungen zu den Variablen getroffen werden, die das Konstrukt inhaltlich erfassen, und zum anderen, wie diese Variablen operationalisiert und gemessen werden sollen.

Lambert und Hill (1994, S. 72 f.) demonstrieren an einem Beispiel, wie die „Erfolgsquote“ derselben Studie in Abhängigkeit von der Wahl verschiedener Erfolgsindikatoren zwischen 69 % und 6,3 % schwankt. Mit Blick auf den zunehmenden Legitimationsdruck der Psychotherapie (z.B. gegenüber Kostenträgern) einerseits, vor allem aber auf die Lage der hilfeschuchenden Patienten andererseits, erscheint ein solcher Zustand unhaltbar.

Um zu einem schulübergreifend gültigen Konsens über die Erfolgsmessung zu kommen, rekurriert Schulte (1993) auf das *allgemeine Krankheitsmodell*: wenn das (gesellschaftlich vorgegebene) Ziel einer Therapie in der Heilung

bzw. Linderung von Störungen mit Krankheitswert liegt, dann sollte „Therapieerfolg“ (u.a.) in einem Rückgang der Beschwerden und ihrer Folgen bestehen. Er verweist zunächst auf *störungsspezifische Maße*, die sich an den im DSM für jede Diagnosekategorie beschriebenen Symptomen orientieren. Vorschläge zu konkreten Meßinstrumenten für verschiedene Störungen stammen z.B. von Strupp, Horowitz und Lambert (1997), im deutschsprachigen Raum von Fydrich und Mitarbeitern (1996).

Darüber hinaus ist auch eine *störungsübergreifende Erhebung der Symptombelastung* etwa mit einer Symptom-Checkliste (z.B. SCL-90-R) notwendig. Schließlich sollten *Maße auf der Ebene der Störungsfolgen* (Demoralisation, Verlust von Kontrolle, Selbstwert und sozialer Einbindung, Krankenrolle) Verbesserungen in diesen Bereichen erfassen. Als Datenquelle kommen hier neben Patienten- und Therapeuteneinschätzungen auch Angaben relevanter Bezugspersonen in Frage.

Die Ausgestaltung einer Psychotherapie sollte sich nicht zuletzt an den individuellen *Zielen* des Patienten orientieren. Diese können andere als Symptomreduktion sein. Dementsprechend müßte eine Erfolgsmessung auch auf individuelle Zielerreichung bezug nehmen (etwa mit der Methode des *goal attainment scaling*; s. z.B. Heuft et al., 1996).

Zunächst müssen also die Variablen und die Instrumente festgelegt werden (s. Kasten 2.1; vgl. auch die exemplarische Begleitforschung im stationären Setting; Dilcher, 2000, oder in psychotherapeutischen Ambulanzen, Grosse Holtforth & Regli, 2000; Sekot & Jacobi, im Druck). Im nächsten Schritt muss nun entschieden werden, wie die erhobenen Daten zu interpretieren sind. Auch hinsichtlich der *Kriterien*, wann ein Ergebnis als „Erfolg“ angesehen werden kann, herrscht keineswegs Einvernehmen. Retrospektive Befragungen von Patienten und Therapeuten nach ihrer Zielerreichung, ihrer aktuellen Zufriedenheit oder nach ihrer Bewertung der Veränderungen können als *direkte* Erfolgsmessungen bezeichnet werden. Sie korrelieren in der Regel nur mäßig mit *indirekten* Outcome-Maßen, beispielsweise dem Rückgang der Symptome oder einer Selbstwertsteigerung, wie sie mit speziellen Fragebögen erfasst wird. Und auch hier muß zwischen dem Ausmaß an Veränderung im Vorher-Nachher-Vergleich und einem möglichst guten Abschlusswert (ungeachtet des Ausgangswertes) differenziert werden.

Kasten 2.1: Auswahl typischer Konstrukte bzw. Fragebogenmaße, die derzeit zur Effektivitätsbeurteilung herangezogen werden (vgl. Dilcher, 2000; Fydrich et al., 1996; Grosse Holtfort & Regli, 2000; Schulte, 1993; Sekot & Jacobi, im Druck).

Symptomebene: Störungsspezifische Maße

- Depressivität: Beck Depressions Inventar (BDI), Hamilton-Skala
- Ängstlichkeit allgemein: State/Trait Angst-Inventar (STAI), Anxiety Sensitivity Index (ASI), Beck Angst Inventar (BAI),
- Phobische Angst: Fear Questionnaire
- Agoraphobische Angst : Mobilitäts-Inventar (MI, allein und begleitet), Agoraphobic Cognitions Questionnaire (ACQ)
- Paniksymptome: Body Sensations Questionnaire (BSQ)
- Soziale Ängste : Social Phobia Scale (SPS), Social Interaction Anxiety Scale (SIAS)
- Zwänge : Hamburger Zwangs-Inventar (HZI), Y-BOCS
- Hypochondrie: Whiteley-Index
- Somatisierung: Screening für somatoforme Störungen (SOMS-2)
- Essstörungen: Fragebogen zum Essverhalten (FEV)
- Partnerschaftsprobleme: Partnerschaftsfragebogen (PFB)
- Alkoholmissbrauch/-Abhängigkeit: Münchener Alkoholismus-Test (MALT)

Übergreifende Symptombelastung:

- Symptomcheckliste (SCL-90-R; Skalen: Somatisierung, Zwanghaftigkeit, Unsicherheit im Sozialkontakt, Depressivität, Ängstlichkeit, Aggressivität/Feindseligkeit, phobische Angst, paranoides Denken und Psychotizismus, sowie Gesamtbelastung)

Störungsfolgen/Beeinträchtigungsebene:

- Items zu Beeinträchtigung bei Arbeit/Ausbildung, Familienleben/häusliche Pflichten, und Partnerschaft/Sozialkontakte
- Global Assessment of Functioning (GAF; Achse V des DSM)
- Fragebögen zu Störungskosten (Vorbehandlungen, Arbeitsunfähigkeit)

globale Ebene:

- Items zu Verbesserungen/Verschlechterung durch die Therapie
- Veraenderungsfragebogen des Erlebens und Verhaltens (VEV)

individuelle Ziele:

- Goal Attainment Scaling (GAS)
- Fragebogen zur Analyse Motivationaler Schemata (FAMOS)

2.2.2.1 Ausmaß der Veränderung: Effektstärken

Das Ausmaß der Veränderung wird in der Regel als *Effektstärke* bezeichnet, mit der Gruppenunterschiede (bzw. Mittelwertsveränderungen) standardisiert ausgedrückt werden (vgl. z.B. Rosnow & Rosenthal, 1988; Steyer, Hannover, Telser & Kriebel, 1997). Die allgemeine Formel für Effektstärken entspricht dem Mittelwertsunterschied zwischen zwei Gruppen relativiert an der (Prä-) Streuung (die sich in kontrollierten Studien in den Gruppen nicht unterscheiden sollte).

$$ES = \frac{M_a - M_b}{s}$$

Formel 2.1: Grundformel zur Effektstärkenberechnung. ES: Effektstärke; M_a , M_b : Mittelwerte der verglichenen Gruppen; s : (Prä-) Streuung

In der vorliegenden Studie wurde ein Prä-Post-Vergleich durchgeführt und nicht auf (z.B. unbehandelte) Kontrollgruppen Bezug genommen. Grawe (1992) fand jedoch, dass bei einer Stichprobe von Untersuchungen die Effektstärken von insgesamt 111 Kontrollgruppen einen „deutlichen Schwerpunkt... um Null und leicht darüber“ haben. Weiter fanden Grawe et al. (1994), dass Effektstärkenwerte in Prä-Post Messungen mit denen der Kontrollgruppenmessung zwischen 0,83 und 0,98 korrelieren. Die Prä-Post-Effektstärke ist also ein äußerst valider Schätzer für Therapieeffekte.

Allerdings können therapieunspezifische Effekte und therapiespezifische Effekte ohne Kontrollgruppen mit unterschiedlichen Behandlungen nicht getrennt werden. Man kann also nicht sagen, was genau an der Therapie die Veränderung bewirkt hat. Dies ist aber nicht Aufgabe der hier vorgenommenen Effektivitätsmessung eines Behandlungsverfahrens im klinisch relevanten Setting, sondern Aufgabe der Prozeß- und der Grundlagenforschung. Einen Beitrag zur Klärung dieser wichtigen Frage kann aber auch die Effektivitätsmessung leisten, sobald die Ergebnisse mehrerer solcher Arbeiten meta-analytisch miteinander verglichen werden. Dazu sind Prä-Post-Effektstärken durchaus geeignet. Dem Problem, dass in Meta-Analysen spezifische Befunde meist wieder verwischt werden, muss dabei durch Standardisierung der Forschungsdesigns und der Erhebungsmethodik über mehrere Studien hinweg und über die Bereitstellung der für eine Meta-Analyse notwendigen Angaben in Publikationen begegnet werden (vgl. Czienskowski, 2000).

Hartmann & Herzog (1995) diskutieren unterschiedliche Erweiterungen der Grundformel, indem etwa unterschiedliche Gruppengrößen Prä/Post, unterschiedliche Streuungen Prä/Post oder die Korrelation der Meßwerte Prä/Post einbezogen werden. Zwei dieser Formeln seien hier vorgestellt:

$$ES = \frac{M_{t1} - M_{t2}}{\sqrt{\frac{(N_{t1} - 1)s_1^2 + (N_{t2} - 1)s_2^2}{N_{t1} + N_{t2} - 2}}}$$

Formel 2.2: Effektstärkenberechnung nach Hartmann, Herzog und Drinkmann (1992), $t1$ ist die Prä-Messung, $t2$ die Post-Messung.

In Formel 2.2 wird mehr Information verarbeitet als in der Grundformel. Erstens werden ggf. bestehende Unterschiede der Stichprobengröße Prä und Post (etwa durch *dropouts*) berücksichtigt. Wichtiger noch ist der Einbezug sowohl der Prä- als auch der Post-Varianz:

- Steigt die Varianz von Prä zu Post, dann muß die Behandlung zu differentiellen Ergebnissen geführt haben. Damit ist bei gleichem durchschnittlichen Effekt die Grundwahrscheinlichkeit für den einzelnen Patienten erhöht, sich nicht bedeutsam zu verbessern. Eine Vergrößerung der Varianz kommt nach Grawe (1990) immer einer Aufforderung gleich, die Wirkungen der Therapie genauer zu betrachten. Eine Therapieform – insbesondere eine, die bereits den Weg in die Versorgung gefunden hat – sollte also keine Varianzvergrößerung zur Folge haben. Deshalb wird in dieser Formel eine solche Varianzvergrößerung mit einer Verminderung der Effektstärke „bestraft“.
- Sinkt die Varianz von Prä nach Post, so bedeutet dies in der Regel bei gleichem durchschnittlichen Effekt, dass sich die Patienten angenähert haben, weil sie sich alle „in die richtige Richtung“ verändert haben. Im Idealfall sinkt die Varianz aufgrund eines Deckeneffektes (alle Patienten im gesunden Bereich) auf Null; diese Tendenz wird in der Formel „belohnt“.

Eine andere Möglichkeit zur Berechnung von Prä-Post-Effektstärken ist die Formel 2.3 nach McGaw & Glass (1980). Sie berücksichtigt neben den Prä-Post Streuungen auch die Korrelation zwischen Prä- und Post-Werten. Niedrige Korrelationen verringern hier die Effektstärke.

$$ES = \frac{M_{t1} - M_{t2}}{\sqrt{(s_{t1}^2 + s_{t2}^2 - 2r_{t1t2}s_{t1}s_{t2})}}$$

Formel 2.3: Effektstärkenberechnung nach McGaw & Glass (1980); $t1$ ist die Prä-Messung, $t2$ die Post-Messung

Hierbei werden Therapien mit niedrigen Prä-Post-Korrelationen bestraft, da dies ebenfalls auf differentielle Effekte schließen läßt. Dies ist aber nur sinnvoll, wenn kein (statistischer oder natürlicher) Deckeneffekt für die niedrige Korrelation verantwortlich ist. Erkennen kann man einen solchen Deckeneffekt an einer Verringerung der Streuung in der Post-Messung bei gleichzeitiger deutlicher Mittelwertsverbesserung. Ein zweites Problem ist, dass die berechnete relativierende Varianz bei extrem hohen Prä-Post-Korrelationen kleiner werden könnte als eine der beiden eingehenden Varianzen. So wird bei hoher Korrelation und geringer Varianzänderung die Effektstärke auch bei geringer Mittelwertänderung überschätzt.

In dieser Studie wird deshalb die Formel 2.2 nach Hartmann, Herzog und Drinkmann (1992) verwendet. Auch sie bestraft Therapien, bei denen sich die Varianz aufgrund der Therapie erweitert mit einer geringeren Effektstärke und bindet das erhöhte Risiko der Betroffenen für eine Verschlechterung durch die Therapie ein. Dies kommt aber nur bei großen Mittelwertänderungen zum Tragen. Je größer die Mittelwertänderung ist, um so mehr kommt die Varianzverminderung aufgrund eines Deckeneffekts zustande, so dass es nur gerechtfertigt ist, die Therapie dafür zu belohnen.

Wie ist nun eine so ermittelte Effektstärke zu bewerten? Leider finden sich in der Literatur wenig konkrete Aussagen, ab welchen Werten Effektstärken wirklich bedeutsam sind. Eine Effektstärke von 1 (also die (positive) Veränderung des Patientenmittelwerts im Vergleich zu vor der Behandlung um eine Standardabweichung) bedeutet, dass es etwa 84% der Patienten nach der Behandlung (statistisch) signifikant besser geht als dem Durchschnittspatienten vorher. Cohen (1988) bezeichnet aus der Sicht des β -Risikos Effektstärken von 0.2 als geringe, Werte von 0.5 als mittlere und ab 0.8 als hohe Effektstärken. Rosenthal (1982, 1990) verweist auf die häufige Unterschätzung von „moderaten“ Effektstärken. Er zeigt, dass eine Effektstärke von 0.85, wie sie von Smith, Glass und Miller (1980) ermittelt wurde, ein äußerst bedeutungsvolles Ergebnis ist. Grawe (1992) zeigt, dass die durchschnittliche Effektstärke für 111 unbehandelte Kontrollgruppen bei 0.10 liegt. Effektstärken über 0.6 waren bei diesen Kontrollgruppen sehr selten gefunden worden (N=7). Man kann also Effektstärken ab 0.6 als relevant, ab 0.8 als substantiell und über 1.0 als hoch ansehen.

2.2.2.2 *Klinische Signifikanz*

Ein Großteil der psychotherapeutischen Forschung ist *konstruktbezogen* derart, dass unter Zuhilfenahme eines theoretischen Konstrukts wie z.B. „Ängstlichkeit“ etwas per Fragebogen gemessen wird. Eine Verbesserung eines Patienten hinsichtlich dieses Konstrukts wird dann im Sinne einer Patientenheilung interpretiert. Diese impliziert aber einen *kategorialen* Ansatz mit den Extrempunkten „gesund“ und „krank“. Diese Kategorien anhand dimensionaler Kon-

struktive sinnvoll zu operationalisieren – also etwa cut-offs für einen Übergang von „krank“ nach „gesund“ zu finden – stellt ein nicht unerhebliches Problem dar.

In der Literatur finden sich verschiedene Vorschläge, wie über die Berechnung statistischer Signifikanztests hinaus ihre klinische Bedeutsamkeit (*clinical significance*; z.B. Jacobson & Truax, 1991; Kazdin, 1994) erfaßt werden kann. Dies ist sinnvoller als eine alleinige Beurteilung der Größe der Veränderungen. Außerdem kann mit der klinischen Signifikanz im Gegensatz zur durchschnittlichen Gruppen-Veränderung auf *individuelle Veränderungswahrscheinlichkeiten* geschlossen werden.

Die klinische Praxis zeigt, dass es wenig Sinn macht, Patienten dichotom in entweder „gesund“ oder „krank“ einzuteilen. So scheint es eine große Gruppe zu geben, die zwar im Rahmen einer Therapie deutliche Fortschritte macht, aber danach immer noch nicht klar der „gesunden“ Population zugerechnet werden kann. Eine Differenzierung stellt die Hinzunahme der Kategorie „subklinisch“ dar. Sowohl Verbesserungen vom „nur“ subklinischen Level aus nach gesund als auch von krank nach „nur noch“ subklinisch können individuell sehr bedeutsam sein. Zudem finden sich viele Patienten auch nach der Therapie trotz individuell erlebten Verbesserungen in Fragebogenmaßen immer noch in einem auffälligen Bereich, und es wäre zu streng, sie deshalb von vorneherein als „nicht klinisch signifikant gebessert“ zu bezeichnen.

In Kasten 2.2 wird ein mögliches Vorgehen zur Konzeptualisierung klinischer Signifikanz in der Erfolgsmessung illustriert.

Besonders beachtet werden sollten eventuelle schädliche (Neben-) Wirkungen der Behandlung. Im Gegensatz zur medizinisch-pharmakologischen Forschung gibt es für die Psychotherapie derzeit noch keine etablierte Forschungsperspektive, was deren möglichen schädliche Auswirkungen betrifft (vgl. Jacobi, 2001; Mays & Franks, 1985). Deshalb wird in der vorliegenden Studie der Anteil dieser Verschlechterungen zumindest berichtet.

Es gibt mehrere Möglichkeiten, die Kategorien bzw. die Klassengrenzen zu operationalisieren. Dazu ist die Orientierung an einem Standard, z.B. „Normales Funktionieren“, nötig. Gibt es keine Normwerte für normale Populationen, so soll der Post-Wert außerhalb der Werte der dysfunktionalen Population in Richtung der funktionalen Population angesiedelt sein. Der Post-Meßwert muß also eher der Normalpopulation als der dysfunktionalen Population angehören (Jacobson et al., 1984).

Diese Forderung ist für bestimmte Störungsformen klinisch nicht sinnvoll. So ist zum Beispiel bei der Therapie von Adipositas die individuelle Festlegung eines Zielgewichtes und die Vermittlung von Strategien dieses beizubehalten wesentlich sinnvoller, als die Orientierung an einem Normgewicht für die Bevölkerung. Für andere Störungsformen, wie Depression und viele Angststörungen ist jedoch eine Orientierung an den Normwerten von funktionalen Populationen durchaus sinnvoll, da die Erreichung dieses Niveaus oft

eine bedeutende Vorhersagekraft für die weitere Entwicklung der Therapierten für den Zeitraum nach der Therapie hat.

Kasten 2.2: Klinisch bedeutsame Veränderungen

nach Th. vor Th.	gesund	subklinisch	krank
gesund	A	F	G
subklinisch	B	E	H
krank	C	D	I

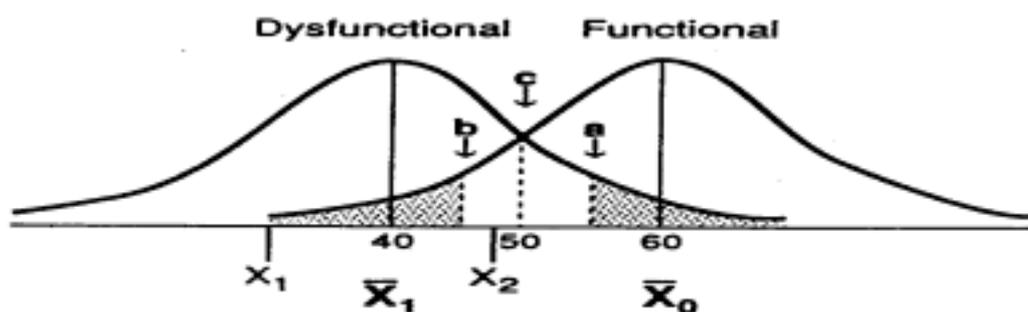
Beispiel für eine Analyse klinisch bedeutsamer Veränderungen:

Zuerst werden cut-off-Werte für die Kategorien „gesund“, „subklinisch“ (Beeinträchtigung und Belastung deutlich vorhanden, aber unterhalb klinisch relevanter behandlungsbedürftiger Störung) und „krank“ (klinisch relevante Störung mit Krankheitswert im sozialrechtlichen Sinne) festgelegt. Diese Kategorien müssen sorgfältig operationalisiert werden.

Dann werden die Patienten entsprechend ihrem Status vor und nach Therapie eingeordnet. Die Überschreitung einer Klassengrenze nach Therapie (alle Kästchen außer A,E,I) sei als klinisch relevante Veränderung bezeichnet. Die Ergebnisse einer Intervention sind zufriedenstellend, wenn sich der überwiegende Teil der Patienten in den Kästchen A bis D befindet. Bedenklich sind Non-Responder (Kästchen E und I) und besonders auch Verschlechterungen oder gar Therapieschäden (Kästchen F bis H).

Um zu entscheiden, ob eine klinisch signifikante Änderung vorliegt, gibt es prinzipiell drei Möglichkeiten: liegt der Post-Wert außerhalb der dysfunktionalen Population, innerhalb der gesunden Population oder mit größerer Wahrscheinlichkeit eher in der gesunden Population. In Abb. 2.3 sind die entsprechenden cut-off-Werte mit a, b und c bezeichnet.

Abbildung 2.3: Vorher- und Nachher-Werte eines Individuums (x) und die drei Cut-off-Punkte für klinisch signifikante Veränderung bezogen auf dysfunktionale und funktionale Gruppen



Welcher dieser drei cut-off-Punkte a, b oder c nun angemessen ist, hängt davon ab, wie konservativ man vorgehen möchte und wie stark sich die beiden Verteilungen überlappen. Überlappen sie sich stark, so ist Kriterium a (außerhalb der klinischen Stichprobe) das konservativste, Kriterium b das weichste. Überlappen sie sich nicht, dann ist Kriterium b (innerhalb der funktionalen Population) strenger.

Die Punkte a und b werden folgendermaßen bestimmt:

$$a = \bar{X}_1 + Ks_1$$

$$b = \bar{X}_0 - Ks_0$$

X_0 und s_0 entsprechen dabei Mittelwert und Standardabweichung der normalen Population. X_1 und s_1 entsprechen dabei Mittelwert und Standardabweichung der dysfunktionalen Population. K ist eine Konstante, die über das Ausmaß der geforderten Differenz entscheidet. Eine mögliche sinnvolle Wahl für diese Konstante ist $K=1.65$, da damit mit 95% iger Wahrscheinlichkeit abgesichert wäre, dass ein Wert jenseits der damit berechneten Grenze auch wirklich nicht mehr in die entsprechende Population gehört.

Kriterium c nimmt in beiden Fällen eine mittlere Position ein und würde sicherstellen, dass der Meßwert eher zur funktionalen als zur dysfunktionalen Population gehört. Berechnet wird dieser Punkt über die folgende Formel:

$$c = \frac{s_0\bar{X}_1 + s_1\bar{X}_0}{s_0 + s_1}$$

X_0 und s_0 entsprechen dabei Mittelwert und Standardabweichung der normalen Population. X_1 und s_1 entsprechen dabei Mittelwert und Standardabweichung der dysfunktionalen Population. Dieser Punkt ist am wenigsten willkürlich, da keine Entscheidung über das Kriterium (innerhalb funktionaler oder außerhalb dysfunktionaler Norm?) oder den Abstand zum entsprechenden Mittelwert (Konstante K) getroffen werden braucht.

Leider liegen oft keine Normwerte für eine gesunde Population vor. In diesem Fall kann man nur Punkt a verwenden. Werte einer dysfunktionalen Stichprobe hat man durch die Prä-Werte immer. Dabei ist allerdings eine Einschränkung zu beachten: Wenn jede Studie eigene dysfunktionale Normen benutzt, so werden die berechneten cut-off-Punkte a und b zwischen den Studien unterschiedlich sein (Jacobson & Truax, 1991; Hollon & Flick, 1988; Wampold & Jensen, 1986). Bei schwer dysfunktionalen Stichproben wäre es dann einfacher, klinisch relevante Veränderungen nachzuweisen.

Ideal wäre daher das Vorliegen von allgemeinen Normwerten für dysfunktionale Stichproben. Diese können gewonnen werden, indem die Mittelwerte und Standardabweichungen über mehrere Studien hinweg gemittelt werden

(Jacobson et al., 1988). Für viele Skalen und Populationen liegen aber noch keine Normen vor, so dass nur ein Rückgriff auf die eigenen Prä-Werte bleibt.

Auf die Operationalisierung der Kategorie „subklinisch“, für die in der Literatur noch keine Empfehlungen vorliegen, wird im entsprechenden Methodenabschnitt 3.2.3 eingegangen.

2.2.2.3 *Stabilität der Effekte über die Zeit*

Neben der Berücksichtigung von Ausmaß und klinischer Signifikanz der Veränderungen ist hinsichtlich der Therapie-Erfolgsmessung zu fordern, dass die vorgefundenen Effekte auf ihre *Zeitstabilität* getestet werden (vgl. Nicholson & Berman, 1983). So kann die Häufigkeit von Rückfällen abgeschätzt oder die Hypothese der „Symptomverschiebung“ nach einer störungsspezifischen Therapie geprüft werden. In diesem Zusammenhang ist auch die Frage weiter zu klären, ob die Wahrscheinlichkeit einer psychischen Störung für Personen, die bereits einmal relevante psychische Probleme aufwiesen, auch in der Zukunft grundsätzlich im Sinne einer allgemeinen Vulnerabilität erhöht ist (unabhängig davon, ob die erstaufgetretene Störung spontan bzw. durch Selbsthilfe oder durch eine beliebige Therapieform und deren kurzfristigem Behandlungserfolg beeinflusst wurde).

Bisher weisen die Ergebnisse aus Katamnesen darauf hin, dass sich ein Teil der Patienten nach der Therapie weiter verbessert (vielleicht aus Gründen von durch die Therapie erhöhter Selbstwirksamkeit), ein Teil die Therapieerfolge aufrechterhält und ein eher geringer Teil erneut Beschwerden entwickelt bzw. bereits während der Therapie schon nur wenig Fortschritte machte. Die psychoanalytische „Symptomverschiebungshypothese“ für Verhaltenstherapie kann damit in ihrer Allgemeinheit als klar widerlegt gelten. Darüber hinaus gibt es aber aus forschungspraktischen und forschungsmethodischen Gründen noch wenig gesichertes Wissen über Langzeit-Wirkungen von Psychotherapie. Auch auf dem Feld der Erhebung, Auswertung und Interpretation langer Katamnesen müsste noch Wissen erarbeitet werden, um Therapieeffekte von den Auswirkungen anderer Einflussfaktoren zu trennen.

Die Frage der Zeitstabilität von Psychotherapieerfolg ist also aufgrund der Seltenheit veröffentlichter Langzeit-Katamnesen noch längst nicht umfassend fundiert beantwortet. Es gibt aber durchaus Hinweise für eine optimistische Einschätzung, wie etwa die Daten der Berner 12-Jahres-Katamnese zeigen (Grawe, 1999).

Gerade im Zusammenhang mit Kostenstudien ist die Dauerhaftigkeit eines Effektes von entscheidender Bedeutung. Die Nachuntersuchungen sollten also um Katamnesen ergänzt werden. Der im Rahmen dieses Projektes mögliche Katamnesezeitraum bezieht sich auf ein Jahr.

In Kasten 2.3 sind die wichtigsten Aspekte der Effektivitätsmessung bei Kosten-Effektivitäts-Studien noch einmal zusammengefasst.

Kasten 2.3: Zusammenfassung: Messung der Effektivität bei Kosten-Effektivitäts-Analysen

Folgendes lässt sich zusammenfassend für die Effektivitätsmessung im Rahmen einer Kosten-Effektivitäts-Analyse festhalten:

- Die Effektivität sollte auf mehreren Ebenen, multimodal und zu mehreren Zeitpunkten (prä, post, Katamnese) erhoben werden.
- Geeigneter als reine Angaben zur statistischen Signifikanz von Veränderungen sind Angaben auf Patientenebene (Anteil gebesserter Patienten, klinische Signifikanz) oder zum Ausmaß der Veränderungen (mit geeigneten Effektstärkemaßen; hier unter Einbezug sowohl der Prä- als auch der Post-Varianz)
- Zur Bildung eines Kosten-Effektivitäts-Quotienten (Therapiekosten/Anteil erfolgreich behandelter Patienten) müssen nachvollziehbare Kriterien benutzt werden, um klinisch relevante Behandlungserfolge zu definieren.

Dabei sollte die Kosten-Effektivität für verschiedene Effektivitäts-Bestimmungen getrennt ermittelt und vergleichend diskutiert werden.

2.3 Kosten-Nutzen-Analysen

2.3.1 Richtlinien zur Durchführung von Kosten-Nutzen-Analysen

Bei Kosten-Nutzen-Analysen im Gesundheitswesen werden die Behandlungskosten dem monetär bezifferbaren Behandlungsnutzen gegenübergestellt. Dabei gibt es aber eine Vielzahl an Perspektiven und Berechnungsarten. Um aufgrund der Analyseergebnisse Vergleiche mit anderen Verfahren zu ermöglichen, bedürfen Kostenanalysen einer gewissen Standardisierung (Weismann, 1997). Für eine gute ökonomische Evaluation haben Drummond et al. (1987) zehn Richtlinien formuliert, die im Folgenden kurz erläutert werden. Die Autoren demonstrieren, wie man diese Leitlinien als „Checkliste“ für eigene oder zur Beurteilung anderer Kosten-Nutzen-Analysen einsetzen kann. Diese Checkliste wird auch in dieser Arbeit Verwendung finden.

- *Aufstellung einer genau definierten Fragestellung in einer beantwortbaren Form*

Dieser Punkt betrifft in Kostenstudien vor allem die klare Herausarbeitung verschiedener Behandlungsalternativen (falls vorhanden, ansonsten Angaben zum Verlauf bei Nicht-Behandlung) und zur Perspektive, die bezüglich Kosten und Nutzen eingenommen wird.

- *Umfassende Darstellung und Beschreibung möglicher zur Verfügung stehender Alternativen*

Wie in jeder sauberen wissenschaftlichen Studie sollten umfassende Angaben gemacht werden, die eine Replikation ermöglichen; auch sollte zur Übertragbarkeit eines Programms auf andere Settings Stellung bezogen werden.

- *Bestimmung der Wirksamkeit der betrachteten Programme und Interventionen*

Die Effizienz uneffektiver Maßnahmen interessiert nicht! Die Effektivität eines Programms soll also vorher oder gleichzeitig mit der Kostenstudie ermittelt werden.

- *Ermittlung aller relevanten Kosten- und Nutzenfaktoren für jede betrachtete Alternative*

Es soll Transparenz darüber geschaffen werden, welche Faktoren genau berücksichtigt wurden (auch wenn sie letztendlich, z.B. aus Gründen der Unsicherheit über Preise oder die Berechnung nicht in die Kosten-Analyse eingeflossen sind).

- *Messung von Kosten und Nutzen in geeigneten Einheiten*

Bei Kosten-Nutzen-Analysen sollen sowohl Kosten wie Nutzen in Geldwerten angegeben werden. Dabei muß deutlich werden, wie man zur Ermittlung der Kosten gelangt ist (Angabe von Preisen, Klarheit darüber, wann Durchschnittswerte benutzt wurden etc.).

- *Kritik der ermittelten Kosten und Nutzen hinsichtlich ihrer glaubwürdigen Verwendbarkeit*

Kosten und Nutzen sollen umfassend diskutiert werden, z.B. bezüglich der momentanen Marktsituation (so schwanken zum Beispiel die Beträge, die der Autor der vorliegenden Arbeit derzeit für eine psychotherapeutische Behandlungseinheit in Rechnung stellen kann, je nach Krankenkasse zwischen DM 49.- und DM 186.-). Auch die veranschlagte Dauer des Nutzens, die Nutzenwerte für Arbeitsunfähigkeit bei Nicht-Erwerbstätigen etc. müssen umfassend begründet werden.

- *Anpassung/Bereinigung der Kosten und Nutzen auf einen Zeitpunkt, wenn sie zu unterschiedlichen Zeiten anfallen (Diskontierung)*

Unter Diskontierung verstehen Wirtschaftler ein Verfahren, mit dessen Hilfe die zeitlichen Unterschiede der Entstehung von Kosten und Nutzen ausgeglichen werden sollen. Diesem Ansatz liegt folgende Überlegung zugrunde: die Kosten eines Verfahrens entstehen immer in der Gegenwart, der Nutzen setzt aber erst in der Zukunft nach Beendigung der Therapie ein. In der Zwischenzeit hätte sich aber der Geldwert, der für die Therapie aufgewendet wurde, erhöht, da er als arbeitendes Kapital zur Verfügung gestanden hätte. Werden z.B. 1990 für eine Therapie 5000,- DM ausgegeben, so hätte dieses Geld (bei einer angenommenen Verzinsung von 3%) 1991 bereits einen Wert von 5150,- DM. Noch schwieriger wird es, wenn die Therapie über mehrere Jahre verläuft. Da das Geld nicht auf einmal für die gesamte Therapie aufgebracht wird, entstehen jedes Jahr (zumeist noch unterschiedlich hohe) Kosten. Diese müs-

sen alle auf ein Bezugsjahr diskontiert werden, bevor sie dem Nutzen gegenübergestellt werden.

Die Berechnung des diskontierten Nutzens (N_x) erfolgte auf der Grundlage des Nutzens nach einem Jahr (N) für bis zu 5 Jahren nach der Therapie. Die Diskontierung erfolgte mit folgender Formel:

$$N_x = N + \frac{N}{(1+r)} + \dots + \frac{N}{(1+r)^{x-1}}$$

Die Diskontierungsrate (r) wird dabei meist mit 5% festgesetzt (Konsensgruppe Gesundheitsökonomie, 1996).

- *Ermittlung der zusätzlichen Kosten und Nutzen eines Programmes (Marginalanalyse)*

Die Marginal- oder Incremental-Analyse hat die Aufgabe, nach Aufstellung der größten Kosten- und Nutzenfaktoren noch nach weiteren Posten zu suchen, die ebenfalls Ressourcen verbrauchen. So erhält man nach und nach eine umfassende Aufstellung aller kostenverursachenden Faktoren und erhöht damit die Genauigkeit der Analyse.

- *Durchführung einer Sensitivitätsanalyse*

Die Sensitivitätsanalyse prüft die Ergebnisse der Kosten-Nutzen-Analyse auf deren Robustheit. Es ist nur allzu leicht denkbar, dass bei der Bestimmung der Kosten Fehler unterlaufen. So können z.B. falsche Durchschnittswerte in die Berechnung eingegangen oder ersatzweise Zahlen verwendet worden sein, die es zum Erhebungszeitraum noch nicht in genauerer Form gab. Kritiker bemängeln dies dann häufig zu Recht. Mit der Sensitivitätsanalyse simuliert man eine gewisse Schwankung der Größen und schaut dabei, ob sich die Aussage in drastischer Weise verändert. Ist dies der Fall, können die Ergebnisse gleich mit besonderen Einschränkungen veröffentlicht werden (z.B. die Therapie lohnt sich aber nur dann, wenn ihre Kosten einen Stundensatz von X nicht überschreiten). Drummond und Mitarbeiter heben hervor, dass in jeder soliden Kostenstudie eine Sensitivitätsanalyse durchgeführt werden muß.

- *Diskussion und Interpretation der Ergebnisse*

Dieser Punkt betrifft Fragen wie: Sind alle methodischen und inhaltlichen Entscheidungen, die in der Studie getroffen wurden, nachvollziehbar dargestellt? Werden Hinweise gegeben, wie die Ergebnisse in andere Settings übertragen werden können? Kann der Leser der Studie entscheiden, welchen der Kosten-Nutzen- oder Kosten-Effektivitäts-Indices er für eine bestimmte Fragestellung für angemessen hält?

Im deutschsprachigen Raum sind von der Konsensgruppe Gesundheitsökonomie (1996) von der Medizinischen Hochschule Hannover in Zusammenarbeit mit der Deutschen Gesellschaft für Klinische Pharmakologie und Therapie ebenfalls Empfehlungen zur Durchführung ökonomischer Evaluation herausgegeben worden. Ihre Forderungen für qualitativ gute Studien decken sich mit den oben dargestellten Richtlinien. Zusätzlich fordern die Wissenschaftler auf, Ergebnisse derartiger Studien zur Diskussion zu stellen und in einschlägigen Fachzeitschriften zu publizieren.

2.3.2 Forschungsstand zu Kosten-Analysen im Psychotherapiebereich (insbes. Angststörungen)

Im folgenden werden umfassend Studien aus dem Bereich der Psychotherapieforschung (insbesondere zu Angststörungen) referiert, bei denen auch Kostenaspekte analysiert wurden. Neben Überblicksarbeiten sind dies Studien zu Krankheitskosten, zur Kosten-Effektivität und zu Kosten-Nutzen-Analysen.

2.3.2.1 Überblicksarbeiten

Einführend zu diesem Abschnitt sollen hier Überblicksarbeiten zu Kostenstudien im Psychotherapiebereich genannt werden, die es dem interessierten Leser ermöglichen, sich in kurzer Zeit einen Überblick zur Thematik zu verschaffen.

Bühringer und Hahlweg (1986) stellen in ihrem Überblicksartikel nicht nur einschlägige Studien vor, sondern geben außerdem auch einen sehr guten Einblick in die verschiedenen Analysetypen. Der Artikel von Candilis et al. (1997) ist zwar kurz, aber auch gut als Einstieg geeignet, da hier v.a. sehr aktuelle Studien zitiert werden. Auch in der Arbeit von Davidson (1996) lassen sich relativ neue Studien zum Thema Lebensqualität und Kostenfaktoren bei Patienten mit Panikstörung finden.

Ein etwas älterer Aufsatz stammt von Berufsverband Deutscher Psychologen (1988). Unter dem Titel „Wirksamkeit und Nutzen psychotherapeutischer Behandlung“ wird hier vor allem die Notwendigkeit von Kostenstudien auf dem Psychotherapiesektor erläutert und weiterhin eine Reihe von Studien dargestellt, die die hohe Belastung des Gesundheitssystems durch unbehandelte Patienten mit psychischen Störungen und die Reduktion dieser Kosten durch Psychotherapie belegen.

Ein Überblicksartikel von Hofmann und Barlow (im Druck) war bereits als Vorabdruck verfügbar. Die Autoren zitieren alle wichtigen Kostenstudien über Angststörungen der letzten Jahre. Schlesinger und Mumford (1980) führten bereits vor fast 20 Jahren eine Meta-Analyse von 11 Studien zu verschiedenen psychischen Störungen durch und aktualisierten diese nochmals 4 Jahre später mit nunmehr 58 Studien (Mumford et al., 1984). Bei ihnen lassen sich in einer sehr gut aufbereiteten Übersicht die wichtigsten Informationen zu den jeweiligen Untersuchungen finden. Die umfangreichste Meta-Analyse im deutschsprachigen Raum stammt von Baltensberger (1996). Sie hat in ihrer Dissertation insgesamt 124 Studien sekundäranalytisch ausgewertet, die sich mit materiellen und immateriellen Nutzenaspekten von Psychotherapien verschiedener Störungen befassen. Die genannten Überblicksarbeiten sind in Tabelle 2.1 zusammengefasst.

Tabelle 2.1: Überblicksarbeiten zum Thema „Kostenanalysen von Psychotherapie“

Autoren	Art der Studie	Anzahl zitierter Kostenstudien
Bühringer & Hahlweg (1986)	Überblicksartikel	14
Candilis & Pollack (1997)	Überblicksartikel	18
Davidson (1996)	Überblicksartikel	12
Berufsverband Deutscher Psychologen (1988)	Überblicksartikel	33
Hofmann & Barlow (im Druck)	Überblicksartikel	20
Schlesinger & Mumford (1980)	Meta-Analyse	11
Mumford et al. (1984)	Meta-Analyse	58
Baltensberger (1996)	Meta-Analyse	124

2.3.2.2 Störungskosten

Zum Inanspruchnahmeverhalten bei medizinischen Leistungen gibt es eine Reihe von Untersuchungen aus dem angloamerikanischen und zunehmend auch aus dem deutschen Raum. Die folgende Darstellung enthält Studien, die danach ausgewählt wurden, ob die untersuchten Patienten irgendeine Art von

Angststörungen aufwiesen. Untersuchungen, die sich ausdrücklich nicht mit Angsterkrankungen befassten (z.B. zu Schizophrenie, Depression oder Substanzmissbrauch) wurden aus der Betrachtung ausgeschlossen und Studien, bei denen Angststörungen wahrscheinlich vertreten waren (z.B. „neurotische Patienten“) eingeschlossen.

Cummings (1977) schätzte, dass ca. 60% aller Arztbesuche auf Patienten entfallen, die neben somatischen Beschwerden auch behandlungsbedürftige emotionale Störungen aufweisen. Reimer et al. (1979) ermittelten eine durchschnittliche Dauer inadäquater Vorbehandlung bis zur Einweisung in eine Fachklinik zwischen 5,5 und 9 Jahren je nach Störung. 92% der Patienten hatten in dieser Zeit mindestens 3 Ärzte konsultiert, 15% mehr als 10 Ärzte, 70% waren ausschließlich somatisch behandelt worden. Obwohl bei dieser Studie die Kosten der medizinischen Fehlbehandlung nicht erhoben wurden, kann man sich leicht vorstellen, dass es sich hier um erhebliche Summen handelt, die an falscher Stelle ausgegeben werden.

Andere Ergebnisse deuten ebenfalls auf inadäquate Behandlung psychischer Störungen hin. Shapiro et al. (1984) stellten fest, dass in den 6 Monaten vor Aufnahme in ihre Studie 70% der Patienten mit seelischen Problemen ambulante Arztpraxen aufgesucht hatten, jedoch nur 18% einen Psychologen oder Psychiater. Sie fanden bei Patienten mit Ängsten und psychosomatischen Störungen einen Prozentsatz adäquater Behandlung von 20%. Ähnlich die Ergebnisse bei Butollo und Höfling (1984): Neurologen und Psychiater wurden von Angstpatienten nur in 9% der Fälle aufgesucht, häufig jedoch praktische Ärzte (51%). 14% der Patienten hatten mindestens einen Klinikaufenthalt aufzuweisen. Ein Viertel hatte bereits eine mehrjährige Psychoanalyse hinter sich, keiner der Patienten dagegen eine verhaltenstherapeutische Behandlung. Simon et al. (1995) kommen zu dem Schluss, dass die Gesamtkosten für das Gesundheitssystem bei Angstpatienten und Depressiven etwa doppelt so hoch liegen wie bei Nichterkrankten. Dabei machen die eigentlichen Psychotherapiekosten nur 10% aus, der Rest entsteht durch ambulante und stationäre medizinische Versorgung.

Das sogenannte „doctor-shopping“ und die z.T. extrem langen Krankheitsverläufe führen außerdem zu einer iatrogenen Chronifizierung. Darunter versteht man die Aufrechterhaltung einer Erkrankung, die durch therapeutische Maßnahmen, gleichsam wie bei einem Teufelskreis hervorgerufen wird (Zielke, 1993).

Weiterhin muß berücksichtigt werden, dass diese Patienten das medizinische System für die somatisch zu behandelnden Patienten blockieren und damit deren Versorgungslage verschlechtern (Bühringer & Hahlweg, 1986). Ringel (1983) und Zielke und Mark (1989) betonen, dass die starke Ausrichtung der Medizin auf den somatischen Bereich eine Unterversorgung von Patienten mit psychischer und psychosomatischer Symptomatik hervorruft.

Markowitz et al. (1989) führten eine Untersuchung mit Kontrollgruppen durch. Sie fanden eine auffallend hohe Prävalenz von 43% der Panikpatienten, die im letzten Jahr medizinische und psychologische Dienste in Anspruch nahmen – im Gegensatz zu 4% bei den Nichterkrankten. 28% der Paniker riefen im genannten Zeitraum den Notarzt – im Vergleich zu 2% in der Kontrollgruppe. Auch bei den Medikamenten gab es deutliche Unterschiede. Im letzten halben Jahr erhielten 42% der Patienten, aber nur 8% der Gesunden leichte Beruhigungsmittel. Antidepressiva bekamen 8% der Erkrankten und nur 1% der Nichterkrankten. Die Autoren beschreiben noch einen weiteren Kostenfaktor, der nicht mit der Inanspruchnahme medizinischer Leistungen zu tun hat. 27% der Personen mit Paniksyndrom lebten von Sozialhilfemaßnahmen – im Gegensatz zu 12% der Gesunden.

Auch Edlund & Swann (1987) untersuchten die Arbeitsfähigkeit bei Panikpatienten. Sie fanden bei 83% der Personen eine verringerte Arbeitsqualität, 66% verloren gar ihre Arbeit ganz. 43% waren mindestens einen Monat lang arbeitsunfähig. Die Kosten für einen Panikpatienten beliefen sich demnach im Durchschnitt auf 44 413 Dollar pro Patient. Diese Auswirkungen der Krankheit auf die Arbeitsfähigkeit mögen sehr drastisch erscheinen. Es ist aber nicht klar, ob es sich hier tatsächlich immer um eine Folgeerscheinung der Erkrankung handelt oder ob nicht vielleicht die Arbeitslosigkeit oder eine schlechte soziale Lage erst zu den Symptomen geführt hat. Hierzu bedarf es weiterer Untersuchungen.

Klerman et al. (1991) berichten, dass Patienten mit Panikanfällen nicht nur das Gesundheitssystem stärker in Anspruch nehmen, sondern häufig auch noch zusätzliche Kosten durch komorbide Erkrankungen wie depressive Syndrome oder Substanzmißbrauch verursachen. In die gleiche Richtung weisen auch Befunde von Leon et al. (1995). 20% der Männer und 30% der Frauen wiesen komorbide Erkrankungen auf. Sie dokumentieren neben der verstärkten Inanspruchnahme von medizinischen Diensten und Notfallhilfen auch noch weitere Kostenfaktoren bei Personen mit Angststörungen, wie erhöhte finanzielle Abhängigkeit, Arbeitslosigkeit und Substanzmißbrauch. Du Pont et al. (1996) machen darauf aufmerksam, dass Angststörungen zwar hohe Kosten im Gesundheitssystem verursachen (31,5% der Kosten aller psychischen Erkrankungen überhaupt), dies aber insgesamt nur ein Viertel aller durch die Krankheit entstehenden Aufwendungen ausmacht. Der größte Teil, nämlich fast drei Viertel der Gesamtkosten, werde durch verlorene Produktivität aufgrund von Arbeitsunfähigkeit verursacht.

In Tabelle 2.2 sind die aufgeführten Studien noch einmal überblicksartig dargestellt. Da die Studien bezüglich Methoden und Ergebnisdarstellungen sehr heterogen sind, sollen zuvor in Kasten 2.4 die Befunde zu den Störungskosten noch einmal zusammengefaßt werden.

In den folgenden Abschnitten werden die Befunde zu den Auswirkungen von Behandlungen auf die Störungskosten bzw. zunächst auf die störungsbezogenen Kostenfaktoren referiert.

Kasten 2.4: Zusammenfassung der Befunde zu den Störungskosten

- Patienten mit psychischen Störungen nehmen das medizinische System sehr viel häufiger in Anspruch als Personen ohne derartige Beeinträchtigungen.
- Sie erhalten dabei selten eine adäquate Behandlung bei einem Psychiater oder Psychologen. Viele Patienten wechseln zudem aus Unzufriedenheit des öfteren ihren Arzt, wodurch Mehrfach- und Wiederholungsuntersuchungen entstehen, die das Gesundheitssystem finanziell unnötig belasten.
- Neben der unmittelbaren finanziellen Belastung für das Gesundheitssystem zu Krankheitsbeginn kann das „doctor-shopping“ zu einer iatrogenen Chronifizierung und damit zu zusätzlichen Folgekosten führen.
- Zudem blockieren diese Patienten das medizinische System für die somatisch zu behandelnden Patienten und verschlechtern damit die allgemeine Versorgungslage, was indirekt möglicherweise an anderen Orten wiederum unnötige Kosten hervorrufen könnte.
- Besonders sind die Kosten durch verminderte Arbeitsfähigkeit hervorzuheben (etwa 75% der Gesamtkosten), und zwar sowohl was Erwerbsunfähigkeit bzw. Arbeitslosigkeit betrifft, als auch Arbeitsunfähigkeit bzw. Krankheitstage (oder verminderte Arbeitsqualität) bei bestehendem Arbeitsverhältnis.

Tabelle 2.2: Übersicht ausgewählter Studien (chronologisch) zu Kosten psychischer Erkrankungen (insbesondere Angststörungen)

Autoren	Diagnose	Datenbasis	Kostenfaktoren	Ergebnisse
Reimer et al. (1979)	chronische psychogene Erkrankungen	N = 100 retrospektive Befragung von Patienten einer psychosomatischen Fachklinik	medizinische Behandlungen, Medikamente	<ul style="list-style-type: none"> • 92% konsultierten mind. 3 Ärzte • 15% mehr als 10 Ärzte • 70% lediglich somatisch behandelt • nichtkausale medizinische Behandlungen tragen zur Chronifizierung bei
Butollo & Höfling (1984)	chronische Ängste	N = 43 Patienten, die sich aufgrund eines Zeitschriftenartikels zur Verhaltenstherapie anmeldeten	psychologische und ärztliche ambulante Behandlungen, stationäre Aufenthalte, Arbeitsunfähigkeit	<ul style="list-style-type: none"> • 51% konsultierten praktische Ärzte • aber nur 9% waren beim Neurologen/Psychiater • 14% hatten Klinikaufenthalte hinter sich • 25% absolvierten mehrjährige Psychoanalyse • kein Patient hatte Verhaltenstherapie bekommen • 50% waren arbeitsunfähig
Shapiro et al. (1984)	verschiedene, u.a. Ängste und psychosomatische Störungen	retrospektive Erhebung über die letzten 6 Monate	psychologische und ärztliche ambulante Behandlungen	<ul style="list-style-type: none"> • 70% hatten ambulante Arztkontakte • aber nur 18% waren beim Psychologen/Psychiater • 20% der psychosomatischen und der Angstpatienten wurden adäquat behandelt • deutliche therapeutische Unterversorgung von Patienten mit seelischen Störungen
Edlund & Swann (1987)	Panik	N = 30 behandlungssuchende Patienten	Arbeitsunfähigkeit	<ul style="list-style-type: none"> • 83% verringerte Arbeitsqualität • 66% arbeitslos • 43% mind. 1 Monat lang arbeitsunfähig

Forts. Tabelle 2.2: Übersicht ausgewählter Studien (chronologisch) zu Kosten psychischer Erkrankungen (insbesondere Angststörungen)

Autoren	Diagnose	Datenbasis	Kostenfaktoren	Ergebnisse
Markowitz et al. (1989)	Panik	Analyse der ECA-Daten (Studie des Epidemiologic Catchment Area zu psychischen Erkrankungen)	psychologische und medizinische Behandlungen, Notarztkontakte, Medikamente staatliche Zuwendungen	<ul style="list-style-type: none"> • 43% nutzten psychol. Und medizin. Versorgung (im Vgl. zu 4% der Nichterkrankten) • 28% riefen den Notarzt (2% der Nichterkrankten) • 42% erhielten Beruhigungsmittel (8% der Gesunden) • 27% bekamen Sozialhilfe (12% der Gesunden) • dtl. höhere Gesundheitskosten als psychisch Gesunde
Leon et al. (1995)	Angststörungen	Analyse der ECA-Daten	psycholog. und medizin. Behandlungen, Notarztkontakte, staatliche Zuwendungen, Arbeitsunfähigkeit	<ul style="list-style-type: none"> • 60% der männl. Paniker z.Zt. arbeitslos • 31% der weibl. Paniker bekommen Sozialhilfe • 36% der männl. Paniker suchten mind. 1 Spezialisten auf • Angstpatienten verursachen sehr hohe soziale Kosten
Simon et al. (1995)	Depressionen und Angststörungen	N = 303 Patienten der medizinischen Basisversorgung, Vergleich mit Nichterkrankten	Inanspruchnahme der medizinischen Grundversorgung	<ul style="list-style-type: none"> • Kosten für Gesundheitssystem bei psychisch Kranken doppelt so hoch • 90% der Kosten für ambulante und stationäre medizinische Versorgung • nur 10% der Kosten für psychologische Behandlungen
DuPont et al. (1996)	Angststörungen	Analyse der ECA-Daten	Nutzung des medizinischen Systems, Arbeitsunfähigkeit	<ul style="list-style-type: none"> • Angststörungen verursachen 31,5% der Kosten aller psych. Erkrankungen im Gesundheitssystem • 75% der krankheitsbezogenen Kosten entstehen durch Arbeitsunfähigkeit

2.3.2.3 Kosten-Effektivität im Sinne der Verringerung von Kostenfaktoren durch psychotherapeutische Behandlungen

Anhand dieser Ergebnisse wird ersichtlich, dass Patienten mit psychischen Störungen im Allgemeinen und Angstpatienten im Besonderen hohe volkswirtschaftliche Kosten verursachen. Es gibt auch bereits eine Anzahl von Untersuchungen, die belegen, dass diese hohen Ausgaben durch Psychotherapie gemindert werden können. Viele dieser Studien bleiben dabei allerdings noch auf der Ebene der Effektivität bzw. stellen lediglich eine Verringerung der *Kostenfaktoren* vom Beginn bis zu einem Zeitpunkt nach der Therapie dar, ohne diese in Geldeinheiten umzurechnen.

In den sogenannten Kaiser-Studien von Follette und Cummings (1967) wurden erstmals umfassend die Auswirkungen psychologischer Behandlung auf die Nutzung medizinischer Einrichtungen untersucht. Es zeigten sich dabei folgende Ergebnisse: die Krankenhaustage konnten um 52%, die Anzahl diagnostischer Maßnahmen um 71% verringert werden. Cummings (1977) schloss daraus, dass es keine empirische Basis für die Befürchtung gäbe, dass die steigende Nachfrage nach psychologischer Behandlung das Gesundheitssystem finanziell gefährden werde. Rosen und Wiens (1979) versuchten, genannte Ergebnisse zu replizieren. Sie fanden einen Rückgang bei den Krankenhaustagen um 35%, bei den Ambulanzbesuchen um 41%, und die Anzahl diagnostischer Maßnahmen verringerte sich um 40%. Auch Goldberg et al. (1970) kamen zuvor schon zu ähnlichen Zahlen. Sie ermittelten eine Reduktion der Ambulanzbesuche um 31% und der diagnostischen Maßnahmen um 30%. Bei diesen drei Untersuchungen wird allerdings immer nur von psychologischer oder psychologisch-psychiatrischer Behandlung gesprochen, ohne dass diese näher spezifiziert wird. Es ist also nicht ganz klar, welche Therapie genau die Patienten erhalten haben, die zu einer solchen Reduktion der Kosten geführt hat.

In einer umfangreichen Meta-Analyse stellen Mumford et al. (1984) die Ergebnisse von 58 kontrollierten Studien zusammen. In 26 Fällen lagen Prä- und Post-Daten vor. Für diese Studien konnte eine durchschnittliche Verringerung der Inanspruchnahme medizinischer Dienste um 33,1% ermittelt werden. 85% aller einbezogenen Studien fanden einen Einsparungseffekt durch die Psychotherapie. Die Einheiten der Erfolgsmessung waren dabei unterschiedlich (z.B. Anzahl der Arztbesuche, verlorene Arbeitstage etc.).

Insgesamt reichten die Befunde von 181,6% Verringerung bis zu 72,4% Erhöhung der medizinischen Leistungen. Gab es eine Kontrollgruppe, so waren deren Werte immer schlechter als die der Versuchsgruppe. Dies ist vor allem als Argument wichtig bei Verfahren, bei denen sich am Ende eine Erhöhung der Kosten ergab, denn im Falle einer Nichtbehandlung würden die Kosten demnach noch stärker ansteigen. Die Autoren replizierten mit der Untersuchung die von Schlesinger et al. (1980) in einer Meta-Analyse von 11 Studien gefundene durchschnittliche Reduktion um 25%.

Ginsberg und Marks (1977) untersuchten die Auswirkungen von Verhaltenstherapie bei „Neurotikern“ und fanden eine Verringerung der stationären Aufenthalte im psychiatrischen Krankenhaus um 86%, der Arztbesuche um 50% und der Krankschreibungen um 56%. Longobardi (1981) verglich „neurotische“ Patienten, die eine kombinierte Behandlung aus Verhaltenstherapie und rational-emotiver Therapie erhielten mit Patienten, die unter unspezifischen medizinischen Symptomen litten. Die Inanspruchnahme medizinischer Leistungen sank auch hier nach der Psychotherapie um 64% in der Versuchsgruppe. Bei der Kontrollgruppe hingegen wurde ein leichter Anstieg festgestellt.

In Tabelle 2.3 sind wiederum die oben dargestellten Studien zusammengefasst. Trotz Heterogenität der Methoden und Ergebnisaspekte kann durchgängig eine Verringerung der Kostenfaktoren nach einer Psychotherapie bzw. die Überlegenheit gegenüber einer Nicht-Behandlung verzeichnet werden.

Im folgenden Abschnitt sind die Studien aufgeführt, in denen diesen Sachverhalt auch monetär bewertet wurde.

Tabelle 2.3: Übersicht über ausgewählte Studien (chronologisch) zur Reduktion von Kostenfaktoren durch Psychotherapie bei psychischen Erkrankungen (insbesondere Angststörungen)

Autoren	Diagnose	Gruppen	Datenbasis	Kosten-Faktoren	Ergebnisse
Follette & Cummings (1967)	versch. psychische Störungen	diagn. Interview kurze PT lange PT ohne Bhdg.	insgesamt N = 304, Patienten des psychiatr.-psychol. Dienstes, Erhebung für 1 Jahr vor und 5 Jahre nach Therapiebeginn	Nutzung medizinischer Einrichtungen	<ul style="list-style-type: none"> • psychol. Behandelte beanspruchten seltener Leistungen nach PT als Unbehandelte • stationäre Tage verringerten sich um 52% • diagn. Maßnahmen sanken um 71%
Goldberg et al. (1970)	Neurosen, Psychosen	a) kurze PT b) ohne Bhdg.	insgesamt N = 256, Erhebung für 1 Jahr vor und 1 Jahr nach der Therapie	Nutzung medizinischer Einrichtungen	<ul style="list-style-type: none"> • 60% der Pat. sign. weniger Arztbesuche • Arztkontakte verringerten sich um 31% • diagn. Maßnahmen sanken um 30%
Ginsberg & Marks (1977)	Neurosen	Verhaltenstherapie	N = 42 , Erhbg. für 3 Monate vor und nach der Therapie, keine Kontrollgruppe	Arztkontakte, stationäre Aufenthalte, Arbeitsunfähigkeit	<ul style="list-style-type: none"> • 86% weniger Psychiatrie-Aufenthalte • 50% weniger Arztbesuche • 56% weniger Krankenschreibungen
Rosen & Wiens (1979)	versch. psychische Störungen	a) Diagnostik + PT b) nur Diagnostik c) ohne Bhdg.	insgesamt N = 468, alle Patienten einer medizin.-psychol. Ambulanz von '70 - '75	Nutzung medizinischer Einrichtungen	<ul style="list-style-type: none"> • a) und b) sign. weniger Leistungen als c) • stationäre Tage verringerten sich um 35% • Ambulanzbesuche sanken um 41% • diagn. Maßnahmen wurden um 40% weniger
Schlesinger et al. (1980)	versch. psychische Störungen	verschiedene Psychotherapien	Metaanalyse von 11 retrospektiven, kontrollierten Studien	Nutzung medizinischer Einrichtungen	<ul style="list-style-type: none"> • Patienten suchten medizin. Einrichtungen in den folgenden 2 Jahren dtl. seltener auf • Nutzung sank im Mittel um 25%
Longobardi (1981)	Neurosen	VT + RET ohne Bhdg.	insgesamt N = 34, Erhebung für 6 Monate vor & nach PT	Nutzung medizinischer Leistungen	<ul style="list-style-type: none"> • Nutzung medizinischer Leistungen sank bei a) um 64%, bei b) leichter Anstieg
Mumford et al. (1984)	versch. psychische Störungen	verschiedene Psychotherapien	Metaanalyse von 58 verschiedenen Studien	Nutzung medizinischer Einrichtungen, Arbeitsunfähigkeit	<ul style="list-style-type: none"> • in 85% der Studien Verringerung der Inanspruchnahme medizin. Leistungen gefunden • Behandelte immer besser als Unbehandelte

2.3.2.4 Kosten-Nutzen-Analysen

In neuerer Zeit gibt es zunehmend auch Kosten-Nutzen-Analysen, bei denen die Kostenfaktoren und deren Verringerung durch die Therapie in Geldeinheiten umgerechnet und den Kosten der Psychotherapie gegenübergestellt wurden. Sie bieten damit eine Basis für eine vergleichende ökonomische Evaluation.

Ginsberg et al. (1984) untersuchten an einer Stichprobe von 46 neurotischen Patienten die Kosteneinsparung, die eine verhaltenstherapeutische Behandlung durch geschulte Krankenpfleger einbringt. Schwestern und Pfleger sind gemeinhin billiger als Psychiater und Psychologen, und so errechneten die Autoren, dass sich die Therapie bei 40 Patienten bereits nach 2 Jahren amortisiert, weil die Kosten an anderer Stelle eingespart werden konnten.

In einer Studie von Gathmann (1990) wurden bei 110 ehemaligen Patienten einer psychosomatischen Abteilung (multimodale Behandlung, die sich keiner Therapieform eindeutig zuordnen lässt) im Rahmen einer Nachuntersuchung folgende Befunde erhoben: Pro Patient nahm die Anzahl der Arztbesuche von durchschnittlich 125 auf 40 im Untersuchungszeitraum ab. Es wurden damit Kosten in Höhe von 12 600,- öS (entspricht ca. 1800,- DM) gespart. Die Aufenthalte im Krankenhaus sanken von 38 Tagen je Patient auf 9 Tage, was einer Kosteneinsparung von 95 410,- öS (13 630,- DM) entspricht. Vor der Therapie waren die Patienten im Mittel 72 Tage arbeitsunfähig, im Zeitraum danach 26 Tage lang. Die vermiedenen Produktionsausfälle betragen demnach 87 750,- öS (12 536,- DM) pro Person. Die Autoren stellten den Nutzen insgesamt den Therapiekosten gegenüber und es ergab sich ein Nettonutzen von 77 897,- öS (11 128,- DM) je Patient.

Salvador-Carulla et al. (1995) verglichen in Spanien die krankheitsbezogenen Kosten von 61 Panikpatienten im Jahr vor und im Jahr während der ambulant durchgeführten Psychotherapie. Die Zahl der Arztbesuche sank dabei um 94%, die der stationären Aufenthalte fiel auf 0. Die Kosten für Laboruntersuchungen verringerten sich von 8 938 Dollar insgesamt auf 4 472 Dollar. Auf der anderen Seite stiegen im Jahr der Therapie natürlich die Ausgaben für psychologische Behandlungen (von 29 156 Dollar auf 46 222 Dollar über alle Patienten hinweg). Dafür sank die Zahl der Arbeitsunfähigkeitstage drastisch von 1050 auf 190, was einer Kosteneinsparung von 51 760 Dollar für die gesamte Gruppe entspricht. Insgesamt sanken die durchschnittlichen Kosten pro Patient um ca. 63% von 1 554 Dollar auf 985 Dollar.

Bei Jacobi & Margraf (in Druck; vgl. Neumer & Margraf, 1996) wird von einer Untersuchung berichtet, die sich mit der Auswirkung kognitiver Verhaltenstherapie bei Angstpatienten beschäftigt. Hier wurden unabhängig vom Therapieerfolg die medizinischen und psychotherapeutischen Gesundheitskosten erfasst. Die Gesamteinsparung belief sich auf rund 8000,- DM pro Patient, was einem Kosten-Nutzen-Verhältnis von 1:5,6 zugunsten der Psychotherapie entspricht.

Breyer (1997) führte eine retrospektive Befragung von tiefenpsychologisch behandelten Patienten durch. Er errechnete eine Gesamtersparnis pro Kopf von 8478,- DM bei Einzel- und 14 330,- DM bei Gruppenpatienten. Dabei machen die Krankenschreibungstage mit knapp 7000,- DM (Einzelpatienten) und gut 9000,- DM (Gruppenpatienten) jeweils den größten Teil aus. Das Kosten-Nutzen-Verhältnis betrug 1:0,25 (Einzeltherapie) bzw. 1:3,32 (Gruppen-therapie).

Die bisher umfangreichste Kostenstudie im deutschen Raum wurde von Zielke (1993) durchgeführt. Es wurden die Wirksamkeit und die Kosten von stationärer Verhaltenstherapie bei psychosomatischen Störungen untersucht. Das Projekt fand multizentrisch in drei Fachkliniken und in Zusammenarbeit mit den örtlichen Betriebskrankenkassen statt, so dass von einem sehr guten Datenmaterial gesprochen werden kann. Insgesamt ermittelten die Forscher ein Kosten-Nutzen-Verhältnis von 1:2,49. Es wurden dabei folgende Kostenfaktoren berücksichtigt: Praxiskontakte, stationäre Aufenthalte, Arbeitsunfähigkeitstage und Medikamente.

Ein Problem für die Vergleichbarkeit dieser Ergebnisse ist die unterschiedliche Berechnungsgrundlage in den Studien, z.B. was den Zeitraum betrifft, für den ein Kosten-Nutzen-Index erstellt wird. Dieser Sachverhalt unterstreicht die Forderung nach der Einhaltung von Leitlinien zur Durchführung und Darstellung von Kosten-Nutzen-Analysen.

Es kann dennoch zusammenfassend festgehalten werden, dass der finanzielle Nutzen in den meisten Studien die Therapiekosten bereits nach ein bis drei Jahren übersteigt. Dies war angesichts der im vergangenen Abschnitt dargestellten Reduktion von Kostenfaktoren durch Psychotherapie auch zu erwarten. Tabelle 2.4 fasst die eben dargestellten Studien noch einmal überblicksartig zusammen.

Es gibt also bereits gute Ansätze für den Einzug ökonomischer Evaluation in die wissenschaftliche Bewertung von Psychotherapie. Es fehlen aber noch Studien, die einen Vergleich zwischen den einzelnen Behandlungsalternativen möglich machen und die qualitativ überzeugen. Für die Zukunft bleibt also die Forderung nach mehr komplexen Beurteilungsverfahren unter Einbeziehung von Kostenaspekten und einer damit einhergehenden Erweiterung der Methodenausbildung und -erforschung auf diesem Gebiet (Bühringer & Hahlweg, 1986).

Tabelle 2.4.: Übersicht über verschiedene Kosten-Nutzen-Analysen (chronologisch) von Psychotherapie bei psychischen Erkrankungen (insbesondere Angststörungen); Angaben in DM zu den Wechselkursen 1999 (Werte in Originalwährungen im Text)

Autoren	Diagnose	Therapie	Datenbasis	Kostenfaktoren	Ergebnisse (pro Patient)
Ginsberg et al. (1984)	Neurosen	VT, durch geschulte Pfleger	insgesamt N = 46, VG und KG (randomisierte Zuweisung), Erhebung für 1 Jahr vor und 1 Jahr nach Therapie	ambulante und stat. Behandlungen, Arbeits-unfähigkeit, weitere (z.B. Reisekosten)	<ul style="list-style-type: none"> durchschnittl. Therapiekosten bei 40 Pat.: 289 engl. Pfund durchschnittl. Nutzen: 152 Pfund => PT-Kosten nach 1,9 Jahren aufgewogen
Gathmann (1990)	Psychosom. Störungen	stationäre Behandlung	N = 110, retrospektive Erhebung 4 Jahre nach Therapieende	ambulante und stat. Behandlungen, Medikamente, Arbeitsunfhgkt.	<ul style="list-style-type: none"> durchschnittl. Nutzen: ca. 28000.- DM Nettonutzen je Patient: ca. 11100.- DM
Zielke (1993)	Neurotische u. psychosom. Erkrankungen	stationäre VT	N = 148, Erhebung für 2 Jahre vor und 2 Jahre nach d. Therapie	psychol. / medizin. Versorgung, stat. Aufenthalte, Medikam., Arbeitsunf.	<ul style="list-style-type: none"> Kosten-Nutzen-Verhältnis je Patient betrug 1 : 3,46; je erwerbstätigen Patient 1 : 2,49
Salvador-Carulla et al. (1995)	Panik	ambulante PT	N = 61, Erhebung für 1 Jahr vor und das Jahr während der Therapie	medizin. und psycholog. Behandlungen, stat. Aufenthalte, Laboruntersuchungen, Arbeitsunf.	<ul style="list-style-type: none"> psychol. Behandlungskosten stiegen alle anderen Kosten sanken insgesamt Verringerung d. Kosten p.P. im Jahr nach Diagnose: 569 Dollar
Margraf, Schneider & Jacobi (in Vorber.)	Panik, Agoraphobie	kognitive VT	N = 66, Erhebung für 2 Jahre vor und 3 Jahre nach der Therapie	medizin. und psycholog. Versorgung, stat. Aufenthalte, Arbeitsunfähigkeit	<ul style="list-style-type: none"> durchschnittl. Nutzen: DM 8124.- durchschnittl. Therapiekosten: DM 1458.- Kosten-Nutzen-Relation: 1 : 5,6
Breyer (1997)	versch. psych. Störungen	ambulante Psychoanalyse	N = 604, retrospektive Befragung	ambulante und stat. Behandlungen, Medikamente, Arbeitsunfähigkeit	<ul style="list-style-type: none"> durchschnittl. Nutzen: DM 8477,80 (Einzel) DM 14 330.- (Gruppe) durchschnittl. Kosten: DM 33 235.- (Einzel), DM 4305.- (Gruppe) K/N-Verhältnis: 1 : 0,25 (Einzel); 1 : 3,32 (Gruppe)

2.4 Einbezug des subjektiven/intangiblen Nutzens in eine Kosten-Nutzen-Analyse

Bei der bislang beschriebenen Kosten-Nutzen-Analyse aus volkswirtschaftlicher Sicht (und aus der Sicht von Kostenträgern) ist der Nutzen definiert als eingesparte zukünftige Aufwendungen für Behandlungen und als Einsparung durch verminderten Produktivitätsausfall. Aber das Hauptziel eines Gesundheitsprogramms sollte nicht die Vermeidung zukünftiger Ausgaben, sondern die Verbesserung der Gesundheit der zu behandelnden Patienten sein. Im Gesundheitsbereich spielt daher besonders der individuell empfundene Nutzen eine große Rolle, der allerdings kaum in Geldeinheiten messbar ist: Für höhere Lebensqualität und -freude gibt es keine festen Preise.

Im Gegensatz zum organmedizinischen Bereich wurde bislang der intangible Behandlungsnutzen (z.B. Zugewinn an Lebensqualität oder die Entlastung Angehöriger) nicht in Kosten-Studien im Psychotherapiebereich mit einbezogen. Im Rahmen der vorliegenden Arbeit wurde mit der Diplomarbeit von Runde (1999) ein geeigneter Fragebogen entwickelt und eine Pilotstudie durchgeführt.

2.4.1 Die Zahlungsbereitschaft (willingness to pay, WTP)

Im Unterschied zu Kosten-Effektivitäts- oder Kosten-Nutzwert-Analysen kann man mit dem Ansatz der *Zahlungsbereitschaft* den Nutzen eines Gesundheitsprogramms in Geld darstellen. Die Methode ist auch unter dem Begriff „willingness to pay“ (WTP) bekannt.

Man unterscheidet die direkte und die indirekte Messung der Zahlungsbereitschaft: bei der indirekten Methode orientiert man sich an tatsächlichem Verhalten und schließt von offenbarten Präferenzen von Personen auf deren Zahlungsbereitschaft; bei der direkten Methode wird ausdrücklich gefragt, wieviel jemand höchstens für eine bestimmte medizinische Intervention zu zahlen bereit wäre. Diese hypothetische Frage kann im Gesundheitswesen *ex post* (nachdem bereits an einer Intervention teilgenommen wurde) gestellt werden oder *ex ante* (beispielsweise als Frage nach der Höhe des Versicherungsbeitrages, den man zu zahlen bereit wäre, falls die Versicherung auch besagte Intervention umschließt).

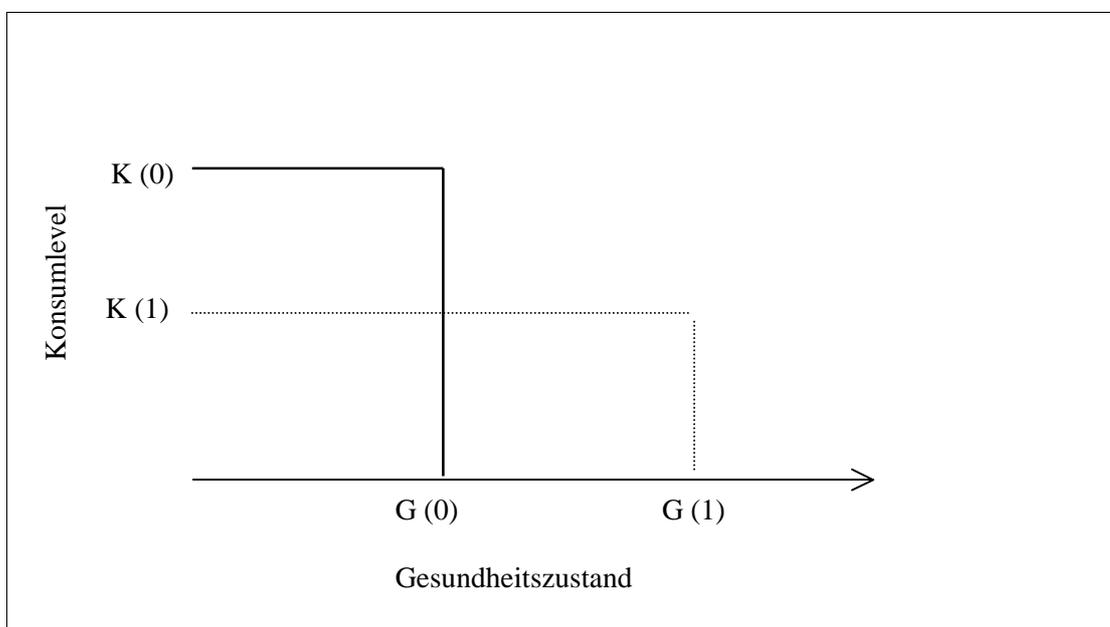
Der WTP-Ansatz fußt auf einer tatsächlichen theoretischen ökonomischen Basis. Diese ist in dem für ökonomische Laien schwer verständlichen Standardwerk von Mitchell und Carson (1989) beschrieben. Die Methode baut auf der Prämisse auf, dass der angegebene Geldbetrag Indikator für den Wert ist, den das Gut für das Individuum hat. Galler (1990, S. 38, Übers. durch Verf.)

stellt die zugrundeliegenden Überlegungen wie folgt dar: Soziale Bewertungen finden dadurch statt, dass Individuen auf einer Kardinalskala ihre Vorlieben angeben und diese Ergebnisse dann zusammenfassen. Im Grunde passiert genau das in der Wirtschaft, wenn Individuen Entscheidungen darüber treffen, wieviel sie für ein bestimmtes Produkt zahlen. Der willingness-to-pay-Ansatz weitet dieses Konzept auf Güter aus, die man nicht auf dem Markt kaufen kann.

Bei der Anwendung im Gesundheitswesen geht es darum zu bewerten, wieviel jemandem eine bestimmte Verbesserung seines Gesundheitszustandes von G_0 bis G_1 wert ist. Dies wird dadurch bestimmt, dass man den Geldbetrag bzw. den Betrag des Konsums bestimmt, den man für diese Verbesserung opfern würde. Wenn K_0 der Konsumbetrag ist, der zu G_0 gehört und K_1 der zu G_1 gehörige Konsumlevel, dann müsste die Differenz $K_1 - K_0$ als (monetäre) Messung der Präferenz des Individuums genutzt werden können“.

Zur Veranschaulichung dieser Aussage kann man sich vorstellen, dass der Flächeninhalt der beiden Vierecke in Abbildung 2.5 (durchgezogene vs. gestrichelte Linie) die Lebensqualität darstellt; sie ist in beiden Versionen gleich hoch. Ein besserer Gesundheitszustand kann also einen niedrigeren Konsumlevel kompensieren und umgekehrt.

Abbildung 2.5: Grafische Darstellung des Zusammenhangs von Gesundheit, Konsum und Lebensqualität (Galler, 1990, S.38)



2.4.2 Die Ermittlung der Zahlungsbereitschaft

Die Zahlungsbereitschaft wird in der Regel direkt ermittelt und kann per Interview oder Fragebogen erhoben werden. Im wesentlichen geht es immer darum, in Gedankenspielen zu erfassen, was die Probanden für einen intangiblen Nutzen (z.B. Verbesserung des Gesundheitszustandes oder Bereinigung eines Umweltschadens) zu zahlen bereit wären.

Da es, ähnlich wie schon bei der Methodik der Kosten-Nutzen-Analyse erläutert, viele Facetten bei der Umsetzung der Erhebung gibt, haben sich auch hier Experten zusammengefunden, um Leitlinien für die Durchführung und Darstellung von WTP-Studien zu erstellen.

Zunächst sollen jedoch noch allgemeine Fragen zur Güte einer solchen Erhebung und Störeinflüsse auf die Validität der Aussagen erörtert werden.

2.4.2.1 Reliabilität und Validität der Frage nach der Zahlungsbereitschaft

Eine akzeptable Reliabilität (Retest) der Methode konnte bisher mehrfach nachgewiesen werden (z.B. Loomis, 1990; O'Brien und Viramontes, 1994). Nach Drummond et al. (1987) macht es jedoch jeweils einen großen Unterschied, ob die Befragten mit dem Untersuchungsgegenstand vertraut sind. Je besser das Wissen der Versuchspersonen, um so zuverlässiger scheinen die Ergebnisse.

Verschiedene Autoren überprüften die Validität des Ansatzes durch den Test unterschiedlicher Hypothesen.

1) Es besteht ein Zusammenhang zwischen der Höhe der Zahlungsbereitschaft und der Höhe des Einkommens.

Dieser Vergleich ergab immer einen signifikanten Zusammenhang (Ryan, 1996, und O'Brien und Viramontes, 1994). Lediglich bei Johannesson et al. (1991) war dieser Zusammenhang nur tendenziell feststellbar.

2a) Es gibt einen Zusammenhang zwischen der Größe des versprochenen Effektes und der Zahlungsbereitschaft.

Diese Hypothese wurde im Umweltbereich überprüft und konnte nicht bestätigt werden. Beispielsweise stellte Devousges (1992), zitiert nach Arrow et al. (1993) fest, dass die Zahlungsbereitschaft, 2000 Vögel vor dem Öltod zu retten, ungefähr gleich groß war wie diejenige, 20000 oder 200000 Vögel vor dem Öltod zu retten. Möglicherweise ist hier aber der Umgang des Messin-

strumentes Mensch mit hohen Zahlen stark überschätzt worden: den Unterschied zwischen 2000 und 20000 Vögeln kann man sich schwer vorstellen. Andere Ergebnisse wären sicherlich mit dem Vergleich niedrigerer Zahlen zu erzielen.

2b) Es gibt bei Anwendungen der Methode im Gesundheitsbereich einen Zusammenhang zwischen dem Schweregrad der Beeinträchtigung und der Zahlungsbereitschaft für eine Intervention, die Verbesserung verspricht.

Diese Hypothese entspricht der Hypothese 2a) bezogen auf den gesundheitlichen Bereich: je stärker die Beeinträchtigung, desto größer auch die Erleichterung bei Besserung. Dieser Zusammenhang konnte (für subjektive Erfolgskriterien) bestätigt werden (O'Brien und Viramontes, 1994, Thompson und Cohen, 1984). Bei Thompson und Cohen (1984) handelte es sich allerdings nur um einen tendenziellen, keinen signifikanten Zusammenhang.

3) Die behauptete Zahlungsbereitschaft entspricht der Höhe der realen Zahlung bei entsprechender Gelegenheit.

Arrow et al. (1993) stellten fest, dass die im Interview ausgedrückte Zahlungsbereitschaft für eine Umweltorganisation signifikant höher war als die Zahlungen, die die Befragten tatsächlich leisteten, als sie um einen Beitrag gebeten wurden. In einer weiteren Studie zahlten von denjenigen, die angegeben hatten, zahlen zu wollen, signifikant weniger Personen, als es tatsächlich ums Zahlen ging. Andererseits war bei den Leuten, die tatsächlich zahlten, der vorher genannte Geldbetrag mit der tatsächlichen Zahlung identisch.

2.4.2.2 Mögliche Störeinflüsse

Arrow et al. (1993) vermuten die folgenden unterschiedlichen Störeinflüsse als Ursache für die teilweise nicht feststellbaren bzw. nicht signifikanten Zusammenhänge. Diesen kann man durch geeignete Frageformulierung entgegenwirken.

- *Strategischer Bias*: Die Befragten entscheiden sich für enorm hohe Beträge, um anstehende staatliche Budget-Entscheidungen zu ihren Gunsten zu beeinflussen, oder sie entscheiden sich für niedrige bzw. Nullbeträge, weil sie möchten, dass auch in Zukunft weiter der Staat/ die Krankenkasse o.ä. die Kosten trägt. Wenn man Psychotherapiepatienten befragt, ist gerade mit strategischen Bias in besonderem Maße zu rechnen: entweder mit der Motivation, Entscheidungsträger dahingehend zu beeinflussen, Psychotherapie

als finanzierenswert zu betrachten (hoher WTP - Wert) oder aber, um eine vermutete Anhebung der Kosten der Psychotherapie zu vermeiden (niedriger Wert). Es ist davon auszugehen, dass diese Motivation vor der Therapie (ex ante) höher ist, als ex post, da derartige Änderungen die Person selbst nicht mehr betreffen werden.

- *Sensitivität gegenüber irrelevanten Faktoren:* Die Höhe der Zahlungsbereitschaft fällt je nach Frageformulierung bzw. Inhalt der vorgegebenen Skala bei geschlossenen Antworten unterschiedlich hoch aus. Will man unterschiedliche Therapieformen mit dieser Methode vergleichen, wäre es daher notwendig, einen standardisierten Fragebogen zu entwickeln.
- *Motivation und Emotion:* Die Meinung, dass man den Wert des Lebens bzw. der Lebensqualität sowieso nicht in Geld bemessen kann (bzw. die Beurteilung der Frage als ethisch unangemessen), kann zur Boykottierung der Frage führen.
- *Nichteinbezug des Budgets:* Die Bereitschaft zur Zahlung einer gewissen Summe wird von den Probanden teilweise wieder korrigiert, wenn gefragt wird, welche anderen Ausgaben dafür in den Hintergrund treten müßten. Das heißt, dass solche Überlegungen nicht immer bei der WTP berücksichtigt werden und deshalb im Erhebungsinstrument angesprochen werden müssen.
- *Verständnis und Akzeptanz:* Besonders wichtig ist es natürlich, Befragten genaue Informationen über das Szenario und die Intervention zu geben. Ob ein zu säubernder Fluß sich von selbst in 100 Jahren oder in 10 Jahren regenerieren wird, ist bei der Bereitschaft, dafür etwas zu zahlen von entscheidender Bedeutung. Dennoch ist es fraglich, ob alle diese Informationen, wenn zur Verfügung gestellt, auch tatsächlich in die Entscheidung mit eingearbeitet werden, ohne von oberflächlichen Heuristiken wie „Diese Umweltunfälle sind meist nie so schlimm, wie man tut“ überlagert zu werden. Bei der Befragung von Psychotherapiepatienten nach der Therapie ist davon auszugehen, dass sie keine derartigen Heuristiken nutzen, da sie genau über Prozeß und Wirkung der Therapie informiert sind.
- *Warm glow:* Die geäußerte Zahlungsbereitschaft kann für den Befragten den Charakter von Spenden haben, bei denen es nicht nur darum geht, eine Organisation in einer bestimmten Höhe, die sich nach dem vermuteten Wert ihrer Arbeit richtet, zu unterstützen, sondern auch um das angenehme Gefühl, dass man sich selbst beteiligt hat („warm glow“). Die Verteilung der Zahlungsbereitschaft für Umweltprogramme ist der bei Spendenaktionen recht ähnlich: Entweder geben die Leute nichts oder aber eine recht hohe

Summe. Daher könnte die Zahlungsbereitschaft weniger den tatsächlichen Wert des Programmes für den einzelnen bedeuten, als vielmehr eine generelle Bereitschaft zu spenden, falls das Programm akzeptiert wird. Dies wäre eine weitere Erklärung für die oben erwähnte Insensitivität gegenüber Quantität bei Umweltprogrammen. Da bei der Psychotherapie hauptsächlich der Befragte selbst profitiert, die Zahlungsbereitschaft also nicht altruistisch motiviert ist und der Wert der Therapie von der Verbesserung des Wohlbefindens des Patienten abhängt, trifft dieser Effekt auf die Zahlungsbereitschaft für Psychotherapie nicht zu.

- *Wichtigkeitshypothese*: Eine andere Erklärung für die Unabhängigkeit der Antwort von der Quantität des Effekts ist die Wichtigkeitshypothese (prominence hypothesis) nach Baron, (1997), die besagt, dass es ein Konzept von der Wichtigkeit eines Gutes gibt, das völlig unabhängig von der Quantität ist. Geldwerte würden also in Hinsicht auf die Wichtigkeit gebildet, die man dem Gut zumißt, unabhängig von dessen Quantität. So stellten Johannesson et al. (1991) fest, dass die Zahlungsbereitschaft für unterschiedliche Behandlungsalternativen stets gleich hoch war, wenn man die Patienten nicht ausdrücklich auf andere Alternativen hinwies. Sie schlußfolgerten, dass die Patienten die Frage im Sinne der Alternativen „Behandlung versus keine Behandlung“ beantworteten.

Aufgrund der beschriebenen Störeinflüsse ist die Methode der Zahlungsbereitschaft umstritten. Einige Kritiker lehnen sie daher völlig ab (vgl. Warner und Luce, 1982). Im ökologischen Bereich hingegen kann man sie eine etablierte Methode nennen. Arrow et al. (1993) kommen zu dem Schluß, dass man bei Beachtung einiger Regeln zu zuverlässigen und gültigen Ergebnissen kommen kann. Daher haben die Autoren im Rahmen des NOAA (National Oceanic and Atmospheric Administration) Panel of Contingent Valuation Richtlinien für eine korrekte Anwendung der Methode entwickelt.

2.4.2.3 Richtlinien für WTP-Studien

Die hier aufgeführten Richtlinien (NOAA Panel of Contingent Valuation) beziehen sich auf Untersuchungen im Umweltbereich. Die Umsetzung dieser Richtlinien sowie die Behandlung eben aufgeführter Störeinflüsse in der vorliegenden Arbeit wird im entsprechenden Methodenteil (4.5) dargestellt.

1. Bei der Auswertung ist im Zweifelsfall stets die Schätzung zu wählen, die die Zahlungsbereitschaft niedriger hält (konservative Schätzung).
2. Das Programm muß akkurat beschrieben werden. Auch an mögliche Alternativen muß erinnert werden; wenn beispielsweise die Zahlungsbereit-

- schaft für die Säuberung eines Flusses ermittelt wird, muß daran erinnert werden, dass es noch andere saubere Flüsse in der Umgebung gibt.
3. Die Befragungen sollten zu unterschiedlichen Zeitpunkten nach dem Umweltunglück durchgeführt werden. Unterscheiden sich die Aussagen zu den unterschiedlichen Zeitpunkten stark voneinander, ist das Ergebnis nicht gültig.
 4. Die Gründe für die Antworten auf die Frage nach der Zahlungsbereitschaft sollten durch offene Fragen eingeholt werden, um die Antworten richtig zu interpretieren.
 5. Als Interpretationshilfe der Antworten sollten weitere Informationen wie Höhe des Einkommens, Einschätzen der Größe der Katastrophe etc. erhoben werden.
 6. Am Ende des Fragebogens ist zu prüfen, inwiefern er verständlich für die Probanden war, und ob er akzeptiert wurde.
 7. Um valide Antworten zu erzielen, sollte man daran erinnern, dass die Probanden den angegebenen Geldbetrag nicht mehr für andere Ausgaben nutzen können.
 8. Die Fragen sollten mit vorgegebenen Antwortalternativen erfolgen. Arrow et al. (1993) gehen von einer höheren Validität geschlossener Fragen aus, da diese eine höhere Ähnlichkeit mit der tatsächlichen Situation eines Käufers hätten: Ein Käufer überlegt nicht, welchen Preis er höchstens ausgeben würde, sondern er sieht den Preis und überlegt, ob ihm die Ware diese Summe wert ist.

Bezüglich der Entscheidung über offene und geschlossene Fragen argumentiert Baron (1995), dass die oben beschriebene Augenscheinvalidität der geschlossenen Fragen nicht bewiesenermaßen zu gültigeren Ergebnissen führt. Er führt verschiedene Studien an, in denen die offenen Fragen sogar Ergebnisse höherer Validität zu geben scheinen: Beispielsweise testete Mc Fadden (1994) die Bereitschaft, Geld zum Schutze von Naturreservaten auszugeben. Bei der offenen Frage war die Zahlungsbereitschaft, 57 Reservate zu schützen, 57% höher als die, nur eines zu schützen, wogegen der Unterschied bei den geschlossenen Fragen nur 33% betrug. Das heißt, dass die Größe der Intervention bei den offenen Fragen genauer berücksichtigt wurde. Bei den geschlossenen Fragen war sogar die Bereitschaft, für ein Reservat zu zahlen höher als die für drei Reservate zu zahlen. Eine Überlegenheit der geschlossenen Fragen hinsichtlich der Validität ist daher nicht als nachgewiesen anzusehen. In einer Studie von Johannesson et al. (1991) zeigte sich, dass die Anzahl der tatsächlich verwendbaren Antworten bei den offenen Fragen wesentlich geringer ist, da hier häufig anstelle eines exakten Betrages Äußerungen wie „viel Geld“ o.ä. angegeben werden. Nachteile der geschlossenen Antworten liegen in der Beeinflussung der Probanden durch vorgegebene Skalen.

2.4.2.4 Weitere Aspekte bei der Ermittlung der WTP

Aus verschiedenen Artikeln stammen die weiteren Hinweise zum Vorgehen. Thompson und Cohen (1990) weisen darauf hin, dass Absicht und Gewichtigkeit des eingereichten Fragebogens den Versuchspersonen erklärt werden müssen.

Lee et al. (1997) schlagen vor, bei der Auswertung der mittleren Zahlungsbereitschaft den Median anstelle des Mittelwerts zu nutzen, da er die zentrale Tendenz konservativer wiedergibt und weniger sensitiv gegenüber extremen Antworten ist.

Fellmann (1996) bezeichnet die Frage nach dem maximalen Verlust an Freizeit, den ein Individuum in Kauf nehmen würde, um an einer Therapie teilzunehmen, als interessante Alternative zur Frage nach dem finanziellen Betrag. Logisch erscheint, dass besonders Einkommensschwache, die nicht viel Geld über das Lebensnotwendige hinaus haben, hier einen Weg finden, ihre Zahlungsbereitschaft auszudrücken. Wenn die Befragten nicht alle am unteren Ende der Einkommensskala liegen, kann sie dazu dienen, die interne Konsistenz der Antworten zu überprüfen.

Im Gesundheitswesen können sowohl direkt Betroffene (also Träger einer bestimmten Symptomatik, Risikopopulationen) als auch nicht Betroffene befragt werden. Bei letzteren wird die Zahlungsbereitschaft für eine Reduktion des Erkrankungsrisikos für eine bestimmte Krankheit erfragt bzw. die Bereitschaft einen höheren Krankenkassenbeitrag zu leisten, um im Krankheitsfalle bestimmte Behandlungsmethoden erstattet zu bekommen. Die Frage wäre, welcher höheren Betrag jemand zahlen würde, wenn eine bestimmte Leistung im Versicherungsschutz enthalten wäre. Gegen diese Befragung nicht Betroffener spricht, dass sie nicht genügend Wissen über die betreffenden Krankheiten und Interventionen haben. Auch das Erkrankungsrisiko können sie nur schwer einschätzen, da sie meist mit der Information, dass z.B. einer von 100 erkranken wird, wenig verbinden können. O'Brien und Gafni (1996) argumentieren aber, dass die Frage nach der Zahlungsbereitschaft an alle gestellt werden soll, die Beiträge an die Krankenkassen zahlen. Nur so könne diese Methode einen Beitrag leisten, tatsächliche Erstattungsfragen zu beantworten.

Befragt man direkt Betroffene und daher besser Informierte, kann man zwar Ergebnisse höherer Validität erwarten (vgl. Kruschke und Petermann, 1997), allerdings spiegelt dies dann nicht die Meinung der Bevölkerung wieder, sondern nur die einer besonderen Gruppe.

2.4.3 Einbezug der WTP in Kosten-Nutzen-Analysen

Natürlich ist das Ergebnis einer Umfrage zur Zahlungsbereitschaft erst dann aufschlußreich, wenn es in Bezug zu den tatsächlichen Kosten gesetzt wird.

Man kann zunächst annehmen, dass dem individuellen Nutzen auch die individuellen intangiblen Kosten gegenübergestellt werden müßten. Man geht aber davon aus, dass sich die Zahlungsbereitschaft zusammensetzt aus dem individuell eingeschätzten Aufwand für eine Behandlung, der in Bezug zum Nutzen gesetzt wird. Bei der Angabe der WTP durch das Individuum sind die intangiblen Kosten der Behandlung also bereits enthalten und in die angegebene Summe reduzierend eingeflossen.

O'Brien et al. (1995), berechnen im Sinne der Kosten-Nutzen-Analyse den Netto-Nutzen eines getesteten Anti-Depressivums, in dem von dem Mittelwert der angegebenen WTP-Summen die Kosten für das Medikament abgezogen werden.

Fellmann (1996) definiert den Gesamtnutzen einer Intervention als den anhand WTP ermittelten Nutzen plus den Nutzen, der sich durch eingesparte Behandlungskosten errechnet. Es werden also sowohl volkswirtschaftlicher als auch subjektiver Nutzen monetär in die Kosten-Nutzen-Analyse einbezogen.

Nach Johannesson et al. (1994) ist aufgrund des experimentellen Status, in dem sich die Methode WTP noch befindet, keine Aussage darüber möglich, inwieweit solche Befunde tatsächlich in Allokationsentscheidungen einbezogen werden können. Danach scheint es nicht sinnvoll, die Summen eingesparte Behandlungskosten, Zahlungsbereitschaft und tatsächliche Kosten zu vermengen und einen gemeinsamen Kosten-Nutzen-Index zu bilden, da sie einen unterschiedlichen Grad von überprüfter Objektivität und Validität aufweisen.

2.4.4 Forschungsstand zur Methode der Zahlungsbereitschaft

Als Grundlage der Ausführungen zum Thema Zahlungsbereitschaft/WTP wurde eine Literaturrecherche in den Datenbanken PsycLit (1981 - 4/1998), Psynindex (1977 - 9/1998), Medline und Econolit (1969 - 8/1998) durchgeführt.

Folgende Stichwörter wurden genutzt:

Zahlungsbereitschaft, willingness to pay, contingent valuation

in unterschiedlicher Weise kombiniert mit

Therapie/ therapy, Psychotherapie/psychotherapy, health.

Zusätzlich wurde den in der Datenbank Medline enthaltenen Querverweisen zu ähnlichen Artikeln sowie den Verweisen in den Artikeln selbst nachgegangen. Keine Einträge gab es zu der jeweiligen Kombination mit Therapie/therapy bzw. Psychotherapie/psychotherapy, das heißt, es gibt bisher noch keine Dokumentationen über die Anwendung der Methode Zahlungsbereitschaft im Zusammenhang mit Psychotherapieevaluation. Tabelle 2.5 zeigt Anwendungen der Methode im organmedizinischen Gesundheitsbereich, bei denen Betroffene zu ihrer Zahlungsbereitschaft bei Therapiealternativen befragt wurden.

Es handelt sich also um eine noch relativ dürftige Forschungslage, was empirische Ergebnisse zur WTP im Gesundheitsbereich betrifft. Es ist aber davon auszugehen, dass das Maß in der Gesundheitsökonomie zunehmend an Bedeutung gewinnen wird (Drummond et al., 1986; Fellmann, 1996).

Zudem erscheint es allein schon deswegen reizvoll, hier für den Psychotherapiebereich Neuland zu betreten, weil die WTP derzeit als die einzige vorhandene Möglichkeit erscheint, eine Kosten-Nutzen-Analyse mit der intangiblen Perspektive abzurunden.

Tabelle 2.5: Übersicht über verschiedene Ergebnisse der Methode Zahlungsbereitschaft im Gesundheitsbereich

Studie und Land	Thema	Absolute WTP (Angaben von Betroffenen; in DM zu den Wechselkursen von 1999)	WTP als Prozentsatz des Nettoeinkommens
Johannesson et al. (1991), Schweden	Behandlung von Bluthochdruck	Mittelwert: 5500 SK Median: 2900 SK	Keine Angabe
O'Brien und Viramontes (1994), Kanada	Behandlung chronischer Lungenkrankheit, die zu 99% die absolute Heilung verspricht und mit der Wahrscheinlichkeit von 1% den Tod bedeutet	Mittel: 113 \$ im Monat Median: 65\$ im Monat	5% des monatlichen Einkommens für die Dauer eines Jahrs = insgesamt 60% eines Monatseinkommens
O'Brien et al. (1995), Kanada	Antidepressivum mit weniger Nebenwirkungen	22\$ pro Monat (zusätzlich zu dem, was bereits bezahlt wird für das jetzige Antidepressivum)	Keine Angabe
Donaldson et al. (1997), GB	Genetisches Trägerscreening für zystische Fibrose	18-25 Pfund für die unterschiedlichen Testalternativen	Keine Angabe
Lee et al. (1997), USA	Autologe Blutgabe: Patienten können vor einer Operation Eigenblut abgeben zur Nutzung im Falle einer nötigen Blutzufuhr im Laufe der Operation	913 \$ insgesamt	Keine Angabe
Ryan (1996), Australien	In-Vitro-Fertilisation	Mittelwert: 2506 \$ Median: 2250 \$	Keine Angabe

3 Kosten-Effektivitäts-Analyse

3.1 Fragestellungen

Es wurde bereits dargestellt, dass es zwar viele Hinweise für die Wirksamkeit von Psychotherapie gibt, aber dass andererseits auch in vielen Studien nicht (primär) unter klinisch relevanten Bedingungen gearbeitet wurde (bzw. aufgrund der Fragestellung nicht unter klinisch relevanten Bedingungen gearbeitet werden konnte). Diese Ergebnisse sollten nun um Effektivitätsstudien im praxisrelevanten Setting bzw. um Ergebnisse zum Erfolg der Umsetzung wissenschaftlich fundierter Behandlungsverfahren in den Versorgungsalltag ergänzt werden.

Weiterhin wurde herausgestellt, dass komplexe Verfahren wie die Kosten-Effektivitäts-Analyse zur Anwendung kommen sollten, so dass mehrere Dimensionen des untersuchten Gegenstandes Psychotherapie berücksichtigt werden.

Die Übersicht über den Forschungsstand legt nahe, dass es insbesondere im ambulanten Bereich an empirischen Ergebnissen zur Kosteneffektivität mangelt. Mit dieser Arbeit wird also – zumindest für einen Ausschnitt aus dem Bereich möglicher Störungen und Therapieverfahren – eine Lücke allein dadurch geschlossen, indem konkrete Kosten-Effektivitätswerte berichtet werden.

Fragestellung 1: *Wie kosteneffektiv ist das untersuchte Treatment?*

„Die“ Kosten-Effektivität ist nur mit anderen Kosten-Effektivitäts-Aussagen vergleichbar, wenn sowohl Kosten- als auch Effektivitätsseite klar definiert und sinnvoll operationalisiert sind. Ferner sollte die Effektivitätsmessung praktikabel sein, d.h. sparsam ausgewählt und so einfach zu erheben und auszuwerten, dass sie in die Maßnahmen eines umfassenden Qualitätssicherungsprogramms integriert werden kann.

Folgende Überlegungen leiteten, ausgehend von Erkenntnissen der Outcome-Forschung (vgl. 2.2.2), die Auswahl und Auswertung der Erhebungen:

- *Erhebung der Effektivität auf mehreren Ebenen und multimodal:* Da eine störungsspezifische Behandlungsform untersucht wird, entspricht es den Behandlungszielen, Maße einzubeziehen, *die störungsspezifische Veränderungen* messen. Daher wurden zum einen etablierte Symptomfragebögen eingesetzt, die den Diagnosekriterien möglichst entsprechen. Dennoch sollte auch die *Komorbidität* kontrolliert werden, um die Breite der

Veränderung abbilden zu können. Dafür eignen sich übergreifende Symptomfragebögen und störungsspezifische Fragebögen anderer Diagnosen. Entscheidend für die Vergabe einer Diagnose sind die Störungsfolgen, also übermäßige *Beeinträchtigungen und Belastungen durch die Störung* für das Individuum. Diese Ebene wurde mit Items zur Beeinträchtigung in wichtigen Lebensbereichen (Arbeitsleben, Freizeit, Familienleben und häusliche Pflichten) erhoben.

Schließlich kann noch auf einer ganz *globalen Ebene* direkt nach der subjektiven allgemeinen Veränderung im Zuge der Therapie gefragt werden. Dieses Maß ist sehr anschaulich und besitzt hohe Augenscheinvalidität. Außerdem liegt es sowohl aus Patienten- als auch aus Therapeutesicht vor, so dass mit dem Therapeutenurteil auch eine Modalität der Fremdeinschätzung genutzt werden kann. Es ist aber weitgehend unklar, was in die Antworten alles einfließt; deshalb sollte es bei Auswertung und Präsentation stets mit anderen der o.g. Maße kombiniert werden.

Individuelle Zielerreichung blieb in der vorliegenden Studie unberücksichtigt, da aufgrund Schwierigkeiten bei der Standardisierung z.B. beim Goal Attainment Scaling: positive und negative Ziele, sowie Ziele unterschiedlicher Erreichbarkeit und Valenz sind insbesondere bei der Aggregation der Ergebnisse in einer großen Stichprobe schwer vergleichbar, und komplexe validierte Zielmessungen wie mit dem Fragebogen zur Analyse Motivationaler Schemata (FAMOS; Grosse Holtfort, 2000) lagen zum Zeitpunkt unserer Untersuchung noch nicht vor.

- *Interessanter als die Ermittlung statistischer Signifikanz: Ermittlung von Effektstärken und von klinisch signifikanten Verbesserungsraten*

Auf den gewählten Ebenen (Symptom, Beeinträchtigung, global) sollten nachvollziehbare Kriterien für die „klinisch signifikante Verbesserung“ bestimmt und verwendet werden. Damit lässt sich aus unterschiedlichen Perspektiven die Rate der Patienten ermitteln, die individuell bedeutsam gebessert waren. Die Effektstärken, die die durchschnittlichen Veränderungen standardisiert darstellen, sind vergleichbar mit anderen Studien, in denen die selben (oder hinreichend ähnliche) Maße eingesetzt worden sind. Sie sollen dementsprechend bewertet werden, insbesondere dahingehend, ob die Effekte aus kontrollierten Therapiestudien hier im klinische Alltag repliziert werden können.

Die Kosten-Effektivität sollte für die verschiedenen Effektivitäts-Bestimmungen getrennt ermittelt und diskutiert werden. Ziele der Fragestellung sind die differenzierte Veranschaulichung des Maßes der Kosten-Effektivität und die Bereitstellung von Vergleichswerten (z.B. für Metaanalysen, die es derzeit für Kosten-Effektivitäts-Studien in umschriebenen Settings noch nicht geben kann).

Fragestellung 2: Wie variabel ist die Kosten-Effektivität in Abhängigkeit von Berechnungsmethoden oder (Sub-) Stichproben?

Außerdem erlaubt der große Datensatz weitere Untersuchungen und Differenzierungen. Von den Einflüssen, die die Kosten-Effektivität moderieren könnten, sollen folgende untersucht werden:

- **Mess-Ebene und Mess-Methode:** Die Variabilität der Kosten-Effektivität aufgrund der in der ersten Fragestellung verwendeten Vielzahl der Effektivitätsmaße soll genauer ausgewertet werden. Zu erwarten ist etwa, dass (bei gleichen Kosten) im globalen Maß bessere Effektivitätswerte zu verzeichnen sind als im übergreifenden Symptomfragebogen. Denn erstens gibt es bei retrospektiven Fragen in der Regel einen *bias* in Richtung Verbesserung (mögliche Erklärung durch Dissonanzreduktion oder soziale Erwünschtheit), wohingegen der Vergleich durch die Differenzbildung von Prä- und Postwerten der Fragebögen den Patienten in weit geringerem Maße bewusst ist. Dies geht mit einer erwarteten kleineren Varianz im globalen Maß einher und beeinflusst damit die Höhe der Effektstärke zusätzlich, da diese mit abnehmender Varianz wächst. Außerdem geht in das globale Maß mehr ein als lediglich eine Symptomreduktion, was im Sinne einer Addition von Einzeleffekten eine eher große Verbesserung erwarten lässt.

Schließlich macht sich die in dieser Arbeit ausführlich untersuchte klinische Signifikanz der Ergebnisse (d.h. auch die Frage des „Anteils erfolgreich behandelter Patienten“) je nach inhaltlicher Betrachtungsweise auch bei der Kosteneffektivität bemerkbar: Was für einen Einfluss auf die Kosteneffektivität hat die Differenzierung der Bewertung der Patientengesundheit um eine subklinische Kategorie?

- **Therapie-Preis:** Analog zur vor allem im Zusammenhang mit Kosten-Nutzen-Analysen relevanten *Sensitivitätsanalyse* sollen die Kosten-Effektivitäts-Werte auch für (ggf. hypothetische) andere Therapiekosten ermittelt werden.
- **Störungsspezifische Variablen:** Es ist aufgrund vieler Vorbefunde zu erwarten, dass sich verschiedene Störungsgruppen sowohl bezüglich der Effektivitätswerte als auch bezüglich der benötigten Therapiedauer unterscheiden. In diesem Zusammenhang sollen hier die Kosten-Effektivität der Therapien in den Subgruppen Panikstörung mit Agoraphobie, Sozialphobie und Zwangsstörung verglichen werden. Außerdem soll festgestellt werden, ob die Krankheitsdauer (und damit die Chronifizierung der Störung) einen Einfluss auf die Kosten-Effektivität durch Moderierung von Therapieerfolg oder Therapiedauer hat.

- **Soziodemografische Variablen** (*Alter, Geschlecht, Bildungsstand*): Üblicherweise ist es sinnvoll, verschiedene Subgruppen zur vergleichen, um Hinweise auf die Generalisierbarkeit der Ergebnisse zu finden. Unterscheiden sich solche Subgruppen bedeutsam, so müssten differentielle Effekte der Therapie oder systematische Unterschiede in der Therapiedauer genauer untersucht werden.

Fragestellung 3: Sind die vorgefundenen Effekte aus dem naturalistischen Setting in ihrem Ausmaß vergleichbar mit Effekten kontrollierter Wirksamkeitsstudien kognitiv-behavioraler Therapie?

Der Datensatz, der den Ergebnissen dieser Studie zugrunde liegt, ist außergewöhnlich dahingehend, dass er einerseits unter Rahmenbedingungen entstanden ist, die eine wissenschaftliche Auswertung ermöglichen. Standardisierte Eingangsdiagnostik, Post- und Follow-Up-Messungen, Einsatz einer großen Bandbreite etablierter Therapieerfolgsmaße, zwar nicht manualisierte, aber doch in Struktur und Strategie hinreichend einheitliche supervidierte kognitiv-behaviorale Therapien ermöglichen einen wissenschaftlichen Zugang, der in der Praxis normaler Alltagsversorgung normalerweise nicht gegeben ist.

Andererseits weisen die untersuchten Treatmentbedingungen ein höheres Maß an externer Validität auf als dies in kontrollierten Wirksamkeitsstudien möglich ist. Die klinische Relevanz der Belastung und Beeinträchtigung durch die Störung (auch im sozialgesetzlichen Sinne der Behandlungsbedürftigkeit), Heterogenität sowohl der untersuchten Patienten (oft ist z.B. Komorbidität ein Ausschlusskriterium in kontrollierten Studien) als auch der durchgeführten Therapien (Selbstselektion der Behandlungsbedingung, individuelle Behandlungsdauer, keine Manualisierung im engeren Sinne, selbstkorrigierende und keine „theoretisch reine“ Behandlungsform, viele verschiedene Therapeutinnen und Therapeuten, verschiedene Behandlungsorte) entsprechen eher dem klinischen Alltag.

Für Settings, in denen die klinische Nützlichkeit liegen bislang kaum empirische Effektivitätsergebnisse vor (s. 2.1.2; vgl. auch Goldfried & Wilson, 1998). Die Implementierung eines unter kontrollierten Bedingungen entwickelten Treatments in den Praxis-Alltag muss allerdings keineswegs ein Sprung ins Ungewisse sein. Eine sogenannte „Benchmarking“-Strategie (entliehen aus Qualitätsmanagement-Konzepten der Wirtschaft) kann dabei helfen: in der Praxis wird zwar auf eine Kontrollbedingung zur Wirksamkeitsprüfung verzichtet, jedoch der Outcome permanent evaluiert. Die Ergebnisse dieser Evaluation werden jetzt denen gegenübergestellt, die zuvor unter kontrollierten Bedingungen erzielt wurden. Liegen sie in einem vergleichbaren Rahmen, so kann die Umsetzung als erfolgreich betrachtet werden.

Entsprechend dieser Logik, die bereits bei Wade et al. (1998) zur Anwendung kam, sollen in dieser Arbeit die Effekte auch jenseits von Kostenge-

sichtspunkten unter der „Benchmarking“-Perspektive bewertet werden; nicht zuletzt weil Kostenstudien bzw. Kosten-Nutzen-Analysen wenig Sinn machen, wenn die Effektivität der untersuchten Maßnahme unklar ist. Die standardisierte Erhebung, die naturalistische Heterogenität von Patienten und Therapien, sowie die Größe des Datensatzes – mit knapp 500 Patienten übersteigt er deutlich die Größe üblicher Stichproben – legen nahe, die gefundenen Ergebnisse mit denen in Meta-Analysen kontrollierter Studien vorgefundenen Effekten zu vergleichen.

3.2 Methoden

3.2.1 Behandlungskosten

Die Kostenseite wurde mit dem üblichen *Betriebskostenansatz* berechnet (vgl. 1.3.2), indem die Zahl der in Rechnung gestellten Behandlungseinheiten (bzw. der Mittelwert der untersuchten Gruppe) mit dem Preis pro Behandlungseinheit multipliziert wurde. Da die Studie multizentrisch war und verschiedene Preise für eine Behandlungseinheit existierten (DM 138.- „Westsatz“, DM 111.80 (bis 1996) bzw. DM 117.30 (ab 1996) „Ostsatz“) wurden entsprechende gewichtete Werte berechnet.

3.2.2 Verwendete Effektivitätsmaße

Die im folgenden aufgeführten Items und Fragebögen finden sich in Anhang. Ausgehend von den Vorüberlegungen aus 2.2.2 wurden folgende Maße eingesetzt:

1) globale Ebene: Item zur globalen Erfolgseinschätzung durch Therapeuten und Patienten

Es wurde in den Nachuntersuchungen erfragt³:

„Im Vergleich zum Beginn der Therapie in der CDS fühle ich mich:

1 = sehr viel besser, 2 = viel besser, 3 = etwas besser, 4 = unverändert, 5 = etwas schlechter, 6 = viel schlechter, 7 = sehr viel schlechter.“

In einem analogen Item wurden auch die Therapeuten nach ihrem Eindruck von der Besserung des Patienten befragt. Damit kommt auch eine Fremdbeurteilungsperspektive zum tragen. Die Nachuntersuchung wurde übrigens nicht immer von den behandelnden Therapeuten durchgeführt (z.B. den ein Therapeut vor der 1-Jahres-Katamnese die Behandlungseinrichtung verlassen hatte); die anderen Diagnostiker, die Nachuntersuchung durchführten, waren aber in

³ Dieses Item wurde sowohl im Nachuntersuchungsfragebogen als auch im Nachuntersuchungs-Interview gefragt. Da sich diesbezüglich bei einer Substichprobe, für die die Interviewdaten eingegeben waren, keine Unterschiede fanden, wurde im folgenden auf die Fragebogendaten zurückgegriffen (da diese im Gegensatz zu den Interviews immer eingegeben worden sind).

der Regel ebenfalls Therapeuten bzw. therapeutisch und diagnostisch geschult. Deswegen wird im Folgenden weiter einfach von „Therapeutenurteilen“ gesprochen.

2) Beeinträchtigungsebene

Vor Therapie und bei Nachuntersuchungen wurden für folgende Bereiche fünfstufige Items eingesetzt:

- Arbeit/Ausbildung
- Freizeit/soziale Kontakte (z.B. Zusammenkünfte mit Freunden, Feste, Ausgehen, Unterhaltung, Theater, Kino, Konzerte, Besuche, Vereine, Sport, Urlaub, etc.)
- Haushalt/Familienleben (z.B. Ihr Verhältnis zu anderen Familienmitgliedern, Hausarbeiten, Rechnungen bezahlen, Einkaufen, Besorgungen, Säubern, etc.)

„Schätzen Sie Ihre (*bei Nachuntersuchung: immer noch bestehende*) Beeinträchtigung (durch die Probleme, derentwegen Sie die Behandlung in unserer Einrichtung aufgesucht haben) im Bereich ... auf der folgenden Skala ein:

0 = gar nicht, 1 = ein wenig, 3 = mäßig, 4 = schwer, 5 = sehr schwer/massiv beeinträchtigt.“

3) Symptomebene: störungsspezifisch und störungsübergreifend

Im Rahmen dieser Dissertation werden bezüglich *störungsspezifischer* Symptome lediglich die Ergebnisse der folgenden drei Instrumente berichtet, da diese sowohl körperliche Symptome als auch Vermeidungsverhalten breit erfassen und als gut etabliert gelten können (d.h. überprüfte hinreichend gute Reliabilitäts- und Validitätswerte aufweisen und international häufig eingesetzt werden; vgl. Fydrich et al., 1996).

Weitere eingesetzte Instrumente (z.B. Agoraphobic Cognitions Questionnaire, Ehlers, Margraf & Chambless, 1993; State-Trait Inventory, Spielberger, 1972; Hamburger Zwangsinventar, Klepsch, Zaworka, Hand, Lünnenschloß & Jauernig, 1993) werden hier der Übersichtlichkeit halber nicht referiert, wurden aber im Rahmen einer Diplomarbeit (Sekot, 1999) gleichwohl auch ausgewertet.

- **BAI: Beck Angstinventar** (Beck et al., 1988; deutsche Version von Margraf & Ehlers, 1995)

Mit dem *Beck Anxiety Inventory* kann die Schwere klinisch relevanter Angst bei Erwachsenen erfasst werden. Das Instrument ist mit dem Ziel konstruiert worden, Ängstlichkeit möglichst unabhängig von depressiven Symptomen zu messen. Es enthält Items zur Klassifikation von Panikanfällen und von generalisierter Angst. Die 21 Items werden auf einer 4-Punkte-Skala dahingehend beurteilt, wie sehr sich die Patienten im Laufe der vergangenen sieben Tage durch das jeweilig angegebene Symptom belastet fühlten. Die Auswertung erfolgt über den Summenwert.

- **BSQ: Body Sensations Questionnaire** (Ehlers et al., 1993)

Der *BSQ* (Ehlers, Margraf, Chambless, 1993) misst die Schwere der Angst vor körperlichen Symptomen. 17 fünfstufige Items erlauben die Einschätzung der Angst vor den häufigsten körperlichen Angstsymptomen, die besonders oft bei Angstanfällen erlebt werden. Ein freies Item erlaubt den Betreffenden das zusätzliche Einfügen eines anderen Körpersymptoms, vor welchem sie Angst haben.

- **MI: Mobilitätsinventar für phobisches Verhalten** (Chambless et al., 1985; deutsche Version von Ehlers et al., 1993)

Das *Mobility Inventory* dient der Erfassung der Ausprägung agoraphobischen Vermeidungsverhaltens bei Patienten mit Paniksyndrom und/oder Agoraphobie oder mit somatoformen Störungen. Der Fragebogen besteht aus 28 Items (plus Zusatz-Items für Situationen, die Patienten individuell angeben können), die sich einmal auf die Bedingung „Vermeidung allein“ (MIa) und zum anderen auf „Vermeidung in Begleitung“ (MIb) beziehen. Auch hier sollen die Patienten auf einer 5-Punkte-Skala beurteilen, wie oft die angegebenen Situationen vermieden werden. Die Auswertung erfolgt über die jeweiligen Summenwerte geteilt durch Anzahl der ausgefüllten Items. Die folgenden Auswertungen beschränken sich auf den relevanteren Teil der Vermeidung allein, also auf den MI-A.

Bei der Erhebung mit *störungsübergreifenden* Instrumente bzw. zur Komorbidität wurde sich ebenfalls auf zwei besonders etablierte Maße beschränkt:

- **SCL-90-R: Symptomcheckliste** (revidierte Version; Derogatis, 1983; Franke, 1995)

Diese Symptomcheckliste dient der Überprüfung des Ausmaßes der psychischen Belastung. Sie besteht aus 90 Items, die in folgende Subskalen eingehen: Somatisierung, Zwanghaftigkeit, Unsicherheit im Sozialkontakt, Depressivität, Ängstlichkeit, Aggressivität/Feindseligkeit, phobische Angst, paranoides Denken und Psychotizismus. Auf einer 5-Punkte-Skala geben die Patienten an, wie ausgeprägt die Symptome sind. Anhand von drei globalen Indizes können Rückschlüsse auf das Antwortverhalten der Patienten gezogen werden. Der *Global Severity Index* (GS), auf den auch in den folgenden Berechnungen zurückgegriffen wird, gibt Auskunft über die grundsätzliche psychische Belastung. Der *Positive Symptom Distress Index* (PSDI) misst die Intensität der Antworten, und der *Positive Symptom Total* (PST) kennzeichnet die Anzahl der Symptome, bei denen ein Leidensdruck vorliegt. Es liegen auch Normtabellen bezüglich Bildungsgrad, Alter und Geschlecht von einer normalen gesunden Stichprobe (n=1006) vor. Das Instrument eignet sich weniger zu einer Differentialdiagnostik als zu einer Feststellung der allgemeinen, summativen psychischen Belastung (Schulte, 1993).

- **BDI: Beck Depressionsinventar** (Beck et al., 1961, deutsche Version von Hautzinger et al., 1993)

Das Beck Depressionsinventar dient der Erfassung der Schwere depressiver Symptome. Mit diesem Fragebogen können Erwachsene zwischen 18 und 80 Jahren untersucht werden. Der Fragebogen enthält 21 Items mit jeweils vier Aussagen, in denen depressive Symptome in aufsteigender Schwere und zunehmender Beeinträchtigung beschrieben sind. Die Patienten sollen aus jeder Gruppe die Aussage ankreuzen, die ihre gegenwärtige Situation am besten beschreibt. Auswertung erfolgt über den Summenwert.

3.2.3 Effektivitätsberechnung und Kriterien klinischer Signifikanz

1) globales Item

Hier wurden zunächst die Häufigkeiten der gegebenen Antworten (bzw. deren %-Sätze) für Patienten- und Therapeuten-Urteil grafisch dargestellt.

Für die Effektstärken-Berechnung kann der Grad der Veränderung, also die Einschätzung, ob es dem Patienten nach der Therapie besser, gleich oder schlechter geht, inhaltlich als Differenzwert betrachtet werden. Allerdings müssen diese Einschätzungen, die auf einer siebenstufigen Skala erfolgten, noch folgendermaßen umkodiert werden:

Kategorie	alter Wert	neuer Wert
„sehr viel besser“	1	3
„viel besser“	2	2
„etwas besser“	3	1
„unverändert“	4	0
„etwas schlechter“	5	-1
„viel schlechter“	6	-2
„sehr viel schlechter“	7	-3

Anhand der umkodierten Werte werden neue Mittelwerte und Standardabweichungen berechnet, mit denen dann die Effektstärke berechnet werden kann. Dabei entspricht der für die Patienteneinschätzung errechnete Mittelwert bereits der Mittelwertsdifferenz im Zähler der Grundformel:

$$ES = \frac{M_a - M_b}{s}$$

Grundformel für die Effektstärkenberechnung nach Smith, Glass & Miller (1980)

Bei der Streuung im Nenner handelt es sich allerdings um eine Streuung von Veränderungswerten und nicht, wie eigentlich bei Effektstärke-Berechnungen üblich, um eine prä-, eine post- oder um eine gepoolte Streuung.

Aus diesen Gründen – keine echte Veränderungsmessung, keine übliche Streuung – stellt der Wert für das globale Item keine „echte“ Effektstärke dar und kann somit nur mit Effektstärken eines solchen Items (z.B. berechnet in einer anderen Stichprobe), nicht aber mit üblichen Effektstärken (z.B. mit den auf den beiden anderen Ebenen berechneten) verglichen werden!

Als „Klinisch signifikante Verbesserung“ wurden die Urteile „sehr viel besser“ und „viel besser“ definiert; erstens weil dies unmittelbar plausibel erscheint und zweitens, weil andere Kriterien wie die Einordnung in Normen bei diesem Einzelitem nicht berechnet werden können.

2) Beeinträchtigungs-Items

Für die Beeinträchtigungs-Items wurde über die Differenzen prä-post zunächst die Effektstärke berechnet. Dabei wird die von Hartmann, Herzog & Drinkmann (1992) empfohlene Formel verwendet, welche die Streuungen einer

Behandlungsgruppe zum Zeitpunkt t_1 und t_2 bei Gewichtung der jeweiligen Stichprobengrößen poolt:

$$ES = \frac{M_{t_1} - M_{t_2}}{\sqrt{\frac{(N_{t_1} - 1)s_1^2 + (N_{t_2} - 1)s_2^2}{N_{t_1} + N_{t_2} - 2}}}$$

Prä-Post-Effektstärke nach Hartmann, Herzog & Drinkmann (1992), t_1 ist die Prä-Messung, t_2 die Post-Messung.

Bei dieser Formel wird die Verringerung der Varianz bei großen Mittelwertsunterschieden (durch Deckeneffekte, wenn sich viele Patienten in Richtung „Normalität“ verbessert haben) belohnt, was inhaltlich angemessener erscheint als die Varianzveränderung von prä nach post nicht zu berücksichtigen (vgl. 2.2.2.1).

Ferner werden neben dieser standardisierten durchschnittlichen Veränderung individuelle Veränderungen betrachtet. In Kosten-Effektivitäts-Matrizen kann abgelesen werden, wie viele Patienten sich jeweils wie verändert haben (s. Ergebnisdarstellung).

Doch was soll nun als „klinisch relevante Verbesserung“ bzw. „erfolgreiche Behandlung“ definiert werden? Da hierfür keine Normwerte zur Verfügung stehen, wird – ähnlich wie bei der globalen Einschätzung – inhaltlich entschieden: klinisch relevant sei eine Beeinträchtigung, die als „schwer“ oder „sehr schwer“ eingeschätzt wird, subklinisch sei eine „mäßige“ und unauffällig „gar keine oder „ein wenig“ Beeinträchtigung. In Anlehnung an die in Abb. 2.1 (S. 25) beschriebenen Darstellung klinisch signifikanter Veränderungen werden die Daten wie in Tab. 3.1 dargestellt umcodiert. Die Überschreitung einer Klassengrenze (Kästchen außer A, E, I) wird dann als klinisch signifikante Veränderung aufgefasst.

Tabelle 3.1: Schema zur Bestimmung klinisch signifikanter Veränderung (Erläuterungen im Text)

Post-Messung Prä-Messung	unauffällig	subklinisch	klinisch relevante Belastung	Summe (Prä)
unauffällig	A	F	G	
subklinisch	B	E	H	
klinisch relevante Belastung	C	D	I	
Summe (Post)				

Die Patienten, die nach der Therapie in die Kategorie „klinisch relevant“ fallen, egal, ob sie vorher „unauffällig“ oder bereits „klinisch relevante“ Symptomatik zeigten, werden als „nicht erfolgreich behandelt“ definiert (Kästchen G-I). Ebenfalls dort zugerechnet werden die Patienten, die vorher gesund sind und nachher eine subklinische Symptomatik aufweisen bzw. subklinisch verbleiben (Kästchen E und F). Besonders kritisch müssen dabei die Verschlechterungen in Kästchen F bis H betrachtet werden.

Dementsprechend sei der Anteil „erfolgreicher Behandlungen“ die Summe der Kästchen A bis D. Die Patienten, die vor *und* nach der Therapie im jeweiligen Bereich unauffällig waren (Kästchen A) werden bei der Bestimmung klinisch signifikanter Veränderung allerdings nicht weiter berücksichtigt, da sie sich durch die Therapie gar nicht mehr verbessern konnten.

3) Symptomebene: störungsspezifische und –unspezifische Fragebogenmaße

Mittelwertsvergleiche wurden mit t-Tests für abhängige Stichproben (bei Gruppenvergleichen für unabhängige Stichproben) berechnet. Allgemein kann man nach Bortz (1985) von einer relativen Robustheit auch bei Verletzung seiner Voraussetzungen (Normalverteilung, Varianzhomogenität) ausgehen. Für den praktischen Einsatz entwickelten Havlicek & Petersen (1974) Richtlinien, denen entsprechend hier in Ausnahmefällen auf den Wilcoxon-Test (abhängig) bzw. den Mann-Whitney-U- (unabhängig) zurückgegriffen wurde. Solche Ausnahmen liegen etwa vor, wenn beide verglichenen Verteilungen gegenläufig schiefverteilt sind oder wenn die Varianzen bei deutlich unglei-

chen Gruppengrößen und normalen oder gleich schiefen Verteilungen nicht homogen waren. Alpha-Adjustierungen mussten aufgrund der Unabhängigkeit der Messfaktoren – bei den verglichenen Variablen handelt es sich nicht um unterschiedliche Treatmentstufen, nicht vorgenommen werden (vgl. Cook & Campbell, 1979).

Die Effektstärke wird analog zur Beeinträchtigungsebene mit der Formel nach Hartmann, Herzog und Drinkmann (1992) berechnet.

Über die statistisch signifikanten Verbesserungen hinaus soll auch hier das Ergebnis auf klinische Signifikanz untersucht werden. Dazu müssen die Patienten prä, post und zur 1-Jahres-Katamnese anhand der Fragebogenwerte in die Kategorien unauffällig, subklinisch und klinisch relevante Belastung eingeordnet werden. Man benötigt für diese Berechnungen Mittelwert und Standardabweichungen gesunder und dysfunktionaler Populationen. Direkte Angaben von den Autoren über entsprechende cut-off-Werte auch für einen subklinischen Bereich lagen nur für das Beck-Depressions-Inventar (BDI; Hautzinger et al., 1993) vor.

Für die anderen Fragebögen werden zunächst die Werte berechnet, für die die Wahrscheinlichkeit, zur unauffälligen oder zur kranken oder dysfunktionalen Gruppe zu gehören 50 % beträgt („Punkt c“; Tab. 3.2 auf folgender Seite; vgl. 2.2.2.2). Für das Beck-Angst-Inventar (BAI), den Body Sensations Questionnaire (BSQ) und das Mobilitätsinventar (MI-A) konnte für die dysfunktionalen Normen auf die Angaben der Autoren zurückgegriffen werden (Beck et al., 1985; Ehlers et al., 1993). Die gesunden Normen wurden hier anhand einer (mit strukturiertem Interview getesteten) gesunden Stichprobe ermittelt⁴.

Nun musste eine Entscheidung darüber getroffen werden, welcher Bereich sinnvoll als „subklinisch“ bezeichnet werden kann. Inhaltlich soll damit analog zum vorangehenden Abschnitt eine für den klinischen Alltag sinnvolle Differenzierung dahingehend geschaffen werden, dass Patienten, die sich nach Therapie deutlich gebessert fühlen, aber immer noch (allerdings nicht mehr behandlungsbedürftige) Symptome aufweisen, nicht aufgrund ihrer Fragebogenwerte als Non-Responder betrachtet werden. Zudem gibt es Patienten, die sich auch unterhalb eindeutig dysfunktionaler Fragebogenwerte stark belastet und beeinträchtigt fühlen und für die ein Übergang in einen eindeutig unauffälligen Bereich subjektiv eine bedeutsame Verbesserung bedeutet. Schließlich ist auch im Rahmen der zunehmend an Bedeutung gewinnenden Frage der Verschlechterungen im Zuge einer Therapie (oder gar Therapieschäden) eine differenziertere Sichtweise zu wünschen: Übergänge vom eindeutig unauffälligen in einen subklinischen oder vom subklinischen in einen eindeutig dysfunktionalen Bereich können bei einer dichotomen Sichtweise „gesund vs. krank“ nicht genauer analysiert werden.

⁴ Vielen Dank an Frau Dr. Monika Frank und Frau Dr. Lydia Fehm für die Bereitstellung dieser Stichprobe.

Es wurde folgendes Vorgehen gewählt: als „eindeutig unauffällig“ wurde der Bereich vom Mittelwert der unauffälligen Gruppe bis zu einer halben Standardabweichung darüber betrachtet. Als „eindeutig dysfunktional“ wurde der Bereich ab einer halben Standardabweichung unterhalb des dysfunktionalen Mittelwertes bestimmt. „Subklinisch“ war damit der Bereich zwischen $M_{\text{unauffällig}} + 0.5s_{\text{unauffällig}}$ bis $M_{\text{dysfunktional}} - 0.5s_{\text{dysfunktional}}$. Vorschläge, die im Rahmen der Diskussion um die klinische Signifikanz gemacht wurden, „eindeutig unauffällig“ bzw. „eindeutig dysfunktional“ als Bereich von ein oder gar zwei Standardabweichungen um den jeweiligen Mittelwert zu definieren (Jacobson & Truax, 1991; Hollon & Flick, 1988; Wampold & Jensen, 1986; vgl. 2.2.2.2), wurden verworfen, da diese Bereich sich anhand der Normwerte oder auch der Daten als viel zu groß erweisen; oft überlappen sich bei dieser Maßgabe die Bereiche sogar.

Die Wahl der halben Standardabweichungen hat also den Vorteil, dass sie auf praktikable Weise die Bereiche „unauffällig“ und „dysfunktional“ eindeutig trennt und dazwischen den gewünschten subklinischen Bereich schafft. In Tabelle 3.2 sind die ermittelten cut-off-Werte aufgeführt.

Da die Operationalisierung der Kategorien „klinisch auffällig“, „subklinisch“ und „unauffällig“ im Rahmen dieser Kosten-Effektivitäts-Analyse zentral ist, wird sie in Kasten 3.2 nochmals zusammengefasst.

Kasten 3.2: Operationalisierung „klinisch relevanter Besserung“ bzw. der Klassengrenzen zwischen „klinisch auffällig“, „subklinisch“ und „unauffällig“ in der vorliegenden Studie

Globale Beurteilung:

„Klinisch relevante Besserung“ = „viel/sehr viel besser“

Beeinträchtigung:

„klinisch auffällig“ = „schwer/sehr schwer beeinträchtigt“

„subklinisch“ = „mäßig beeinträchtigt“

„unauffällig“ = „gar nicht/ein wenig beeinträchtigt“

Fragebogenmaße:

BDI: von Autoren vorgegebene cut-offs

„klinisch auffällig“ = Wert > 18

„subklinisch“ = Wert > 11 und < 18

„unauffällig“ = Wert < 11

Forts. Kasten 3.2

andere Fragebögen, für die keine Normen vorlagen:

„klinisch auffällig“ = Wert $> 1/2$ Standardabweichung unterhalb klinischer Population

„subklinisch“ = Wert $> 1/2$ Standardabweichung oberhalb gesunder Population und $< 0,5$ Standardabweichungen unterhalb klinischer Population

„unauffällig“ = Wert $< 1/2$ Standardabweichung oberhalb gesunder Population

aus dieser Definition ermittelte Werte für BAI, BSQ, SCL, MI-A finden sich in Tabelle 3.2

Tabelle 3.2: Punkte der gleichen Wahrscheinlichkeit für unauffällig und dysfunktional und die Grenzen für eine subklinische Kategorie

	\bar{X}_{gesund}	S_{gesund}	$\bar{X}_{\text{dysfunktional}}$	$S_{\text{dysfunktional}}$	C	$\bar{X}_{\text{gesund}} + 0,5S_{\text{gesund}}$ (untere Grenze subklinisch)	$\bar{X}_{\text{dysfunktional}} - 0,5S_{\text{dysf.}}$ (untere Grenze dysfunktional)
BAI	4,24 (N=134)	4,75	23,5 (N: k. A.)	12,36 (k.A.von Autoren; deswegen unserer Stichprobe)	9,59	6,62	17,32
BSQ	1,55 (N=133)	0,57	2,6 (N: k. A.)	0,72	2,01	1,84	2,24
MI-A	1,33 (N=97)	0,41	2,7 (N: k.A)	1,04	1,72	1,54	2,18
SCL	29,7 (N=1006)	22,5	116,1 (N=285)	61,2	52,92	40,95	85,5
BDI	unauffällig: <10					subklinisch: 11-17	dysfunktional >= 18

3.2.4 Darstellung der Kosten-Effektivität

Die Kosten-Effektivität wird nun auf zweierlei Weise dargestellt:

- „Nebeneinanderstellen“ verschiedener Veränderungsstufen und der zugehörigen Kosten (in Tabellen oder Kosten-Effektivitäts-Matrizen), um einen vorläufigen Eindruck zu bekommen.

Diese eher „impressionistische“ Betrachtungsweise wird ergänzt durch den

- *Kosten-Effektivitäts-Index*, bei dem die Kosten durch die jeweilige Rate „erfolgreich behandelter“ Patienten geteilt werden. Dieser Index berücksichtigt also auch die nicht erfolgreichen Therapien und ist damit aussagekräftiger bezüglich der Kosten eines Programmes oder der Behandlungen in einer Einrichtung als die Durchschnittskosten einer einzelnen als erfolgreich erachteten Therapie (vgl. 2.2.1).

Der in dieser Arbeit verwendete Index definiert eine „erfolgreiche Behandlung“ als klinisch relevante Verbesserung, wie sie für die globale Einschätzung, die Beeinträchtigungen und die Fragebogenmaße beschrieben wurden. Dabei sind auch Verbesserungen von klinisch relevanter Beeinträchtigung in einen subklinischen Status einbezogen; außerdem wird dabei die Berechnung von dem Anteil der Patienten bereinigt, der sich im jeweiligen Maß bereits vorher auf einem unauffälligen Niveau befindet. Die hier verwendete Formel lautet also:

$$\text{Kosteneffektivität} = \frac{\text{mittlere Therapiekosten aller Patienten (DM)}}{\text{Rate „klinisch relevant gebesserter“ Patienten}}$$

3.3 Rahmenbedingungen und Stichprobe

3.3.1 Therapieform, Therapeuten, klinische Diagnostik

Für eine Beschreibung der Christoph-Dornier-Stiftung (CDS) sei auf die Darstellung der Christoph-Dornier-Stiftung im Internet hinweisen. (<http://www.christoph-dornier-stiftung.de/>). Weitere Informationen finden sich in Fiegenbaum und Fünfstück (1997). Dass die Therapien unter klinisch relevanten Bedingungen stattfanden, wurde bereits unter 2.2.1 ausgeführt.

Alle Therapierenden sind Diplompsychologen mit verhaltenstherapeutischer Ausbildung, die die Zulassung des zuständigen Gesundheitsamtes haben, eigenverantwortlich Heilkunde im Bereich der Psychotherapie ausüben. Sie bilden sich an anerkannten Ausbildungsinstituten in Verhaltenstherapie weiter oder nehmen nach abgeschlossener Weiterbildung regelmäßig an verhaltenstherapeutischen Fortbildungsveranstaltungen teil.

Die Therapierenden sind in der Regel gleichzeitig Promotionsstipendiaten der Stiftung. Regelmäßig werden die jeweils aktuellen Fälle von den Institutsleiterinnen, den Supervisoren der jeweiligen Weiterbildungsgänge oder vom Vorstand der CDS supervidiert. Die Behandlungsmethode kann als kognitive Verhaltenstherapie bezeichnet werden (störungsspezifisch; vgl. entsprechende Kapitel in Margraf, 2000). Ein zentrales Behandlungselement bei der Therapie von Angststörungen ist dabei die Konfrontation mit angstauslösenden Reizen bei gleichzeitiger Verhinderung üblichen Sicherheits- und Vermeidungsverhaltens. Ergänzend eingesetzt werden störungsspezifische kognitive Strategien (Schneider & Margraf, 1998).

Zur Diagnose wird das „Diagnostische Interview bei psychischen Störungen“ durchgeführt (DIPS; Margraf, Schneider & Ehlers, 1991). Es handelt sich um ein strukturiertes Interview, welches die Anwendung von Algorithmen der internationalen Klassifikationssysteme ICD oder DSM unterstützt, indem es die wesentlichen Symptome für die Erstellung der Diagnose berücksichtigt. Zudem wird darin bereits therapierelevante Information erhoben. Im Rahmen der vorliegenden Arbeit werden Patienten untersucht, die (mindestens) eine Angststörung nach DSM-III-r bzw. DSM IV aufweisen.

Neben dem diagnostischen Interview werden zur Diagnostik, sechs Wochen nach Therapieende (Post-Zeitpunkt) sowie in 1-Jahres-Katamnesen umfangreiche Fragebogendaten erhoben.

3.3.2 Stichprobe

Aus den Ambulanzen der Stiftung in Dresden, Braunschweig, Marburg und Münster wurden die bis Mitte 1998 gesammelten Daten über alle Behandelten zusammengetragen und ausgewertet. Der Stichtag, an dem keine Daten mehr in den Datensatz aufgenommen wurden, war der 20. Juni 1998. Der endgültige Datensatz umfasste 1490 Personen (157 aus Braunschweig, 418 aus der Ambulanz in Dresden, 681 aus Marburg und 14 aus dem ambulanten Institut in Münster⁵).

Da die Christoph-Dornier-Stiftung hauptsächlich auf den Bereich der Therapie von Ängsten und Zwängen spezialisiert ist und die sonstigen Behandlungsschwerpunkte zwischen den einzelnen Instituten variieren, beschränkt sich die Stichprobe auf die Personen, die in allen Instituten mit ähnlichen Methoden behandelt werden. Es wurden daher nur Personen in die Stichprobe aufgenommen, die eine der Primärdiagnosen aus dem Angst- und Zwangsbe- reich nach DSM-III-r bzw. DSM-IV erhielten⁶. Von der Stichprobe ausge- schlossen wurden Patienten ohne Angststörung als Primärdiagnose und Patien- ten, für die keine Nachuntersuchungsergebnisse vorlagen (definiert über feh- lenden Wert im Fragebogen SCL-90-R zur Postmessung). Dies waren vor al- lem Patienten, die zu einem frühen Zeitpunkt aus Kostengründen oder aus Zweifel am Therapiekonzept die Therapie abbrachen. So ergab sich eine Zahl von 493 behandelten Personen (67 aus Braunschweig, 117 aus Dresden, 295 aus Marburg und 14 aus Münster).

3.3.2.1 Soziodemographische Variablen

Zur Stichprobenbeschreibung werden in folgender Tabelle 3.3 die Verteilung in den Variablen Alter, Geschlecht, Bildungsstand und ausgeübter Beruf dar- gestellt.

Alter- und Geschlechtsverteilung sind für Psychotherapie- Patientenpopulationen typisch (vgl. Linden et al., 1993; Vessey & Howard, 1993). Akademische Berufe sind im Vergleich zu Linden et al. (1993), die

⁵ Aufgrund eines Computerschadens in Münster waren dort alle eingegebenen Daten verlorenegan- gen und konnten nicht mehr in absehbarer Zeit neu eingegeben werden

⁶ Die störungsspezifischen Therapien der verschiedenen Angst- und Zwangsstörungen weisen struk- turelle Ähnlichkeiten auf (vgl. Margraf, 1996), die es gerechtfertigt erscheinen lassen, diese gemein- sam zu untersuchen. Nicht mit aufgenommen wurde allerdings die im DSM-, nicht aber im ICD- System zu den Angststörungen gerechnete Posttraumatische Belastungsstörung, da deren Therapie aufgrund der in der Regel diffusen Hyperarousal- und Dissoziationssymptome (vgl. Maercker, 1997) und vor allem wegen des grundsätzlich unterschiedlichen empirischen Realitätsgehalts der Befürch- tungen eine andere Strategie verfolgt (vgl. Ehlers & Clark, 2000). Für die Diagnose der Generalisier- ten Angststörung lagen nur fünf Datensätze vor, so dass diese Kategorie (für die ebenfalls von den phobischen Störungen verschiedenen therapeutische Ansätze empfohlen werden; Turowsky & Bar- low, 1996) ebenfalls nicht in die Auswertungen einging.

eine große Studie zur Epidemiologie von Verhaltenstherapie-Antragstellern durchführten, allerdings überrepräsentiert. Dies mag mit dem damaligen rechtlichen Status der Behandlungseinrichtungen (Kostenerstattungsverfahren) zusammenhängen, das den Patienten gegenüber dem Delegationsverfahren mehr Initiative und bisweilen eine Zuzahlung abverlangte.

Tabelle 3.3: Soziodemographische Stichprobendaten (N=493) zum Zeitpunkt vor Therapie

Alter	$\bar{x} = 35,7$	$s = 9,4$; Quartile = 29/35/42
Geschlecht	männlich:	38 %
	weiblich:	62 %
höchster Schulabschluss	keiner:	1,2 %
	Hauptschule:	18,3 %
	Realschule:	28,4 %
	Fachabitur/Abitur:	26,9 %
	abgeschlossenes Fach-/Hochschulstudium:	25,2 %
derzeitige Berufsgruppe	Arbeiter:	3,1 %
	Facharbeiter/Handwerker:	5,5 %
	Angestellter/Beamter	
	einfacher/mittlerer Dienst:	29,3 %
	gehobener Dienst:	14,4 %
	Selbständiger:	12,8 %
	Schüler/Azubi/Student:	14,9 %
	Hausfrau:	13,3 %
Familienstand	arbeitslos:	4,4 %
	berentet:	2,2 %
	verheiratet	51,2 %
	ledig	37,8 %
	geschieden/getrennt	8,4 %
	verwitwet	0,7 %
	k.A.	1,9 %

3.3.2.2 Störungsspezifische Variablen

In der folgenden Tabelle 3.4 wird ein Überblick über die folgenden störungsspezifischen Variablen gegeben: Primärdiagnose, Anzahl weiterer Diagnosen, Krankheitsdauer, Gesamtbelastung durch Symptome (als Indikator diente der SCL-90-R-Gesamtwert), Einschränkung der Mobilität (als Indikator diente der Wert im Mobilitäts-Inventar MI-A, der die Vermeidung von Situationen ohne Begleitung erfasst).

Tabelle 3.4: Störungsspezifische Stichprobendaten (N=493) zum Zeitpunkt vor Therapie

Verteilung der Primär Diagnosen (Angststörungen; N=493)	Panikstörung (ohne Agoraphobie):	5 %
	Panikstörung mit Agoraphobie	69 %
	Soziale Phobie:	18 %
	Spezifische Phobie:	2 %
	Zwangsstörung:	6 %
Anzahl weiterer Diagnosen: 192	davon	
	weitere Angststörung: affektive Störung:	50 %
	somatoforme Störung:	30 %
	sonst. Psych. Störung:	12,5 %
		7,5 %
Krankheitsdauer	Mittelwert:	9,1 Jahre
	Median:	6 Jahre
	Quartile:	3/6/13 Jahre
Gesamtbelastung durch Symptome (SCL-90-R Gesamtwert, Prä-Zeitpunkt)	1. Roh-Werte	
	Mittelwert:	97
	Standardabweichung:	54
	Quartile:	55/90/129
	2. T-Werte	
	Mittelwert:	68
	Standardabweichung:	10,7
Quartile:	61/71/78	
Einschränkung Mobilität / Vermeidung (MI-A, Prä-Zeitpunkt)	Mittelwert:	2,9
	Standardabweichung:	1,1
	Quartile:	2,0/2,8/3,8

Der Schwerpunkt bei den Diagnosen liegt deutlich bei der Agoraphobie mit Panikstörung. Die diagnostische Kategorisierung wird durch hohe Werte in den entsprechenden Unterskalen der SCL-90-R bestätigt (Sekot, 1999). Auch der Gesamtwert in der SCL-90-R liegt im von den Autoren angegebenen Bereich für „hohe bis sehr hohe Belastung“ und liegt im Bereich der dort angegebenen Vergleichsgruppe mit „neurotischen Störungen“ (Franke, 1995).

Zusatz- bzw. Nebendiagnosen wurden auffallend wenige vergeben (vgl. z.B. Wittchen, 1991, der nur bei 14% der Panikpatienten keine Komorbidität fand). Dies hat unserer Einschätzung nach zwei Ursachen. Erstens sind die Behandlungseinrichtungen für die Behandlung von Agoraphobien bekannt, so dass sich – insbesondere im Untersuchungszeitraum Mitte der 90er Jahre – vor allem Patienten mit eindeutigem Krankheitsbild dort vorstellten. Wichtiger aber noch für die niedrige identifizierte Komorbiditätsrate scheint das Diagnoseverhalten der Therapeuten zu sein: der Fokus lag in der Regel auf der therapie- und kassenantragsrelevanten Primärdiagnose. Dies entspricht der Annah-

me bei Wittchen und Vossen (2000), dass klinische Diagnosen implizit stärker auf den Querschnittsbefund bezogen sind und so oft bereits weiter zurückliegende oder durch das aktuelle Beschwerdebild überdeckte Störungen vernachlässigt werden. Zudem wurden Achse-II-Störungen gar nicht diagnostiziert. Es kann also davon ausgegangen werden, dass in einer exakten epidemiologischen Studie in dieser Patientenpopulation mehr Komorbidität gefunden worden wäre.

Dafür spricht auch der Schweregrad der Störungen, wie er sich in den Fragebogeninstrumenten Symptomcheckliste und Mobilitäts-Inventar niederschlägt. Der Mittelwert im MI-A befindet sich über dem von den Autoren angegebenen Wert für Angstpatienten ($M=2,7$; Ehlers et al., 1993) und fast vier Standardabweichungen über dem Wert einer zur Vergleichsgründen in der CDS untersuchten gesunden Population ($M=1,3$; $s=0,41$)!

Die vorliegende Studie baut also auf eine für derartige Behandlungseinrichtungen durchaus typische schwer belastete und beeinträchtigte, für viele Jahre erkrankte, typischerweise komorbide (auch wenn ungenügend erhoben, s.o.) Patientenpopulation auf, die als charakteristisch für schwer erkrankte, behandlungsbedürftige Patienten dieser Diagnosegruppe gelten kann. Insofern ergeben sich bezüglich der späteren Ergebnislage keine erkennbaren Einschränkungen bezüglich der Aussagekraft hinsichtlich der Repräsentativität der Stichprobe (vgl. auch 2.1.2, 2.2).

3.4 Kosteneffektivitätsstudie: Ergebnisse

3.4.1 Kosten

Die Therapien wurden entweder im damaligen Kostenerstattungsverfahren über die Krankenkassen (damals sog. „TK-Regelung“) oder privat abgerechnet. In beiden Fällen wurden für alle Behandlungseinheiten (BE) à 50 Minuten die selben Tarife veranschlagt; dabei gab es keine Unterschiede bezüglich Erstgespräch/probatorischen Sitzungen, diagnostischer oder therapeutischer Sitzungen.

Die Patienten verteilen sich wie folgt auf die Tarife: 16,1% der Patienten wurden zum Osttarif von 111,70 DM behandelt, 5,9% der Patienten zum Osttarif von 117,30 DM (der Preis wurde 1997 angehoben), und bei 74,2% wurde pro Behandlungseinheit der Westtarif von 138,- DM berechnet.

Von N=25 Patienten (3,8%) lagen keine oder nur ungenaue Angaben über die Anzahl der BEs vor, so dass diese Werte zur Vermeidung von Verzerrungen nicht für die Berechnungen verwendet, sondern als Missingwerte betrachtet werden.

Die Patienten kontaktierten die Therapeuten im Durchschnitt 11,4 mal. Die Verteilung der Kontakte streut jedoch beträchtlich: von einem Sitzungstag bis zu maximal 110 Sitzungstagen. Da nicht bei jedem Patienten jedes einzelne Therapiedatum dokumentiert ist, kann anhand der Anzahl der Kontakte allein jedoch keine Aussage dazu getroffen werden, ob es sich bei den Therapien um Block- oder kontinuierliche Behandlungen handelt.

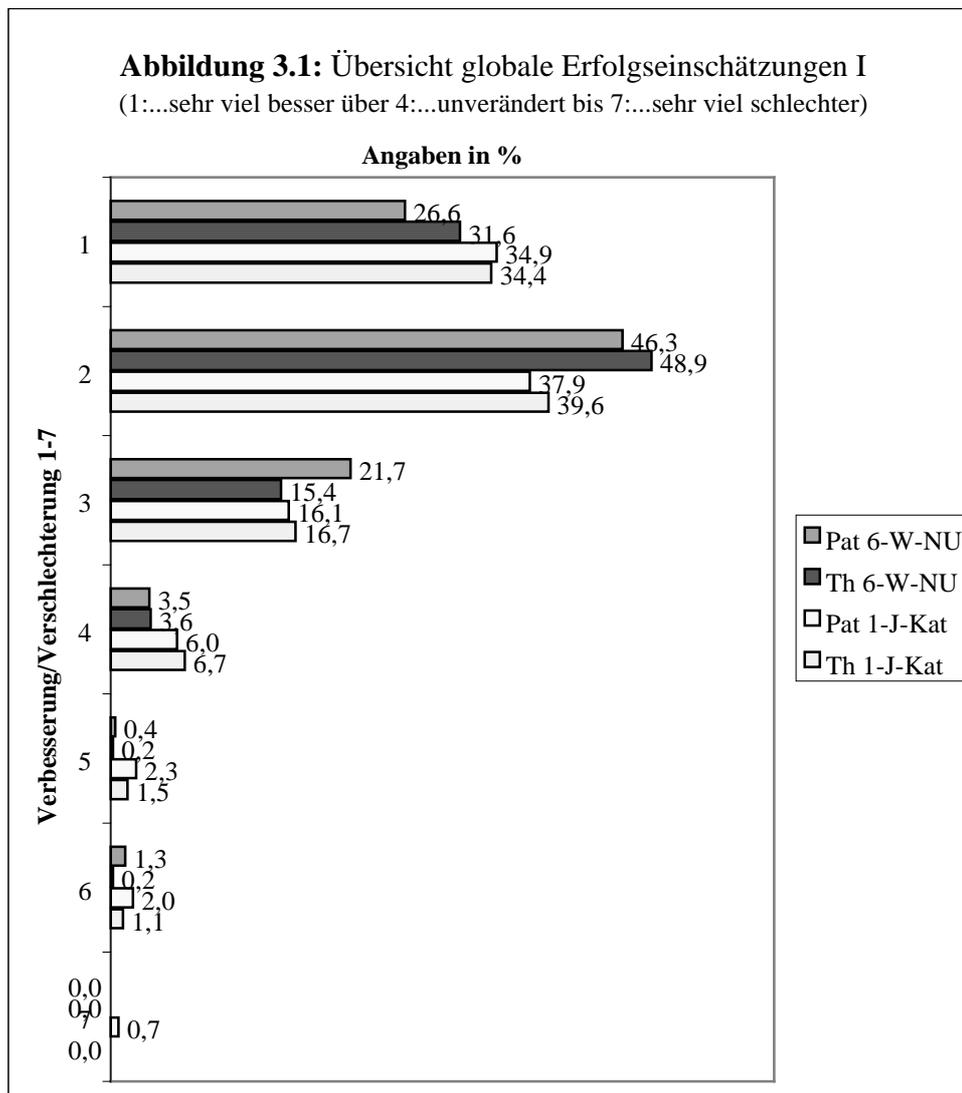
Die durchschnittliche Therapie eines Patienten in dieser Stichprobe umfasste 49 BEs, wobei auch hier eine breit streuende Verteilung zu finden ist ($s=22,83$). Das Minimum liegt bei 10,9 BEs, der Maximalwert wird mit 209,1 BEs erreicht. Solche Ausreißerwerte wurden auf ihre Richtigkeit geprüft und deshalb im Datensatz belassen.

Auf Grundlage der durchschnittlichen BEs und der verschiedenen Tarife wurden 6587,- DM als durchschnittliche Therapiekosten errechnet ($s=3001,-$ DM).

3.4.2 Effektivität (1) – globale Therapieerfolgseinschätzung durch Patienten und Therapeuten

3.4.2.1 Ausmaß der Effekte im Vergleich zu den Kosten

Im folgenden wird dargestellt, wie sich die globalen Erfolgseinschätzungen auf die einzelnen Kategorien zum Post-Zeitpunkt (sechs Wochen nach Therapieende) und zur 1-Jahres-Katamnese verteilen. Abbildung 3.1 sowie Tabelle 3.5 zeigen die entsprechenden Patienten- und Therapeuteneinschätzungen⁷.



⁷ Für die unterschiedlichen Gruppen waren unterschiedlich viele Urteile verfügbar: $N_{\text{Pat 6-W-NU}}=451$, $N_{\text{Th 6-W-NU}}=415$, $N_{\text{Pat 1-J-Kat}}=298$, $N_{\text{Th 1-J-Kat}}=255$. Mögliche systematische Verzerrungen durch fehlende Werte bzw. Dropouts werden in Abschnitt 3.5 untersucht; Hinweise auf Selektionseffekte fanden sich nicht.

Tabelle 3.5: Übersicht über signifikante Veränderungen auf der Basis der globalen Erfolgseinschätzungen

	\bar{x}	s	N	Vergleich Pat 6-W-NU (t-Test)	Vergleich Th 6-W-NU (t-Test)	Vergleich Pat 1-J-Kat (t-Test)	Vergleich Th 1-J-Kat (t-Test)
Pat. 6-W-NU	2,1	0,93	451		** (p=0,00)	n.s. (p=0,69)	n.s. (p=0,61)
Th. 6-W-NU	1,9	0,82	415	** (p=0,00)		* (p=0,02)	n.s. (p=0,11)
Pat. 1-J-Kat	2,1	1,2	298	n.s. (p=0,69)	* (p=0,02)		n.s. (p=0,10)
Th. 1-J-Kat	2,0	1,0	255	n.s. (p=0,61)	n.s. (p=0,11)	n.s. (p=0,10)	

Die mittleren Einschätzungen (Tab. 3.1) bleiben zwischen 6-Wochen-Nachuntersuchung und 1-Jahres-Katamnese stabil. Die signifikanten Unterschiede zwischen der Fremdbeurteilung durch die Therapeuten zum Post-Zeitpunkt und den Patientenurteilen resultieren nicht aus einer großen Mittelwertsdifferenz, sondern sind eher auf die Kombination niedriger Varianz und großer Stichprobengröße zurückzuführen. Der Median war für die Patienten wie für die Therapeuteneinschätzung mit jeweils 2.0 gleich.

Aufgrund der Gleichheit bzw. Stabilität der Urteile wird in der folgenden Übersicht in Tabelle 3.6 die Verteilung der Kosten auf die Therapieerfolgs-Kategorien lediglich für die Therapeuten-Urteile zur 1-Jahres-Katamnese bezogen und nicht für alle vier Einschätzungen einzeln aufgeführt. Die Therapeuten-Urteile wurden ausgewählt, damit neben den weiteren subjektiven Modalitäten der Therapieerfolgsmessung (Beeinträchtigung der Patienten und Fragebogenmaße) auch die Fremdbeurteilungsperspektive zum Tragen kommt. Die 1-Jahres-Katamnese wurde der Post-Messung vorgezogen, weil zur Beurteilung des Therapieerfolges und seiner Stabilität Katamnesedaten Post-Messungen überlegen sind.

Pfennigbeträge wurden aufgerundet und die ersten beiden Ziffern fett markiert, um Unterschiede hervorzuheben. Außerdem sind in der Tabelle die Verteilung der Kosten pro Besserungskategorie angegeben.

Tabelle 3.6: Verteilung der Kosten auf die Therapieerfolgseinschätzungen der Therapeuten in der 1-Jahres-Katamnese; die grau unterlegte Zeilen entsprechen den „klinisch relevant gebesserten“ Patienten (zusammen 74%)

Kategorie	%	Mittlere Kosten (DM)	Verteilung
1...sehr viel besser	34,4	6470,- (N=87)	
2...viel besser	39,6	6506,- (N=101)	
3...etwas besser	16,7	6574,- (N=43)	
4...unverändert	6,7	6808,- (N=17)	entfällt, da zu geringes N
5...etwas schlechter	1,5	8658,- (N=4)	entfällt, da zu geringes N
6...viel schlechter	1,1	7303,- (N=3)	entfällt, da zu geringes N
7...sehr viel schlechter	-	-	-

Legende zur Verteilung der Kosten (aufsteigend im Uhrzeigersinn):

	unter 3000,-
	3001,- bis 6000,-
	6001,- bis 9000,-
	9001,- bis 12000,-
	mehr als 12000,-

Es scheint zwar eine Tendenz dahingehend zu geben, dass durchschnittlich betrachtet die weniger erfolgreichen Therapien teurer waren: die durchschnittlichen Therapiekosten der 74% erfolgreich behandelten Patienten betragen DM 6488.- und die der nicht erfolgreichen 26% DM 7336.- (durchschnittliche Kosten über alle Patienten: DM 6587.-). Jedoch sind die Kategorien 4-7 nur schwach besetzt. Auch in den Kategorien 1-3 kommen sowohl „teure“ als auch „billige“ Therapien ähnlich häufig vor, und die „mittel teuren“ Therapien zwischen 3000.- und 9000.- machen jeweils etwa 75% aus.

Es lässt sich also kein Trend dahingehend ausmachen, dass Therapieerfolg und Therapieumfang (bzw. Therapiekosten) zusammenhängen. Die Korrelation zwischen Kosten und globaler Therapeuteinschätzung in der 1-Jahres-Katamnese beträgt 0,06 und ist nicht signifikant ($p=0,34$).

Die *Effektstärke* für die globale Therapieerfolgseinschätzung durch die Therapeuten 1 Jahr nach Therapieende beträgt nach der in 3.2.4 genannten Umkodierung mit der klassischen Formel $ES = 1.95$. Diese ist als ausgesprochen hoch zu betrachten (vgl. 2.2.2.1). Aus in 3.2.4 genannten methodischen Gründen kann diese aus einer retrospektiven Veränderungsmessung ermittelte ES allerdings nicht ohne weiteres mit „normalen“ Prä-Post-Effektstärken verglichen werden.

3.4.2.2 *Klinisch signifikante Veränderungen und Kosten-Effektivitäts-Index*

Als *klinisch signifikant gebessert* gelten die Patienten, die bei der 1-Jahres-Nachuntersuchung per Fremdeinschätzung (Therapeutenurteil) die Einschätzung „1...sehr viel besser“ oder „2...viel besser“ angaben (zusammen 74%, vgl. Tabelle 3.6). Berechnet man nun die Kosteneffektivität in Anlehnung an die Formel von Koester et al. (1982), dann ergibt sich folgendes Bild (Pfenningbeträge werden aufgerundet):

$$\begin{aligned} \text{Kosteneffektivität} &= \frac{\text{mittlere Therapiekosten aller Patienten (DM)}}{\text{Rate klinisch signifikant gebesserter Patienten}} \\ &= \frac{6587,- \text{ DM}}{0,74} = 8901,- \text{ DM} \end{aligned}$$

Die Therapie mit dem Ergebnis einer klinisch signifikanten Verbesserung in der Modalität „globale Verbesserungseinschätzung durch die Therapeuten“ kostete demnach 8901,- DM.

Das Augenmerk liegt hier auf dem definierten Behandlungsergebnis, an dem die durchschnittlichen Gesamtkosten für alle Behandlungen relativiert werden. Es handelt sich also um eine eher konservative Berechnung, da in die-

sem Fall die klinisch signifikant gebesserten Patienten den entstandenen finanziellen Aufwand für alle Patienten mittragen, die das Ergebnis nicht erreicht haben.

3.4.3 Effektivität (2) – Störungsfolgen: Beeinträchtigungsebene

3.4.3.1 Ausmaß der Effekte im Vergleich zu den Kosten

Für den Aspekt der Beeinträchtigungen liegen Angaben in drei unterschiedlichen Bereichen vor: Arbeit/Ausbildung, Freizeit/Sozialleben und Familienleben/häusliche Pflichten. Zunächst wird geprüft, wie stark der Zusammenhang der drei Beeinträchtigungsskalen ist; dafür wurden die Werte der Prä-Messung verwendet. Die Skalen Beeinträchtigung bei Arbeit, Freizeit und Familie korrelieren dabei wie folgt (Tab. 3.7):

Tabelle 3.7: Interkorrelationsmatrix für Beeinträchtigungsskalen

Beeinträchtigung in...	Arbeit	Freizeit	Familie
Arbeit	1.0	0.26	0.33
Freizeit		1.0	0.49
Familie			1.0

Die Tabelle zeigt nur geringe bis mittlere positive Zusammenhänge. Das deutet darauf hin, dass Patienten, die in einem Bereich sehr stark beeinträchtigt sind, dies nicht zwingend in den jeweiligen beiden anderen Bereichen sein müssen. Ebenso gibt es demnach Patienten, die in einem Bereich vermutlich gar nicht beeinträchtigt sind, dafür aber in einem anderen. Aus diesem Grund wird im folgenden getrennt auf jede einzelne Beeinträchtigungsskala eingegangen.

Zum Gebrauch der im folgenden für alle Bereiche aufgeführten Kosten-Effektivitäts-Matrizen siehe Kasten 3.1.

Kasten 3.1: Kosten-Effektivitäts-Matrizen

Die Darstellung der Veränderungen nach Therapie und der entsprechenden Kosten-Effektivität in den Beeinträchtigungsbereichen geschieht mit Hilfe von Kosten-Effektivitäts-Matrizen. Diese sind wie folgt zu lesen:

Aus den Randsummen ist zu erkennen, wie viele Patienten insgesamt vor und nach der Therapie wie stark beeinträchtigt waren. In den Zellen ist angegeben, wie sich die Veränderungen genau verteilen. Dabei stehen im weißen Bereich unterhalb der hellgrauen Diagonalen die Verbesserungen, im dunkelgrauen Bereich darüber die Verschlechterungen. Außerdem sind in den Zellen die jeweiligen durchschnittlichen (der Übersichtlichkeit halber auf jeweils DM 500.- gerundeten) Therapiekosten aufgeführt.

Zwei Beispiele (Daten der Kosten-Effektivitäts-Matrix in Abb. 3.2 auf nächster Seite):

- Die durchschnittlichen Therapiekosten für die 45 Patienten, die sich bezüglich der Arbeitsfähigkeit von „schwer beeinträchtigt“ zu „ein wenig beeinträchtigt“ verbessert haben, betragen DM 6500.-.
- Die durchschnittlichen Therapiekosten für die 33 Patienten, die sich bezüglich der Arbeitsfähigkeit „schwer beeinträchtigt“ lediglich zu „mäßig beeinträchtigt“ verbessert haben, betragen DM 6000.-.

Bei der 6-Wochen-Nachuntersuchung liegen N=403 Datensätze zur Beeinträchtigungen in Freizeit und Sozialleben, sowie N=404 zum Bereich Familienleben und häusliche Pflichten vor. Im Bereich Beeinträchtigung bei Arbeit/Ausbildung ist die Anzahl im Vergleich zu den anderen beiden Bereichen mit N=292 geringer. Das hängt damit zusammen, dass Patienten, die zum Zeitpunkt des Ausfüllens der Fragen weder ein Arbeitsverhältnis haben noch in Ausbildung stehen, diese Skala nicht beantworten müssen.

Arbeit/Ausbildung

Abbildung 3.1: Beeinträchtigungen bei Diagnostik und Post-Messung zusammen mit den jeweils entstandenen Kosten (DM-Beträge auf 500,- gerundet); Erläuterung zum Lesen von Kosten-Effektivitäts-Matrizen in Kasten 3.1

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)					Summe (Prä)
	0 gar nicht beeinträchtigt	1 ein wenig beeinträchtigt	2 mäßig beeinträchtigt	3 schwer beeinträchtigt	4 massiv beeinträchtigt	
0 gar nicht beeinträchtigt	12 6500,-	2 4500,-				14 (4,8%)
1 ein wenig beeinträchtigt	20 6000,-	6 7000,-	7 5500,-			33 (11,3%)
2 mäßig beeinträchtigt	28 6000,-	41 7000,-	12 6500,-	6 6000,-	2 6500,-	89 (30,5%)
3 schwer beeinträchtigt	23 6000,-	45 6500,-	33 6000,-	16 6500,-	2 7000,-	119 (40,8%)
4 massiv beeinträchtigt	7 7500,-	7 6500,-	11 (10) 7500,-	9 9500,-	3 5500,-	37 (12,7%)
Summe (Post)	90 (30,8%)	101 (34,6%)	63 (21,6%)	31 (10,6%)	7 (2,4%)	292 (100,0%)

Die Tabelle zeigt, dass sich N=49 Patienten (16,8%; mittelgraue Felder der Diagonale) in bezug auf Beeinträchtigung in Arbeit /Ausbildung weder verbessert noch verschlechtert haben.

19 Patienten (6,5%; dunkelgraue Felder oberhalb der Diagonale) gaben nach der Therapie eine größere Beeinträchtigung als vorher an.

Die hellen Felder unter der Diagonale zeigen, dass sich 76,7% (N=224) Patienten nach der Therapie hinsichtlich der Beeinträchtigung im Bereich Arbeit/Ausbildung verbessert haben.

Setzt man nun die Kosten zu diesen Ergebnissen in Beziehung, so fällt augenscheinlich auf, dass sich keine klaren Zusammenhänge zwischen Therapieerfolg und Therapiekosten finden lassen.

Die erfassten Veränderungen lassen sich auch als Effektstärke ausdrücken; danach ergibt sich für den Messzeitpunkt sechs Wochen nach Therapieende eine Prä-Post-Effektstärke von ES=1.27, ein Jahr nach Therapieende liegt sie bei ES=1.34. Die Veränderung ist also nicht nur kurzfristig, sondern bleibt bei der Ein-Jahres-Katamnese stabil. Diese Effekte kosteten durchschnittlich 6600,- DM (gerundet).

Freizeit/Sozialleben

Abbildung 3.2: Beeinträchtigungen bei Diagnostik und Post-Messung zusammen mit den jeweils entstandenen Kosten (DM-Beträge auf 500.- gerundet) ; Erläuterung zum Lesen von Kosten-Effektivitäts-Matrizen in Kasten 3.1

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)					Summe (Prä)
	0 gar nicht beeinträchtigt	1 ein wenig beeinträchtigt	2 mäßig beeinträchtigt	3 schwer beeinträchtigt	4 massiv beeinträchtigt	
0 gar nicht beeinträchtigt	4 4500,-	1 8000,-				5 (1,2%)
1 ein wenig beeinträchtigt	10 5000,-	13 5000,-	3 6500,-	1 8000,-		27 (6,7%)
2 mäßig beeinträchtigt	29 6000,-	47 6000,-	32 5500,-	5 6000,-		113 (28,0%)
3 schwer beeinträchtigt	43 6000,-	79 7000,-	51 7000,-	20 7000,-	2 4500,-	195 (48,4%)
4 massiv beeinträchtigt	12 7500,-	20 7500,-	23 9500,-	3 9500,-	5 6000,-	63 (15,6%)
Summe (Post)	98 (24,3%)	160 (39,7%)	109 (27,0%)	29 (7,2%)	7 (1,7%)	403 (100,0%)

18,4% (N=74; mittelgraue Felder der Diagonale) der Patienten waren nach der Therapie noch genauso wie vorher beeinträchtigt, und 3,0% (N=12; dunkelgraue Felder oberhalb der Diagonale) der Patienten gaben post eine größere Beeinträchtigung als vor der Behandlung an. Eine Verbesserung hinsichtlich der Beeinträchtigung in Freizeit und Sozialleben gaben 78,7% (N=317; helle Felder unterhalb der Diagonale) der Patienten an. Auch hier fällt auf, dass sich keine klaren Zusammenhänge zwischen Therapieerfolg und Therapiekosten finden lassen.

Die Prä-Post-Effektstärke beträgt für den Meßzeitpunkt sechs Wochen nach Therapieende ES=1.61, ein Jahr nach Therapieende liegt sie bei ES=1.55. Diese Effekte sind höher als bei der Beeinträchtigung im Bereich Arbeit/Ausbildung, und sie bleiben auch langfristig stabil. Für diese genannten Veränderungen mussten durchschnittlich 6600,- DM aufgewendet werden (gerundeter Wert).

Familienleben/Häusliche Pflichten

Abbildung 3.3: Beeinträchtigungen bei Diagnostik und Post-Messung zusammen mit den jeweils entstandenen Kosten (DM-Beträge auf 500.- gerundet); Erläuterung zum Lesen von Kosten-Effektivitäts-Matrizen in Kasten 3.1

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)					Summe (prä)
	0 gar nicht beeinträchtigt	1 ein wenig beeinträchtigt	2 mäßig beeinträchtigt	3 schwer beeinträchtigt	4 massiv beeinträchtigt	
0 gar nicht beeinträchtigt	39 5500,-	5 6000,-	1 3500,-		1 5000,-	46 (11,4%)
1 ein wenig beeinträchtigt	46 6000,-	31 5500,-	8 5000,-	1 2500,-		86 (21,3%)
2 mäßig beeinträchtigt	69 6500,-	53 6500,-	21 7000,-	9 7500,-	1 5000,-	153 (37,9%)
3 schwer beeinträchtigt	31 6000,-	39 7000,-	20 9500,-	10 8000,-	1 6500,-	101 (25,0%)
4 massiv beeinträchtigt	4 7000,-	5 6000,-	6 9000,-	2 11000,-	1 9500,-	18 (4,5%)
Summe (Post)	189 (46,8%)	133 (32,9%)	56 (13,9%)	22 (5,4%)	4 (1,0%)	404 (100,0%)

Bei 25,2% (N=102; mittelgraue Felder der Diagonale) der Patienten ergaben sich keine Veränderungen der Beeinträchtigungen nach der Therapie, und 6,7% (N=27; dunkelgraue Felder oberhalb der Diagonale) verschlechterten sich hinsichtlich der Beeinträchtigung im Bereich des Familienlebens. 68,1% (N=275; helle Felder unterhalb der Diagonale) der Patienten waren nach der Therapie weniger beeinträchtigt als vorher. Es konnten wie in den anderen beiden Bereichen auch hier keine klaren Zusammenhänge zwischen Erfolg und Kosten gefunden werden.

Drückt man die Veränderung wiederum als Prä-Post-Effektstärke aus, so ergibt sich für den Messzeitpunkt sechs Wochen nach Therapieende $ES=1.12$, und ein Jahr nach Therapieende ein Effekt von $ES=1.03$. Die im Bereich Familie hervorgerufenen Effekte sind jeweils geringer als die in den anderen beiden Bereichen festgestellten positiven Veränderungen, sie bleiben aber dennoch langfristig stabil. Es mussten auch hier durchschnittlich 6600,- DM (gerundet) aufgewendet werden.

Tabelle 3.8 fasst die Effektstärken noch einmal zusammen. Alle prä-post-Effektstärken sind als hoch zu bezeichnen (vgl. 2.2.2.1) und bleiben im Jahr nach Therapieende stabil.

Tabelle 3.8: Effektstärken in den Beeinträchtigungsmaßen

Beeinträchtigung in...	Effektstärken Prä/6-Wochen-NU	Effektstärken Prä/1-Jahres-Kat.
Arbeit/Ausbildung	1.27	1.34
Freizeit/Sozialleben	1.61	1.55
Familienleben/Häusliche Pflichten	1.12	1.03

3.4.3.2 *Klinisch relevante Veränderungen, „erfolgreiche Besserung“ und Kosten-Effektivitäts-Index*

Analog zur globalen Therapieerfolgseinschätzung, in der als klinisch relevant gebessert die Patienten mit der Einschätzung „1...sehr viel besser“ oder „2...viel besser“ definiert wurden, soll nun in der Modalität der Beeinträchtigungen ein geeignetes Kriterium für die klinische Relevanz der Verbesserungen gefunden werden.

Wie im Abschnitt 3.2.3 erläutert, wird eine von uns inhaltlich festgesetzte Umkodierung der 5-stufigen Beeinträchtigungs-Skala vorgenommen, die folgende Zuordnung vorsieht:

- Die Werte „0...gar nicht beeinträchtigt“ und „1...etwas beeinträchtigt“ werden zum Kriterium „unauffällig“,
- der Wert „3...mäßig beeinträchtigt“ zum Kriterium „subklinisch“ und
- „4...schwer beeinträchtigt“ und „5...sehr schwer/massiv beeinträchtigt“ zum Kriterium „klinisch relevante Beeinträchtigung“ zusammengefasst.

In den Abbildungen sind Verbesserungen, Nicht-Veränderungen und Verschlechterungen wiederum unterschiedlich grau hinterlegt. Nach dieser neuen Kodierung verteilen sich die Patienten vor/nach Therapie wie folgt (Abbildungen 3.5-3.7):

Abbildung 3.4: Beeinträchtigungen bei Arbeit/Ausbildung

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)			Summe (Prä)
	unauffällig	subklinisch	klinisch relevant	
unauffällig	40	7	0	47 (16%)
subklinisch	69	12	8	89 (31%)
klinisch relevant	82	44	30	156 (53%)
Summe (Post)	191 (65%)	63 (22%)	38 (13%)	292 (100%)

Abbildung 3.5: Beeinträchtigungen bei Freizeit/Sozialleben

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)			Summe (Prä)
	unauffällig	subklinisch	klinisch relevant	
Unauffällig	28	3	1	32 (8%)
Subklinisch	76	32	5	113 (28%)
Klinisch relevant	154	74	30	258 (64%)
Summe (Post)	258 (64%)	109 (27%)	36 (9%)	403 (100%)

Abbildung 3.6: Beeinträchtigungen Familienleben/Häusliche Pflichten

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)			Summe (Prä)
	unauffällig	subklinisch	klinisch relevant	
unauffällig	121	9	2	132
subklinisch	122	21	10	153
klinisch relevant	79	26	14	119
Summe (Post)	322	56	26	404

Als „*klinisch relevant gebessert*“ werden nun gemäß 3.2.2 die Patienten angesehen, die bei der Prä-Messung in die Kategorie klinisch relevant oder subklinisch *und* bei der Post-Messung in die Kategorie unauffällig fallen. In diese Gruppe werden auch diejenigen Patienten mit aufgenommen, die sich von kli-

nisch relevant nach subklinisch gebessert haben. Die Patienten, die vor *und* nach der Therapie unauffällig waren, werden nicht weiter berücksichtigt, da sie zwar durchaus erfolgreich behandelt worden sein mögen, sich aber in dem jeweiligen Bereich durch die Therapie nicht verbessern konnten.

Diejenigen Patienten, die nach der Therapie in die Kategorie „klinisch relevant“ fallen (egal, ob sie vorher unauffällig, subklinische oder bereits klinisch relevante Symptomatik zeigten) werden als „*nicht erfolgreich behandelt*“ definiert. In diese Gruppe werden auch diejenigen Patienten mit aufgenommen, die subklinisch geblieben sind oder sich von unauffällig nach subklinisch verschlechtert haben.

Die nicht erfolgreich behandelten lassen sich nun noch einmal dahingehend aufteilen, dass sich ein Teil klinisch relevant verschlechtert, der andere subklinisch oder klinisch relevant geblieben ist. Diese Unterteilung ist insofern sinnvoll, als dass für die Bewertung der Unbedenklichkeit einer Behandlungsmethode Verschlechterungen besonders zu beachten sind (möglicherweise Therapieschäden, vgl. 2.2.2.2). In Tabelle 3.9 sind die nach diesen Gesichtspunkten neugruppierten Daten für die drei Bereiche zusammengefasst.

Hätte man auf die Kategorie „subklinisch“ verzichtet und damit also unter anderem auch die Verbesserungen von klinisch relevant zu subklinisch als „*nicht erfolgreich behandelt*“ betrachtet, so würde die Rate klinisch relevant gebesserter Patienten im Bereich Arbeit 60%, im Bereich Freizeit 48% und im Bereich Familienleben 71% betragen (gemittelt 60%). Wie in den Abschnitten 2.2.2.2 und 3.2.2 herausgestellt, ist es aber ausgesprochen streng, individuell als sehr bedeutsam erlebte Veränderungen von klinisch relevant zu subklinisch als „Non-Responder“ oder gar als Misserfolge zu betrachten.

Tabelle 3.9: Prozentsätze der drei Veränderungskategorien bei den Beeinträchtigungsskalen (Nachkommastellen gerundet); diese Übersicht ist von den Patienten bereinigt, die bereits vor der Therapie jeweils keine klinisch relevante Beeinträchtigung in Arbeit, Freizeit, Familie angaben und die sich auch nicht weiter verschlechterten⁸.

Skala	klinisch relevant gebesserte Patienten	subklinisch oder klinisch relevant geblieben	klinisch relevant verschlechtert
Arbeit	77%	17%	6%
Freizeit	81%	17%	2%
Familie	80%	12%	8%
gemittelter Prozentsatz	79%	15%	6%

Die Rate der klinisch relevant gebesserten Patienten (79%) kann nun wiederum in Beziehung zu den Gesamtkosten gesetzt werden. Dabei ergibt sich folgende Kosten-Effektivität:

$$\begin{aligned} \text{Kosteneffektivität} &= \frac{\text{mittlere Therapiekosten aller Patienten (DM)}}{\text{Rate klinisch relevant gebesserter Patienten}} \\ &= \frac{6587,- \text{ DM}}{0,79} = 8338,- \text{ DM} \end{aligned}$$

Für einen nach der Modalität „Beeinträchtigung“ klinisch relevant gebesserten Patienten mussten also durchschnittlich 8338,- DM aufgewendet werden⁹. Es wird bei dieser konservativen Berechnung der Kosteneffektivität wieder berücksichtigt, dass die klinisch relevant gebesserten Patienten den entstandenen finanziellen Aufwand für alle Patienten mittragen, die dieses Ergebnis nicht erreicht haben.

⁸ Hätte man diese Aufstellung nicht um die jeweiligen vorher *und* nachher unauffälligen Patienten bereinigt, so hätte der Anteil erfolgreich behandelter Patienten im Bereich Arbeit 80%, im Bereich Freizeit 82% und im Bereich Familienleben 86% betragen (gemittelt 83%); die Bereinigung entspricht also einem konservativen Vorgehen.

⁹ Bei der – eben als unangemessen streng bezeichneten – Vorgehensweise ohne die Kategorie „subklinisch“ würde dieser Wert DM 10875,- betragen.

3.4.4 Effektivität (3) – Fragebogenmaße

3.4.4.1 Ausmaß der Effektivität

Im folgenden werden die Veränderungen in ihrem Ausmaß graphisch dargestellt (Abb. 3.8-3.12). Die einzelnen Werte bzw. die statistischen Signifikanzen sind aus den Abbildungen unmittelbar ersichtlich und sind in Tabelle 3.10 zusammengefasst..

Abbildung 378: Mittelwerte im BAI prä/post/1-Jahres-Katamnese

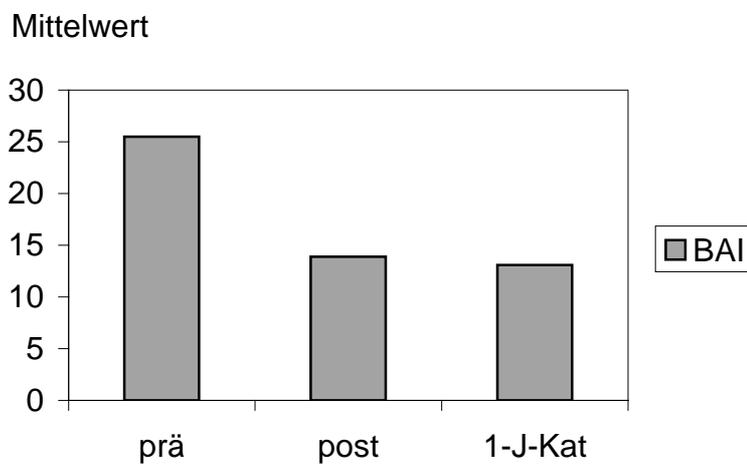


Abbildung 3.8: Mittelwerte im BSQ prä/post/1-Jahres-Katamnese

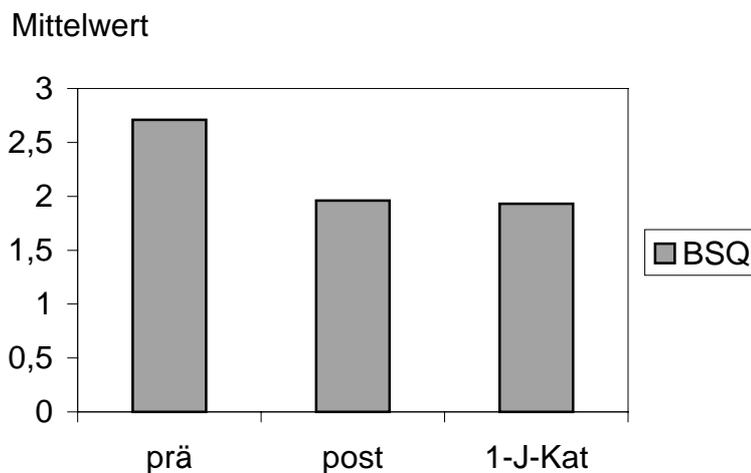


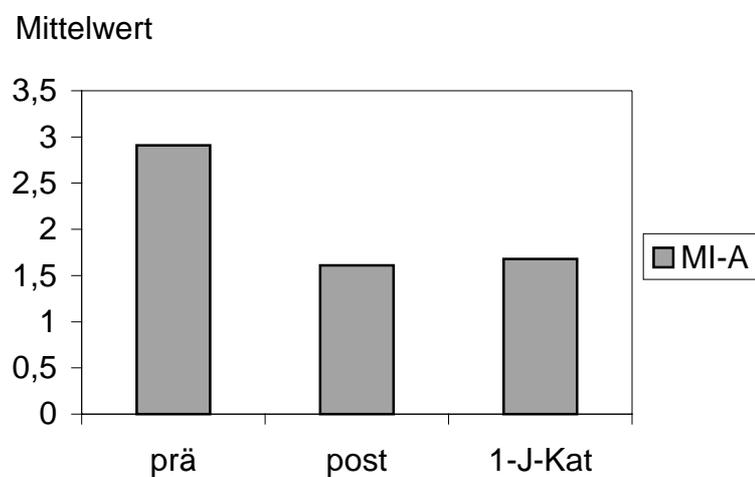
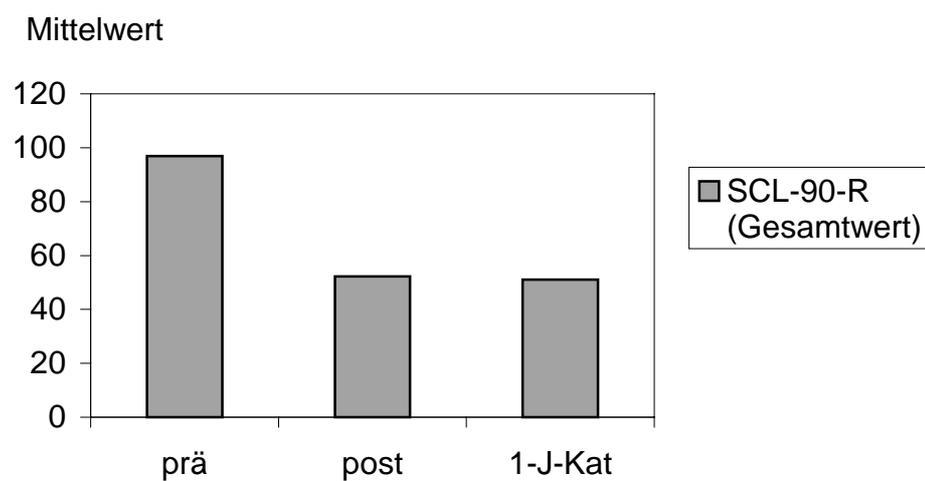
Abbildung 3.9: Mittelwerte im MI-A prä/post/1-Jahres-Katamnese**Abbildung 3.10: Mittelwerte SCL-90-R (Gesamtwert) prä/post/1-Jahres-Katamnese**

Abbildung 3.11: Mittelwerte im BDI prä/post/1-Jahres-Katamnese

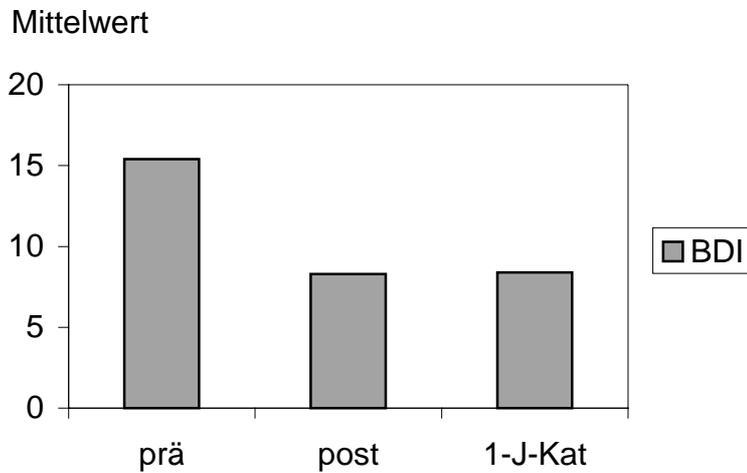


Tabelle 3.10: Mittelwertsunterschiede bei den Fragebogenmaßen mit Signifikanzen

	\bar{X} (prä)	\bar{X} (post)	\bar{X} (1-J-Kat.)	Vergleich prä/post	Vergleich prä/1-J- Kat. (t-Test)	Vergleich post/1-J- Kat. (t-Test)
BAI	25,5 (s=12,35)	13,9 (s=10,58)	13,1 (s=10,58)	** (Wil- coxon-Test) p=0,000	** (Wilcoxon- Test) p=0,000	n.s. (t-Test) p=0.31
BSQ	2,71 (s=0,75)	1,96 (s=0,67)	1,93 (s=0,67)	** (Wilcoxon- Test) p=0,000	** (Wilcoxon- Test) p=0,000	n.s. (t-Test) (p=0.11)
MI-A	2,91 (s=1,13)	1,61 (s=0,65)	1,68 (s=0,76)	** (Wilcoxon- Test) p=0,000	** (Wilcoxon- Test) p=0,000	* (t-Test) p=0.04
SCL-90- R (Gesamt- wert)	96,9 (s=53,96)	52,3 (s=44,03)	51,0 (s=43,91)	** (Wilcoxon- Test) p=0,000	** (Wilcoxon- Test) p=0,000	n.s. (t-Test) p=0.60
BDI	15,4 (s=8,75)	8,3 (s=7,97)	8,4 (s=7,96)	** (Wilcoxon- Test) p=0,000	** (Wilcoxon- Test) p=0,000	n.s. (t-Test) p=0.57

Diese Veränderungen können wiederum auch als Effektstärken ausgedrückt werden (Tab. 3.11).

Die Kosten, die diesen Veränderungen zuzuordnen sind, betragen wiederum durchschnittlich rund DM 6600.-, und es besteht wie in den vorangegangenen Abschnitten auch kein offensichtlicher Zusammenhang zum jeweiligen Ausmaß des Erfolges (sämtliche Korrelationen sind nicht signifikant).

Tabelle 3.11: Effektstärken für störungsspezifische und störungsübergreifende Maße (Formel nach Hartmann, Herzog & Drinkmann (1992); vgl. 3.2.4)

Fragebogen	ES (Diagnostik- 6-WNU)	ES (Diagnostik-1-J-Kat)
BAI	0,98	1,06
BSQ	1,05	1,08
MI-A	1,34	1,23
störungsspezifisch gemittelt	1,12	1,12
SCL-90-R-Gesamtwert	0,90	0,93
BDI	0,85	0,84
störungsübergreifend gemittelt	0,88	0,89

3.4.4.2 Klinisch relevante Veränderungen, „erfolgreiche Besserung“ und Kosten-Effektivitäts-Index

Analog zu den Beeinträchtigungsmaßen und mit den unter 3.2.2 beschriebenen Methoden werden die Fragebogendaten nun so gruppiert, dass sich klinisch signifikante Verbesserungen und Verschlechterungen nachvollziehen lassen.

In den Abbildungen 3.12-3.22 wird dargestellt, wie viele Patienten vor und nach der Therapie welchem Status zuzuordnen sind. Zunächst erscheinen die als eher ungünstig erachteten dichotomen Einteilungen „gesund vs. krank“ (nach dem cut-off-Punkt c in der Mitte zwischen beiden Populationen), dann

folgen die Einteilungen unter Hinzunahme der subklinischen Kategorie (vgl. 3.2.2, Tab. 3.1).

Aus den Randsummen ist zu erkennen, wie viele Patienten vor und nach Therapie welchem Status zuzuordnen sind. In den inneren Feldern ist abzulesen, wie die Randsummenveränderungen zustande gekommen sind. Auf Kostenangaben in den Zellen wie in den Kosten-Effektivitäts-Matrizen im Abschnitt zur Beeinträchtigung in verschiedenen Lebensbereichen wurde verzichtet, da sich wie zuvor keine bemerkenswerten Zusammenhänge finden lassen. Dafür werden anschließend Kosten-Effektivitäts-Indizes berechnet.

Abbildung 3.12: Einordnung der Patienten in die dichotomen Kategorien „unauffällig“ und „klinisch relevant belastet“ bezüglich des BAI prä (Zeilen) und post (Spalten); cut-off=11 entsprechend Autorenangaben

Prä-Messung (Diagnostik)	unauffällig	klinisch relevant	Summe (Prä)
unauffällig	29	9	38
klinisch relevant	171	274	445
Summe (Post)	200	283	483

Zusammenfassung BAI: N=171 gebessert, N=303 gleichgeblieben, N=9 verschlechtert

Abbildung 3.13: Einordnung der Patienten in die dichotomen Kategorien „unauffällig“ und „klinisch relevant belastet“ bezüglich des BSQ prä (Zeilen) und post (Spalten):

Prä-Messung (Diagnostik)	unauffällig	klinisch relevant	Summe (Prä)
unauffällig	80	12	92
klinisch relevant	205	173	378
Summe (Post)	285	185	470

Zusammenfassung BSQ: N=205 gebessert, N=253 gleichgeblieben, N=12 verschlechtert

Abbildung 3.14: Einordnung der Patienten in die dichotomen Kategorien „unauffällig“ und „klinisch relevant belastet“ bezüglich des MI-A prä (Zeilen) und post (Spalten)

Prä-Messung (Diagnostik)	unauffällig	klinisch relevant	Summe (Prä)
unauffällig	79	5	84
klinisch relevant	230	131	361
Summe (Post)	309	136	445

Zusammenfassung MI-A: N=230 gebessert, N=210 gleichgeblieben, N=5 verschlechtert

Abbildung 3.15: Einordnung der Patienten in die dichotomen Kategorien „unauffällig“ und „klinisch relevant belastet“ bezüglich der SCL-90-R (Gesamtwert, GSI) prä (Zeilen) und post (Spalten)

Prä-Messung (Diagnostik)	unauffällig	klinisch relevant	Summe (Prä)
unauffällig	106	6	112
klinisch relevant	189	189	378
Summe (Post)	295	195	490

Zusammenfassung SCL-90-R (GSI): N=189 gebessert, N=295 gleichgeblieben, N=6 verschlechtert

Abbildung 3.16: Einordnung der Patienten in die dichotomen Kategorien „unauffällig“ und „klinisch relevant belastet“ bezüglich des BDI prä (Zeilen) und post (Spalten)

Prä-Messung (Diagnostik)	unauffällig	klinisch relevant	Summe (Prä)
unauffällig	286	12	298
klinisch relevant	134	51	185
Summe (Post)	420	63	483

Zusammenfassung BDI: N=134 gebessert, N=337 gleichgeblieben, N=12 verschlechtert

Es folgen die entsprechenden Angaben unter Hinzunahme der Kategorie „subklinisch“

Abbildung 3.17: Einordnung der Patienten in die Kategorien „unauffällig“, „subklinisch“ und „klinisch relevant belastet“ bezüglich des BAI prä (Zeilen) und post (Spalten)

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)			Summe (Prä)
	unauffällig	subklinisch	klinisch relevant	
unauffällig	11	7	3	21
subklinisch	56	54	14	124
klinisch relevant	60	151	127	338
Summe (Post)	127	212	144	483

Zusammenfassung BAI: N=267 gebessert, N=192 gleichgeblieben, N=24 verschlechtert

Abbildung 3.18: Einordnung der Patienten in die Kategorien „unauffällig“, „subklinisch“ und „klinisch relevant belastet“ bezüglich des BSQ prä (Zeilen) und post (Spalten)

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)			Summe (Prä)
	unauffällig	subklinisch	klinisch relevant	
unauffällig	56	2	6	64
subklinisch	50	15	7	72
klinisch relevant	130	67	137	334
Summe (Post)	236	84	150	470

Zusammenfassung BSQ: N=247 gebessert, N=208 gleichgeblieben, N=15 verschlechtert

Abbildung 3.19: Einordnung der Patienten in die Kategorien „unauffällig“, „subklinisch“ und „klinisch relevant belastet“ bezüglich des MI-A prä (Zeilen) und post (Spalten)

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)			Summe (Prä)
	unauffällig	subklinisch	klinisch relevant	
unauffällig	60	1	3	64
subklinisch	55	19	1	75
klinisch relevant	146	89	71	306
Summe (Post)	261	111	73	445

Zusammenfassung MI-A: N=290 gebessert, N=150 gleichgeblieben, N=5 verschlechtert

Abbildung 3.20: Einordnung der Patienten in die Kategorien „unauffällig“, „subklinisch“ und „klinisch relevant belastet“ bezüglich der SCL-90-R (Gesamtwert) prä (Zeilen) und post (Spalten)

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)			Summe (Prä)
	unauffällig	subklinisch	klinisch relevant	
unauffällig	72	4	2	78
subklinisch	100	42	10	152
klinisch relevant	76	104	80	260
Summe (Post)	248	150	92	490

Zusammenfassung SCL-90-R: N=280 gebessert, N=194 gleichgeblieben, N=16 verschlechtert

Abbildung 3.21: Einordnung der Patienten in die Kategorien „unauffällig“, „subklinisch“ und „klinisch relevant belastet“ bezüglich des BDI prä (Zeilen) und post (Spalten)

Prä-Messung (Diagnostik)	Post-Messung (6-Wochen-Nachuntersuchung)			Summe (Prä)
	unauffällig	subklinisch	klinisch relevant	
unauffällig	146	3	2	151
subklinisch	104	33	10	147
klinisch relevant	67	68	50	185
Summe (Post)	317	104	62	483

Zusammenfassung BDI: N=239 gebessert, N=229 gleichgeblieben, N=15 verschlechtert

Die Angaben aus den vorangegangenen Abbildungen werden nun in %-Werte verrechnet, wobei diejenigen, die bereits vor der Therapie keine klinisch relevanten werte hatten und die sich auch nicht weiter verschlechterten, aus der Analyse ausgeschlossen wurden. Somit erhalten wir die Rate erfolgreicher Behandlungen (Tab. 3.12-3.13):

Tabelle 3.12: Klinisch relevante Veränderungen, dichotome Betrachtungsweise (Nachkommastellen gerundet); diese Übersicht ist von den Patienten bereinigt, die bereits vor der Therapie keine klinisch relevanten Werte hatten und die sich auch nicht weiter verschlechterten¹⁰

Skala	klinisch relevant gebesserte Patienten	klinisch relevant geblieben	klinisch relevant verschlechtert
BAI	38 %	60 %	2 %
BSQ	53 %	44 %	3 %
MI-A	62 %	36 %	2 %
gemittelter Prozentsatz störungsspezifisch	51 %	47 %	2 %
SCL-90-R (Gesamtwert)	49 %	49 %	2 %
BDI	68 %	26 %	6 %
gemittelter Prozentsatz störungsübergreifend	59 %	38 %	4 %
Gesamtmittel	55 %	43 %	2 %

¹⁰ Hätte man diese Aufstellung nicht um die vorher *und* nachher unauffälligen Patienten bereinigt, so hätte der Anteil erfolgreich behandelter Patienten im Gesamtmittel 64 %, der klinisch relevant gebliebenen 34 % und der verschlechterten 2 % betragen.

Die folgenden analogen Angaben mit subklinischer Kategorie sind optimistischer, was die Rate erfolgreicher Behandlungen betrifft – und was inhaltlich als sinnvoller erachtet wurde (2.2.2.2). Zudem unterliegen die Werte zwischen den verschiedenen Maßen geringeren Schwankungen als die der dichotomen Betrachtungsweise.

Tabelle 3.13: Klinisch relevante Veränderungen unter Berücksichtigung einer subklinischen Kategorie (Nachkommastellen gerundet); diese Übersicht ist von den Patienten bereinigt, die bereits vor der Therapie keine klinisch relevante Beeinträchtigung in Arbeit, Freizeit, Familie angaben und die sich auch nicht weiter verschlechterten ¹¹

Skala	klinisch relevant gebesserte Patienten	subklinisch oder klinisch relevant geblieben	klinisch relevant verschlechtert
BAI	56 %	38 %	6 %
BSQ	60 %	37 %	3 %
MI-A	75 %	23 %	2 %
gemittelter Prozentsatz störungsspezifisch	63 %	33 %	5 %
SCL-90-R (Gesamtwert)	67 %	29 %	4 %
BDI	71 %	25 %	4 %
gemittelter Prozentsatz störungsübergreifend	69 %	27 %	4 %
Gesamtmittel	66 %	30 %	4 %

¹¹ Hätte man diese Aufstellung nicht um die vorher *und* nachher unauffälligen Patienten bereinigt, so hätte der Anteil erfolgreich behandelter Patienten im Gesamtmittel 71 %, der klinisch relevant gebliebenen 26 % und der verschlechterten 3 % betragen.

Was ist nun ein möglichst realistischer und informationshaltiger Wert für die Rate erfolgreich gebesserter Patienten zur Berechnung des Kosten-Effektivitäts-Index?

Der Gesamtmittelwert unter Berücksichtigung der subklinischen Kategorie erscheint uns am solidesten/robuststen zu sein, da er sowohl störungsspezifische als auch störungsübergreifende Maße integriert und zudem differenzierter ist als die dichotome Beurteilung. Letztere erscheint sowohl was die klinische Alltagsbeobachtung betrifft als auch angesichts des globalen Urteils zum Behandlungserfolg durch Patienten und Therapeuten zu streng zu sein, wird aber gleichwohl auch aufgeführt, da sie die konventionelle Betrachtungsweise klinischer Signifikanz repräsentiert.

$$\begin{aligned}\text{Kosteneffektivität 1} &= \frac{\text{mittlere Therapiekosten aller Patienten (DM)}}{\text{Rate klinisch relevant gebesserter Patienten (incl. subklinisch)}} \\ &= \frac{6587,- \text{ DM}}{0.66} = 9980,- \text{ DM}\end{aligned}$$

$$\begin{aligned}\text{Kosteneffektivität 2} &= \frac{\text{mittlere Therapiekosten aller Patienten (DM)}}{\text{Rate klinisch relevant gebesserter Patienten (dichotom)}} \\ &= \frac{6587,- \text{ DM}}{0.55} = 11976,- \text{ DM}\end{aligned}$$

3.5 Variabilität des Kosten-Effektivitäts-Index in Abhängigkeit von Berechnungsmethoden und (Sub-) Stichproben

Bislang wurden die Kosten-Effektivitäts-Werte für die ganze Stichprobe berichtet. Hier soll die Kosten-Effektivität nun für einige Untergruppen bzw. hinsichtlich verschiedener Arten der Auswertung differenzierter dargestellt werden.

3.5.1 Messebene

Wie in 2.2.2 ausgeführt gibt es eine Vielzahl von Operationalisierungen und Kriterien zur Bewertung von Therapieerfolg. Im vergangenen Abschnitt wurden für drei Modalitäten getrennt Kosten-Effektivitäten bestimmt. Es liegt in der Natur der Sache, dass dabei die Kosten-Effektivitäts-Indizes eine gewisse Variabilität aufweisen, denn gäbe es *das* Therapie-Erfolgsmaß, so hätte man es wohl bereits entdeckt bzw. sich bereits darauf geeinigt.

So bleibt also die Bezugnahme auf verschiedene Maße – die allerdings möglichst etabliert und damit auf dem aktuellen methodischen Stand und vergleichbar sein sollten. Außerdem sollte man bei der Angabe „der“ Kosten-Effektivität einer Maßnahme transparent vorgehen. Deshalb werden in Tabelle 3.14 die durch unterschiedliche Messungen ermittelten Kosten-Effektivitäts-Indizes nochmals zusammenfassend aufgeführt.

Als Therapiekosten wurde gemäß 3.3.1 der durchschnittliche Betrag von DM 6587.- eingesetzt. Die Besserungsraten schwanken je nach Messebene bzw. Berechnungsgrundlage zwischen 51 % und 79 % und somit die Kosten-Effektivität als Quotient von Therapiekosten und Besserungsrate (also unter Einbezug nicht erfolgreicher Therapien) zwischen DM 8338.- und DM 12916.-.

Tabelle 3.14: Zusammengefasste Besserungsraten und Kosten-Effektivität in den drei verwendeten Modalitäten des Therapieerfolgs

Skala	klinisch relevant gebesserte Patienten	Kosten- Effektivitäts- Index (DM)
globales Maß	74%	8901.-
Beeinträchtigung in Arbeit/Freizeit/Familienleben (gemittelt; unter Einbezug einer subklinischen Kategorie)	79 %	8338.-
störungsspezifische Fragebogenmaße (gemittelt; unter Einbezug einer subklinischen Kategorie)	63 %	10456.-
störungsübergreifende Fragebogenmaße (gemittelt; unter Einbezug einer subklinischen Kategorie)	69 %	9546.-
störungsspezifische Fragebogenmaße (gemittelt; dichotome Auswertung)	51 %	12916.-
störungsübergreifende Fragebogenmaße (gemittelt; dichotome Auswertung)	59 %	11164.-
Gesamtmittel	66 %	9980.-

3.5.2 Therapie-Preis

Auch hinsichtlich des Therapiepreises, also des Zählers des Kosten-Effektivitäts-Index, sollte die Kosten-Effektivität auf Schwankungen untersucht werden. Hierfür werden die unterschiedlichen Therapiekosten (abhängig von Therapiepreis und Zahl der Behandlungseinheiten) dem globalen Maß gegenübergestellt (Tab. 3.15).

Verglichen werden die Behandlungsinstitute Braunschweig, Dresden und Marburg (für Münster lagen lediglich 14 Fälle vor). Der Therapieerfolg unterschied sich in den Instituten nicht signifikant; die durchschnittliche Besserungsrate im globalen Urteil betrug gemäß 3.3.2.1 74 %.

Tabelle 3.15: Kosten-Effektivität in Abhängigkeit von Therapiekosten der jeweiligen Behandlungseinrichtung

Institut	durchschnittlicher Therapiepreis (DM)	Kosten-Effektivitäts-Index (DM)
Braunschweig	5934.-	8019.-
Dresden	5044.-	6816.-
Marburg	7452.-	10070.-

Es findet sich auch hier eine deutliche Variabilität, die vor allem auf den in Marburg signifikant größeren Therapieumfang zurückzuführen ist: Die Therapien dauerten einschließlich Diagnostik in Marburg durchschnittlich 53, in Braunschweig und Dresden hingegen 42 Behandlungseinheiten. Dass der Therapiepreis in Dresden niedriger als in Braunschweig war, liegt an unterschiedlichen Behandlungssätzen für Ost und West.

Diese gefundenen Unterschiede können nicht auf störungsspezifische Unterschiede bei den behandelten Patienten erklärt werden.

3.5.3 Störungsspezifische Variablen

Es ist aufgrund der Vorbefunde zu erwarten, dass sich verschiedene *Störungsgruppen* sowohl bezüglich der Effektivitätswerte als auch bezüglich der benötigten Therapiedauer unterscheiden. In diesem Zusammenhang sollen hier die Kosten-Effektivität der Therapien in den Subgruppen Agoraphobie (mit Panikstörung), Sozialphobie und Zwangsstörung verglichen werden.

Tabelle 3.16 gibt einen Überblick über Kosten, Effektivität und Kosten-Effektivitäts-Index bei den unterschiedlichen Störungsgruppen.

Tabelle 3.16: Kosten-Effektivität in Abhängigkeit vom Störungsbild

Diagnose	durchschnittlicher Therapiepreis (DM) der Behandlungen bei dieser Diagnose	%-Satz „sehr viel besser“ oder „viel besser“ (Patientenurteil post)	Kosten-Effektivitäts-Index (DM)
Agoraphobie (incl. Panikstörung)	6523.-	78,3	8331.-
Sozialphobie	5259.-	53,3	9867.-
Zwangsstörung	9549.-	72,7	13136.-

Hier tritt wiederum der Sinn eines zweidimensionalen Maßes zur Therapiebeurteilung wie der Kosten-Effektivität zutage. Die Behandlung der Sozialphobie ist zwar am billigsten, erreicht aber vergleichsweise schlechtere Besserungsraten. Die Behandlung von Zwangsstörungen waren teurer, erzielte aber eine höhere Besserungsrate. Die Kosten-Effektivität der Agoraphobiebehandlung ist bei mittlerem Preis und höchsten Besserungsraten am günstigsten.

Ferner wurden die störungsspezifischen Variablen Krankheitsdauer und allgemeine Symptombelastung (ausgedrückt durch den Gesamtwert der SCL-90) bezüglich ihres Einflusses auf die Kosten-Effektivität untersucht. Dazu sollen zuerst die Zusammenhänge zwischen diesen Variablen und Therapieumfang (und damit Therapiepreis) sowie dem Therapieerfolg ausgedrückt im globalen Urteil (Patient post) betrachtet werden.

Die Interkorrelationsmatrix (Tab. 3.17) zeigt keine Einflüsse der Krankheitsdauer, wohl aber der Symptomschwere zu Therapiebeginn auf Therapieerfolg und –kosten auf. Deshalb wurde eine Berechnung der Kosten-Effektivität für hoch- vs. niedrige Eingangsbelastung berechnet (Median-Split beim SCL-Wert von 90; Tab. 3.18). Die Therapien bei den höher belasteten Patienten dauerten länger und waren weniger erfolgreich; sie mussten somit also – durchaus erwartungsgemäß – auch eine ungünstigere Kosten-Effektivität aufweisen.

Tabelle 3.17: Korrelationen zwischen Krankheitsdauer, SCL-90-Gesamtwert (prä) als Indikator für die Schwere der Eingangssymptomatik, Therapiekosten und Therapieerfolg (globales Urteil Patient post); Pearson-Koeffizient, ** = 2-seitig signifikant mit $p < 0.01$

	Krankheitsdauer	allg. Schwere Eingangssymptomatik	Therapiekosten	Therapieerfolg
Krankheitsdauer		-0,09	0,11	0,06
allg. Schwere Eingangssymptomatik			0,38**	0,13**
Therapiekosten				0,03
Therapieerfolg				

Tabelle 3.18: Kosten-Effektivität in Abhängigkeit von der Schwere der Eingangssymptomatik

Schwere der Eingangssymptomatik	durchschnittlicher Therapiepreis (DM) der Behandlungen	%-Satz „sehr viel besser“ oder „viel besser“ (Patientenurteil post)	Kosten-Effektivitäts-Index (DM)
leicht (SCL-90-Gesamtwert ≤ 90)	5649.-	79 %	7151.-
Schwer (SCL-90-Gesamtwert > 90)	7129.-	67 %	10640.-

3.5.4 Soziodemographische Variablen

Es wurden keine signifikanten Zusammenhänge zwischen Alter und Bildungsstand einerseits und Therapiepreis und Therapieerfolg andererseits gefunden. Ein geschlechtsspezifischer Zusammenhang (Frauen hatten im Durchschnitt längere und damit teurere Therapien) verschwand, wenn man die Schwere der Eingangssymptomatik kontrollierte (vgl. 3.5.3).

3.5.5 Variabilität der Kosten-Effektivitäts-Indices: Zusammenfassung

Die in den einzelnen Modalitäten berechneten Kosten-Effektivitäts-Indices schwanken folgendermaßen:

- Unter der Perspektive der Störungsfolgen (Beeinträchtigungen in verschiedenen Lebensbereichen) sind die Erfolgsraten am größten und damit die KE am besten.
- Die zweitbeste KE wurde mit der globalen Therapieeinschätzung ermittelt.
- Danach folgt die KE auf Grundlage der Verbesserungsraten bei der Symptombelastung; dabei ist erwartungsgemäß die KE bei den störungsspezifischen Maßen besser als bei den störungsübergreifenden (dort wurden auch breite, aber nicht so ausgeprägte Verbesserungen gefunden wie im durch die Therapie angezielten Symptombereich).
- Die Hinzunahme einer subklinischen Kategorie erhöht die Besserungsraten und damit auch die KE gegenüber einer dichotomen Betrachtungsweise „gesund“ vs. „krank“.

Die Therapiekosten setzen sich aus der Anzahl der Behandlungseinheiten und dem Preis für eine Therapiestunde zusammen. Dementsprechend wurden bei drei unterschiedlichen Ambulanzen unterschiedliche KE gefunden, obwohl die Erfolgsrate sich nicht unterschied:

- Dresden: günstigste KE, da am billigsten und vergleichsweise wenige Behandlungseinheiten;
- Braunschweig: mittlere KE, da teurer Westsatz, aber vergleichsweise wenige Behandlungseinheiten;
- Marburg: ungünstigste KE, da teurer Westsatz und vergleichsweise viele Behandlungseinheiten.

Störungsspezifische Aspekte moderieren sowohl Therapiedauer (und damit die Therapiekosten) als auch die Erfolgsrate. Folgende Unterschiede bei der KE wurden hier gefunden:

- Panikstörung mit Agoraphobie: günstigste KE, da zwar mittlerer Preis, aber hohe Erfolgsrate;
- Sozialphobie: mittlere KE, da zwar niedrigster Preis, aber auch niedrigste Besserungsrate;
- Zwangsstörung: ungünstigste KE, da zwar gute Erfolgsrate, aber deutlich teurere (weil umfangreichere) Behandlungen. Dies entspricht der Befundlage in diesem Bereich (Reizkonfrontation mit Reaktionsverhinderung bei Zwangspatienten dauert in der Regel länger, bis Erfolg eintritt, als die Reizkonfrontation bei Phobikern).

Die Krankheitsdauer hatte keinen Einfluss auf die KE, wohl aber die Eingangs-Symptomschwere. Bei den schwereren Fällen waren die Erfolgsraten etwas niedriger und die Therapiekosten deutlich höher, was sowohl über als auch unter dem Bruchstrich zu einer Verschlechterung der KE beiträgt.

Soziodemographische Variablen hatten keinen Einfluss auf die KE (außer weibliches Geschlecht, da Frauen im Schnitt eine höhere Symptombelastung aufwiesen).

3.6 Überprüfung von Selektionseffekten durch Dropouts

Im Rahmen dieser Studie war es nicht möglich zu kontrollieren, dass alle behandelten Patienten nachuntersucht wurden, oder dass gar alle Patienten, die frühzeitig die Therapie abbrachen, im Sinne einer *intention-to-treat-study* weiterverfolgt wurden. Letzteres ist für eine Versorgungsstudie wie dieser auch nicht erforderlich: die Aussagen sollen von vorneherein auf Patienten begrenzt sein, die eine kognitiv-verhaltenstherapeutische Behandlung aufgesucht und „in voller Dosis“ erhalten haben. In jedem Fall muss aber Störhypothesen nachgegangen werden, die die Validität der Ergebnisse in Frage stellen könnten, wie etwa der möglichen Unterstellung, dass systematisch „schwierige“ Patienten nicht nachuntersucht worden sein könnten.

Zentral ist auch die Frage, ob es eine systematische Verzerrung dahingehend gibt, dass Patienten, die einen „echten“, aber erfolglosen Behandlungsversuch unternommen haben, seltener nachuntersucht wurden.

3.6.1 Unterscheiden sich nachuntersuchte und nicht nachuntersuchte Patienten?

Um zu überprüfen, ob die nicht nachuntersuchten Patienten des Ausgangsdatensatzes mit „echtem“ Therapieversuch, die aufgrund fehlendem SCL-Postwert nicht in diese Studie mitaufgenommen wurden, eventuell einer anderen Population entstammen als die Nachuntersuchten, sollen diese beiden Gruppen hinsichtlich ihrer Prä-Werte verglichen werden. Folgende Variablen wurden diesbezüglich analysiert:

Soziodemografische Variablen:

- Alter
- Geschlecht
- Bildungsstand
- ausgeübter Beruf

Störungsspezifische Variablen:

- Krankheitsdauer
- BAI
- BSQ
- MI-A
- alle Subskalen der SCL-90-R
- BDI

In den Tabellen 3.19 und 3.20 sind die Werte unserer Stichprobe im Vergleich mit denjenigen aus der Gesamtdatenbank der CDS angegeben, die zwar auch eine Angststörung aufwiesen, aber wegen fehlender Nachuntersuchung nicht aufgenommen worden waren. Für diese Vergleichsgruppe (N = 223) wurde außerdem gefordert, dass ein „echter“ Therapieversuch unternommen worden war (definiert über eine Therapiestundenzahl größer 10).

Tabelle 3.19: Vergleich der Stichprobenpopulation (N=493) und einer ebenfalls behandelten, aber nicht nachuntersuchten Vergleichsgruppe (N=223) in soziodemographischen Variablen

		Untersuchungsgruppe	Vergleichsgruppe	Signifikanz
Alter		$\bar{x} = 35,7$ $s = 9,4$	$\bar{x} = 36,2$ $s = 9,2$	n.s. (t-Test) $p = 0,79$
Geschlecht	männlich: weiblich:	38 % 62 %	33 % 67 %	n.s. (Chi-Quadrat) $p = 0,11$
höchster Schulabschluss	keiner: Hauptschule: Realschule: Fachabitur/Abitur: abgeschlossenes Fach-/Hochschulstudium:	1,2 % 18,3 % 28,4 % 26,9 % 25,2 %	0,5 % 17,6 % 28,6 % 30,8 % 20,3 %	n.s. (Chi-Quadrat) $p = 0,64$
derzeitige Berufsgruppe	Arbeiter: Facharbeiter/ Handwerker: Angestellter/Beamter einfacher/ mittlerer Dienst: gehobener Dienst: Selbständiger: Schüler/Azubi/ Student: Hausfrau: arbeitslos: berentet:	3,1 % 5,5 % 29,3 % 14,4 % 12,8 % 14,9 % 13,3 % 4,4 % 2,2 %	1,3 % 3,2 % 25,2 % 11,0 % 11,0 % 21,3 % 18,1 % 6,5 % 2,6 %	n.s. (Chi-Quadrat) $p = 0,08$

Tabelle 3.20: Vergleich der Stichprobenpopulation (N=493) und einer nicht nachuntersuchten Vergleichsgruppe (N=223) in störungsbezogenen Variablen zum Prä-Zeitpunkt

	Untersuchungsgruppe	Vergleichsgruppe	Signifikanz
Krankheitsdauer	M = 9,1 s = 8,3	M = 9,3 s = 9,1	n.s. (t-Test, 2-seitig) p=0,79
BAI	M = 25,5 s = 13,9	M = 23,1 s = 12,5	n.s. (t-Test, 2-seitig) p=0,16
BSQ	M = 2,71 s = 0,76	M = 2,74 s = 0,79	n.s. (t-Test, 2-seitig) p=0,73
MI-A	M = 2,91 s = 1,13	M = 2,97 s = 1,2	n.s. (t-Test, 2-seitig) p=0,50
SCL-90-R Gesamtwert	M = 96,9 s = 53,9	M = 104,9 s = 60,0	n.s. (t-Test, 2-seitig) p=0,10
SCL-90-R Somatisierung	M = 12,5 s = 8,5	M = 12,5 s = 8,2	n.s. (t-Test, 2-seitig) p=0,96
SCL-90-R Zwanghaftigkeit	M = 10,3 s = 7,3	M = 10,9 s = 7,9	n.s. (t-Test, 2-seitig) p=0,36
SCL-90-R Unsicherheit im Sozialkontakt	M = 10,0 s = 7,3	M = 10,5 s = 7,5	n.s. (t-Test, 2-seitig) p=0,48
SCL-90-R Depressivität	M = 16,1 s = 10,5	M = 17,6 s = 11,3	n.s. (t-Test, 2-seitig) p=0,12
SCL-90-R Ängstlichkeit	M = 14,8 s = 8,3	M = 14,8 s = 8,4	n.s. (t-Test, 2-seitig) p=0,99
SCL-90-R Aggressivität / Feindseligkeit	M = 4,5 s = 3,8	M = 4,8 s = 4,6	n.s. (Mann-Whitney-U-Test, 2-seitig) p=0,93
SCL-90-R Phobische Angst	M = 11,4 s = 7,9	M = 11,3 s = 7,9	n.s. (t-Test, 2-seitig) p=0,94
SCL-90-R paranoides Denken	M = 4,6 s = 4,2	M = 5,3 s = 4,9	n.s. (Mann-Whitney-U-Test, 2-seitig) p=0,22
SCL-90-R Psychotizismus	M = 6,1 s = 5,4	M = 7,1 s = 6,8	n.s. (Mann-Whitney-U-Test, 2-seitig) p=0,26
BDI	M = 15,4 s = 8,8	M = 16,8 s = 9,6	n.s. (t-Test, 2-seitig) p=0,10

Anstelle von 2-seitigen t-Tests für unabhängige Stichproben wurden Mann-Whitney-U-Tests berechnet, wenn die Varianzen der Gruppen unterschiedlich waren. Da sich weder in soziodemographischen noch in Störungsvariablen signifikante Unterschiede finden lassen, kann man davon ausgehen, dass nachuntersuchte und nicht nachuntersuchte Patienten in den Behandlungseinrichtungen aus derselben Population stammen¹².

Ein weiteres dahingehendes Argument ist der Vergleich mit einer früheren Studie auf der Grundlage von CDS-Daten (Frank & Fiegenbaum, 1993), bei der annähernd 100% der dort behandelten Patienten nachuntersucht wurden: zu den dortigen Ergebnissen zeigen sich keine Differenzen.

3.6.2 Hinweise auf Unsystematik von Drop-Outs aufgrund des Settings

Auch offensichtliche Ressourcenprobleme der Behandlungseinrichtung sprechen dafür, dass keine Verzerrung aufgrund nicht durchgeführter Nachuntersuchungen bestehen. Folgende Gründe lassen sich für nicht durchgeführte Nachuntersuchungen anführen, die sich alle nicht systematisch auf die Ergebnisse auswirken dürften:

- Verzögerte Eingabe und Schwierigkeiten des Datentransfers zwischen den vier Instituten

Die Eingabe der Daten erfolgt in der Regel durch Praktikanten und studentische Hilfskräfte. Es gibt regelmäßig Phasen, in denen die zur zügigen Eingabe nötigen Stellen unterbesetzt sind, so dass Verzögerungen entstehen. Außerdem gibt es bisweilen Probleme bei der Abgleichung der Daten zwischen den Instituten sowie EDV-bedingte Schwierigkeiten aufgrund des Fehlens eines zentralen Systemadministrators und Datenwirts. Es ist daher möglich, dass mittlerweile Daten hinzugekommen sind (bzw. in nächster Zeit hinzukommen), die bereits erhoben sind, auf die wir aber keinen Zugriff hatten.

- Keine finanziellen Mittel zur Eintreibung von Daten

Sind Patienten zum Nachuntersuchungstermin nicht erreichbar, besteht aufgrund mangelnder Ressourcen nicht die Möglichkeit, Aufwand für Nachforschungen zu betreiben, um sie doch noch zu erreichen. Zudem kann den Patienten keine Aufwandsentschädigung für die etwa zwei Stunden angeboten

¹² Wenn man unsere Stichprobe in die Gruppe der Patienten aufteilt, die zur 1-Jahres-Katamnese nachuntersuchten vs. nicht nachuntersucht wurden, in ergibt sich das selbe Bild von nicht-signifikanten Unterschieden in t- bzw. Mann-Whitney-U-Tests (einzige Ausnahme: die Nachuntersuchten wiesen zum Prä-Zeitpunkt einen signifikant höheren MI-A-Durchschnittswert auf)

werden, die das Ausfüllen der Fragebogenbatterie dauert (plus die Zeit für das Nachuntersuchungsinterview). So kann die Motivation bei Verweigerern von Nachuntersuchungen „aufgrund Zeitmangel“ nicht gehoben werden.

- Vergessen der Nachuntersuchung nach Weggang des behandelnden Therapeuten

Die Fluktuation der behandelnden Therapeutinnen und Therapeuten ist relativ hoch, da der Verbleib in der Einrichtung im Zusammenhang mit der wissenschaftlichen Tätigkeit auf drei bis fünf Jahre begrenzt ist. Da in der Regel für die Nachuntersuchungen diejenigen verantwortlich sind, die die Therapie durchgeführt haben, kann es trotz Kontrolle der Institutsleiterinnen bisweilen vorkommen, dass beim Weggang von Therapeuten kein Ersatz-Diagnostiker für die Nachuntersuchung eingesetzt wird.

Und schließlich spricht auch die Größe der Stichprobe dafür, dass die hier präsentierten Ergebnisse repräsentativ sind: bei mehreren hundert nachuntersuchten Patienten müssten schon ziemlich viele hypothetische erfolglose Therapien systematisch nicht nachuntersucht bzw. nicht in den Datensatz eingegeben worden sein, um die Validität der Ergebnisse bedeutsam in Frage zu stellen.

3.2 Einordnung der Effekte in den Forschungsstand

Um die vorgefundenen Ergebnisse dieser Effektivitätsstudie in den Forschungsstand zu kontrollierten Wirksamkeitsstudien einzuordnen und die Frage zu klären, ob die Implementation störungsspezifischer kognitiv-behavioraler Psychotherapie in den klinischen Alltag gelungen ist, sollen sie den Ergebnissen einer jüngst durchgeführten Meta-Analyse gegenübergestellt werden. So kann im Sinne eines „*benchmarking*“-Ansatzes ein Kriterium aufgestellt werden, an dem sich die zu untersuchende Behandlung messen lassen muss (wenn es schon keine Vergleichs- bzw. Kontrollgruppe gibt). Ein Beispiel für diesen Ansatz findet sich bei Wade et al. (1998).

Ruhmland und Margraf (2001) berichten Effektstärken für verschiedene Behandlungsformen bei der Therapie von Angststörungen und beziehen dabei den Forschungsstand bis 1997 mit ein. Aufgrund des Übergewichts der Primärdiagnose Agoraphobie mit Panikstörung in unserer Stichprobe sollen die Ergebnisse zu dieser Diagnose aus der Meta-Analyse als Referenzwerte herangezogen werden. Dabei ist natürlich zu bedenken, dass ein Vergleich mit kontrollierten Studien in seiner Aussagekraft dort begrenzt wird, wo unklar ist, ob die untersuchten Behandlungen wirklich substanziell ähnlich sind. Dies stellt übrigens ein allgemeines zentrales Problem bei Meta-Analysen dar (vgl. z.B. Czienskowski, 2000)

Bei Ruhmland und Margraf wurden die entsprechenden Effektstärken mit der gleichen Formel berechnet wie in der vorliegenden Arbeit, was insofern von Bedeutung ist, dass es in bestimmten Fällen bei unterschiedlichen Berechnungsarten zu variablen Ergebnissen kommen kann (Hartmann & Herzog, 1995).

In der Tabelle 5.1 finden sich unter „CDS“ folgende Effektstärken (vgl. Tab. 3.11, Tab. 3.8), die der Gruppierung bei Ruhmland & Margraf (2001) am ehesten entsprechen:

- Hauptsymptomatik: gemittelte Effektstärken in den störungsspezifischen Fragebögen
- Depressivität: Effektstärke BDI
- Beeinträchtigung: gemittelte Effektstärken in den drei untersuchten Beeinträchtigungsbereichen
- Symptome insgesamt: Effektstärke SCL-90-R

Die reine Konfrontationsbehandlung hat die besten Effekte vorzuweisen. Ob diese signifikant größer als in unserer Studie sind, konnte nicht ermittelt werden, da uns eine diesbezügliche statistische Absicherung (Unterschied einer einzelnen Effektstärke zu einer Stichprobe anderer Effektstärken) nicht be-

kannt ist. Die Ähnlichkeit der hier gefundenen Effektstärken („CDS“) mit den Vorbefunden aus kontrollierten Studien zu Konfrontation, kognitiv-behavioraler und kognitiver Therapie, sowie die (auch wenn nicht statistisch abgesichert augenscheinlich deutliche) Überlegenheit gegenüber den Werten von nondirektiver Therapie und Wartelisten kann als starker Hinweis dafür gewertet werden, dass der Transport des empirisch fundierten Verfahrens in die Praxis erfolgreich war.

Tabelle 5.1 : Panikstörung mit Agoraphobie: Vergleich der Ergebnisse der vorliegenden Studie („CDS“) mit der Metaanalyse von Ruhmland & Margraf (2001); Erläuterungen im Text

	Haupt-symptomatik	Depressivität	Beeinträchtigung	Symptome insgesamt
CDS				
Effektstärke:	1.12	0.85	1.33	0.90
Konfrontation				
Mittelwert (SD):	1.64 (1.16)	0.83 (1.03)	2.11 (3.56)	0.99 (0.06)
Anzahl der Effektstärken:	29	18	7	2
Kog.-behav. Therapie				
Mittelwert (SD):	1.19 (0.56)	0.84 (0.64)	0.89 (0.35)	0.96 (0.46)
Anzahl der Effektstärken:	49		7	10
Kognitive Therapie				
Mittelwert (SD):	0.92 (0.20)	0.91 (0.35)	--	--
Anzahl der Effektstärken:	3	2		
Nondirektive Therapie				
Mittelwert (SD):	0.59 (0.37)	0.56 (0.27)	0.78 (0.54)	--
Anzahl der Effektstärken:	4	3	2	
Warteliste				
Mittelwert (SD):	0.02 (0.39)	0.19 (0.24)	--	--
Anzahl der Effektstärken:	11	6		

Die Diskussion der Kosten-Effektivitätsstudie findet sich im Anschluss an die nun folgende Kosten-Nutzen-Analyse.

4 Kosten-Nutzen-Analyse

4.1 Fragestellungen

Nachdem die Effektivität der untersuchten Therapieform – der kognitiv-verhaltenstherapeutischen Behandlung von Angststörungen – im vorangegangenen Teil belegt wurde und den Kosten der Therapie gegenübergestellt wurde, soll diese Therapieform nun rein unter monetären Gesichtspunkten analysiert und bewertet werden. Wie eingangs unter 1.3 aufgeführt, gibt es für Kosten-Nutzen-Analysen verschiedene Perspektiven zu berücksichtigen. Gemäß den Vorbefunden zu den Kosten psychischer Störungen wurden folgende Kosten- und Nutzenfaktoren zur Analyse ausgewählt:

1. Kosten für die Volkswirtschaft durch Produktivitätsausfall durch störungsbedingte Arbeitsunfähigkeit sowie für das Gesundheitssystem durch erhöhte Inanspruchnahme von Gesundheitsleistungen. Der Nutzen einer Behandlungsmaßnahme errechnet sich hier über die Verringerung solcher Kostenfaktoren.
2. Subjektiver Nutzen für den Patienten, monetär ausgedrückt über die Zahlungsbereitschaft (Willingness to pay, WTP). Da dieses Maß noch nicht allgemein im medizinischen (und noch weniger im Psychotherapeutischen) Bereich etabliert ist, wird dieser Nutzen gesondert untersucht bzw. erst in einem zweiten Schritt dem unter 1. errechneten Nutzen zugeschlagen.

Beide Punkte zusammengenommen decken den Großteil der möglichen Perspektiven und der direkten und indirekten tangiblen und intangiblen Kosten ab. Die hier durchgeführte Kosten-Nutzen-Analyse ist dem Forschungsgebiet der Evaluationsforschung zuzuordnen und nur zum Teil der Versorgungsforschung. Deshalb werden in dieser Arbeit gesamtgesellschaftliche soziale Fragestellungen nicht berücksichtigt, wie etwa die Zugänglichkeit psychosozialer Maßnahmen für alle Bevölkerungsschichten oder Planspiele dahingehend, wie die Versorgung psychischer Störungen bundesweit unter Kostengesichtspunkten optimalerweise organisiert werden müsse.

Ferner beschäftigt sich diese Arbeit mit methodischen Aspekten von Kosten-Nutzen-Analysen in diesem Feld. Dies betrifft den Umgang mit Schätzwerten, Sensitivitätsanalysen und den Vergleich längs- und querschnittlicher Herangehensweise. Ebenso wie bei der Kosten-Effektivitäts-Analyse gibt es nämlich in Abhängigkeit der gewählten Methodik immer mehrere Kosten-Nutzen-Indices. Deshalb stellt sich auch hier in besonderem Maße die Forde-

rung nach Transparenz, wenn die Ergebnisse mit denen anderer Behandlungsmaßnahmen verglichen werden sollen.

Fragestellung 1: *In welchem Verhältnis stehen Behandlungskosten und wirtschaftlicher Nutzen des untersuchten Treatments zueinander?*

Verändern sich die durch die Störung verursachten Kostenfaktoren nach der Therapie? Können Einsparungen nach einer Psychotherapie die Kosten dieser Therapie wieder aufwiegen, und wenn ja, wann? Schwierigkeiten ergeben sich hier vor allem bei der monetären Bezifferung der Störungskosten, da – vor allem wegen datenschutzrechtlicher Beschränkungen und sehr heterogener Preise – keine vollständig exakten Daten herangezogen werden können. Bei der Interpretation muss berücksichtigt werden, dass bei dieser Fragestellung vorausgesetzt wird, dass die Einsparungen in der Zeit nach der Therapie ohne die untersuchte störungsspezifische Psychotherapie nicht zustande gekommen wären. Diese Annahme wird unterstützt durch die deutlichen und stabilen positiven Effekte, die in vorangegangener Kosten-Effektivitäts-Analyse ermittelt wurden.

Fragestellung 2: *Wie verändert die Hinzunahme des WTP-Ansatzes das Kosten-Nutzen-Verhältnis?*

Wieviel % des Haushalts-Nettoeinkommens wären Patienten bereit, für die absolvierte Therapie zu zahlen, wenn es keine Krankenkassen gäbe? Mit dem WTP-Ansatz wird eine Kosten-Nutzen-Analyse vor allem um intangible Aspekte ergänzt, die bei der konventionellen Vorgehensweise in Fragestellung 1 unberücksichtigt bleiben. Für die Ermittlung der WTP muss mangels eines etablierten Maßes ein eigenes Erhebungsinstrument entwickelt werden, das die dem Forschungsstand entsprechenden Anforderungen angemessen auf den Bereich der Psychotherapie überträgt.

Fragestellung 3: *Wie können Kosten-Nutzen-Analysen (im Psychotherapiebereich) vereinfacht werden?*

Dies betrifft die Frage nach der „Kosten-Nutzen-Analyse der Kosten-Nutzen-Analyse“. Der Mangel an Kostenstudien liegt nicht zuletzt auch in der Aufwendigkeit der Durchführung von Kostenstudien begründet. Mit dieser Fragestellung soll das Verhältnis von Aufwand bei der Datenerhebung und Auswertung und dem Gehalt bzw. der Verwertbarkeit der Ergebnisse aufgegriffen werden.

4.2 Methoden

4.2.1 Berechnung der Behandlungskosten

Die Kostenseite wurde entsprechend zur Kosten-Effektivitäts-Analyse mit dem üblichen *Betriebskostenansatz* berechnet (vgl. 1.2.3), indem die Zahl der in Rechnung gestellten Behandlungseinheiten (bzw. der Mittelwert der untersuchten Gruppe) mit dem Preis pro Behandlungseinheit multipliziert wurde. Da die Studie multizentrisch war und verschiedene Preise für eine Behandlungseinheit existierten (DM 138.- „Westsatz“, DM 117.30 „Ostsatz“) wurden entsprechende gewichtete Werte berechnet.

Der durchschnittliche Behandlungsumfang betrug in dieser Stichprobe 41 Behandlungseinheiten¹³. Die durchschnittlichen Therapiekosten betragen DM 5371.-

4.2.2 Nutzenberechnung I: Bestimmung der Kostenfaktoren vor und nach der Behandlung

Zunächst müssen die durch die Störung verursachten Kostenfaktoren vor Therapie sowie nach Therapie weiterbestehende Kostenfaktoren bestimmt werden. Entsprechend den Vorbefunden (vgl. 2.3.2) beschränkten wir uns auf die wichtigsten Kostenfaktoren

- ambulante Behandlungen (Psychotherapie, Fach- Haus- und Notarztkontakte),
- stationäre Behandlungen, sowie
- Arbeitsunfähigkeit.

Da andere (sog. „Marginal“-) Kosten (wie etwa durch Medikamente) relativ dazu gering sind, deren Erhebung aber unverhältnismäßigen Aufwand erfordert, wurden sie aus der Analyse ausgeschlossen.

Optimalerweise würde man die Angaben z.B. über die Inanspruchnahme medizinischer Leistungen anhand von Daten der Krankenkassen überprüfen. Die exakte Ermittlung der Behandlungen und der Arbeitsunfähigkeit über Kostenträger war aus rechtliche Datenschutzbestimmung des SGB-V unmög-

¹³ Dies ist ein geringerer Behandlungsumfang als in der Stichprobe zur Kosten-Effektivität (Teil 3 dieser Arbeit), die in den Jahren früher behandelt wurde. Dies läßt sich im wesentlichen auf die restriktivere Kostenerstattungspolitik der Krankenkassen gegen Ende der 90er Jahre zurückführen: die späteren Therapien *mussten* mit weniger Stunden auskommen.

lich, unter anderem da diese Daten nicht personenbezogen gespeichert werden. Die Ermittlung über die Kostenträger hätte nur in einem Großprojekt in Zusammenarbeit mit Betriebskrankenkassen (vgl. BKK-Studie; Zielke, 1993) oder mit Privatkassen stattfinden können. Da aber eine solche Kooperation nicht möglich war (fast niemand der untersuchten und behandelten Patienten war in einer BKK oder privat versichert), musste ein hinreichend valider Ersatz geschaffen werden.

Die einzige Möglichkeit bestand in der direkten Befragung der Patienten. Im Anhang finden sich die von uns entwickelten eingesetzten Fragebögen, die in den Diplomarbeiten von Sommer (1998) und Pieper (1999) validiert wurden.

Dabei wurde folgendermaßen vorgegangen: Zunächst wurden die nötigen Items im Interview auf Verständlichkeit und „Augenscheinvalidität“ geprüft, indem Patienten und Therapeuten der Behandlungseinrichtung dazu befragt wurden. So entstand der Fragebogen zur Erhebung von Krankheitskosten, in dem direkt und konkret nach den Kostenfaktoren gefragt wird (z.B. „Haben Sie aufgrund der Beschwerden und Symptome, derentwegen Sie die Behandlung in unserer Einrichtung aufgesucht haben, jemals vorher eine Psychotherapie durchgeführt?“ „Wenn ja, wann, bei wem und mit wie vielen Sitzungen?“).

Nun galt es, Störhypothesen zu testen bzw. nach systematischen Verzerrungen gegenüber der „objektiven“ Krankenkassendaten-Methode zu suchen. Die Möglichkeit, die Patienten könnten vor Therapie systematisch Behandlungen und Arbeitsunfähigkeitszeiten erfinden, die in der Vergangenheit gar nicht stattgefunden haben, wurde mangels Plausibilität ausgeschlossen.

Die Überlegungen gingen eher in Richtung einer *Unterschätzung* der Störungskosten durch vergessene Behandlungen. Eine solche Unterschätzung schien zudem für den Zeitraum vor Therapie für wahrscheinlicher als für die Befragung über das Jahr nach der Therapie (z.B. da die Patienten bei der 1-Jahres-Nachuntersuchung die Fragen nach Behandlungen bereits kannten, mit dem Therapieende einen klareren Anker hatten als „jemals“ oder „in den letzten zwei Jahren“). Falls also Gedächtniseffekte eine Rolle spielen, so eher in einem konservativen Sinne, indem die Einsparungen nach Therapie unterbewertet werden. Aufgrund der Problematik von Gedächtniseffekten wurden nur die Angaben für die zwei Jahre vor Therapie ausgewertet, und nicht die Angaben zum gesamten Störungszeitraum (der durchschnittlich etwa neun Jahre betrug).

Ferner wurde überprüft, ob es bedeutsame Unterschiede hinsichtlich der Erhebungsmethodiken Interview vs. Fragebogen gab bzw. ob die deutlich effizientere Fragebogenmethode dem aufwendigen, sorgfältig individuell angepassten Interview vorgezogen werden konnte. Die Vergleiche zwischen Fragebogen und Interviews, in denen die Patienten noch einmal detailliert befragt

wurden, um die Erinnerungsleistung zu verbessern, ergaben keine bedeutsame Unterschiede.

Somit scheint eine hinreichende Validität des eingesetzten Fragebogens (bei leichter Tendenz zur Unterschätzung des Psychotherapienutzens) gegeben zu sein, so dass man damit eine Aufstellung über angefallene Kostenfaktoren vor und nach Therapie vornehmen kann.

4.2.3 Nutzenberechnung II: Umrechnung der Kostenfaktoren in monetäre Einheiten

Die Zuordnung von Geldwerten zu den einzelnen Posten soll im folgenden beschrieben werden. Alle genutzten Preise beziehen sich auf ein einziges Jahr, um einen Einfluss der Inflation auf die Ergebnisse auszuschließen (vgl. Warner und Luce, 1986). Da das Jahr 1996 zum Zeitpunkt der Berechnung das letzte Jahr war, für das beim statistischen Bundesamt Angaben für alle Posten vorlagen, fiel die Wahl auf dieses Jahr.

1. Ambulante Arztkosten und Psychotherapien

Die Kosten für ambulante Arztkontakte und Psychotherapiesitzungen konnten nicht exakt ermittelt werden. Zunächst war geplant, für eine Teilstichprobe eine genaue Abrechnung über einen längeren Zeitraum bei einem Kostenträger zu ermitteln, um eine empirisch gestützte Schätzgrundlage zu erhalten. Dies stellte sich aber (trotz mündlicher Zusage einer örtlichen Krankenkasse zum Zeitpunkt des Projektbeginns) aus datenschutzrechtlichen Gründen als unmöglich heraus (vgl. SGB V, §§304, 305). Selbst als Versicherter hat man über seine eigenen Leistungsdaten nur Zugriff auf das letzte Geschäftsjahr. Außerdem werden bei Primär- und Ersatzkassen die Abrechnungsdaten nicht patientenbezogen, sondern Arzt- und Apothekenbezogen gespeichert. Die einzige Möglichkeit einer exakten Erhebung wäre die Kooperation mit einer Betriebskrankenkasse (vgl. Zielke, 1993) oder einer privaten Krankenkasse gewesen. Die Privatversicherten oder Angehörige einer Betriebskrankenkasse machten aber in der uns verfügbaren Stichprobe nur einen verschwindend geringen Anteil aus.

Ein Wert für z.B. „den durchschnittlichen Preis eines Arztbesuches“ war weder bei anderen Krankenkassen, der Kassenärztlichen Bundesvereinigung oder in gesundheitsökonomischen Zentren zu erfahren. Daher mussten die entsprechenden Schätzwerte selbst gebildet werden. Dabei wurde auf statistische Angaben der kassenärztlichen Bundesvereinigung und des Statistischen Bundesamtes zurückgegriffen.

Folgende Werte werden nun zur Berechnung der Kosten ambulanter Gesundheitsleistungen (1996) eingesetzt:

Tabelle 4.1: Kosten für ambulante Behandlungen (in DM)

Psychotherapiesitzung	120,00
Facharztbesuch	60,00
Hausarztbesuch	30,00
Ruf des Notarztes	400,00

Da diese Werte nicht exakt sind, sollen sie einer Sensitivitätsanalyse unterzogen werden (Drummond et al., 1986). Wie sich zeigte, verändert sich die Kosten-Nutzen-Relation nicht, auch wenn die tatsächlichen Werte 50% oder 150% dieser Schätzwerte betragen würden (s. 4.4.4). Wenn aus einer neuen Quelle neue, exakte Zahlen zu erfahren wären oder im Gesundheitssystem ein neues Preisgefüge entsteht (wie z.B. derzeit nach Inkrafttreten des Psychotherapeutengesetzes), ließe sich zudem die Kosten-Nutzen-Analyse leicht neu berechnen – auch wenn sich wohl kaum etwas substantiell ändern würde.

2. Stationäre Kosten

Die stationären Kosten mussten für die Psychiatrie, die Psychosomatik, die Reha-Klinik und die Kurklinik bestimmt werden. Krankenhauskosten setzen sich zusammen aus dem Basispflegesatz, der in allen Kliniken derselbe ist und aus dem Abteilungspflegesatz. Diese Daten standen lediglich für die Psychiatrie zu Verfügung.

Der tägliche Pflegesatz (Basispflegesatz plus Abteilungspflegesatz) betrug im Jahr 1996 392,43 DM (Angabe des Verbandes der privaten Krankenversicherungen, schriftliche Mitteilung vom 12.06.99).

Für die anderen genannte Klinikarten gibt es keine bundesweiten einheitlichen Pflegesätze. Nach Angaben des Verbandes deutscher Rentenversicherungsträger schwanken die Tagessätze zwischen unter 100 DM und über 700 DM am Tag, je nach Klinik, da die unterschiedlichsten Patienten in einzelnen Kliniken behandelt werden (z.B. verursacht ein Patient in der Intensivstation mehr Kosten als ein psychisch Kranker). Daher wurden die Preise der einzelnen von den Patienten belegten Kliniken (Rehaklinik, Kurklinik und Psychosomatik) ermittelt und ein Durchschnittspreis daraus berechnet. Die Tagessätze für jede konkret von den Patienten genannte Klinik (also Klinik X in Bad Y etc.) wurde beim Verband der privaten Krankenversicherungen in Köln für das Jahr 1996 erfragt. Aus den Angaben von 14 Patienten konnte nicht entnommen werden, um welche Klinik es sich handelte, für 42 konkrete Kliniken aber liegt der Preis vor. Ebenso fielen manche Kliniken nicht in die oben genannten Rubriken bzw. es konnte aus den Angaben nicht entnommen werden, um wel-

che Klinikart es sich handelte. Diese fielen unter „Sonstige“, und die Kosten wurden mit dem Durchschnittssatz aller Kliniken einschließlich des Pflegesatzes der Psychiatrie berechnet. Da es sich ausschließlich um Patienten mit derselben Symptomatik handelt, scheint dies gerechtfertigt.

Die Tagesanzahl des Aufenthalts wurde jeweils mit diesen Preisen verrechnet. Es wurde davon abgesehen, die tatsächlich entstandenen Kosten pro Patienten zu berechnen (also Patient X in Klinik Y für Z Tage), da dies zu Verfälschung des Verhältnisses Prä/ Post führen kann. Die Klinikpreise variieren so beträchtlich, dass eine zufällige Akkumulation der besonders hohen oder besonders niedrigen Preise in der Post-Stichprobe, in der nur N=8 Patienten einen Klinikaufenthalt hatten, zu einer Unter- bzw. Überschätzung des Wertes führen könnten.

Die errechneten Durchschnittswerte für die einzelnen Kliniktypen betragen:

Tabelle 4.2: Preise der unterschiedlichen Kliniktypen (in DM)

Psychiatrie	392,43
Psychosomatik	271,38
Rehaklinik	252,15
Kurklinik	183,56
Sonstige	274,88

3. *Kosten der Arbeitsunfähigkeit*

Die Kosten des Arbeitsausfalls setzen sich aus unterschiedlichen Werten zusammen. Sie bestehen bis zum 42. Tag der Krankschreibung im Produktivitätsausfall, ab dem 42. Tag im Krankengeld und im Lohnsteuerausfall. Nach Auskunft des Deutschen Institutes für Wirtschaftsforschung kann man die entgangene Arbeitsproduktivität generell als Maß für den volkswirtschaftlichen Schaden von Arbeitsunfähigkeit nutzen. Nach Auskunft des statistischen Bundesamtes lag 1996 die Arbeitsproduktivität pro Erwerbstätigenstunde bei 56,00 DM (Bruttoinlandsprodukt je Erwerbstätigenstunde). Laut Institut für Arbeitsmarkt - und Berufsforschung kann man als Schätzwert der durchschnittlichen Arbeitstätigkeit pro Tag die tarifliche Wochenarbeitszeit nutzen und durch fünf dividieren. Die tarifliche Wochenarbeitszeit ist in West- und Ostdeutschland unterschiedlich.

Tabelle 4.3: Tarifliche Wochenarbeitszeit in Ost- und Westdeutschland

	West	Ost
Vollzeit	37,56 h	39,5 h
Teilzeit	17,42 h	22,87 h

Unter Berücksichtigung des Anteils der Herkunft aus Ost und West und der Vollzeit und Teilzeit Arbeitenden bzw. der Erwerbslosen ergab sich ein gewichteter Mittelwert des volkswirtschaftlichen Schadens pro Tag Arbeitsunfähigkeit von 187,91 DM (über alle Patienten hinweg).

Eine weitere Möglichkeit der Berechnung besteht darin, nur die erwerbstätigen Probanden zu berücksichtigen. Aus dem Anteil der Erwerbstätigen der hier untersuchten Stichprobe wird die prozentuale Anzahl der Teil- und Vollzeitarbeitenden und die Herkunft aus Ost oder West ermittelt. Hier ergibt sich ein gewichteter Mittelwert von 394,77 DM als Maß für den volkswirtschaftlichen Schaden eines Tages Arbeitsunfähigkeit pro Erwerbstätigem.

4.2.4 Sensitivitätsanalyse

Die Sensitivitätsanalyse überprüft den Einfluss der geschätzten Parameter auf das Kosten-Nutzen-Verhältnis. Ein Unsicherheitsfaktor sind die Kosten für ambulante Behandlungen. Da es hierfür keine exakten Angaben gibt, muss auch hier eine Sensitivitätsanalyse durchgeführt werden. Anhand von unverhältnismäßig hohen bzw. geringen Preisen soll untersucht werden, ob sich das gefundene Kosten-Nutzen-Verhältnis ändert. Ebenso sind die Kosten der Arbeitsunfähigkeit, die anhand der bundesweiten Durchschnittswerte berechnet werden, in die Analyse einzubeziehen. Da keine anderen Angaben vorliegen, wird die Sensitivitätsanalyse jeweils mit 50% bzw. 150% der zunächst genutzten DM-Werte durchgeführt.

4.2.5 Diskontierung

Bei der Diskontierung wurden die Kosten, die in verschiedenen Jahren anfallen, auf ein Bezugsjahr relativiert (vgl. 2.1). Der gleiche Geldwert ist in der Zukunft nicht so viel wert wie in der Gegenwart, da er sich im Laufe der Zeit durch Verzinsung hätte steigern können. Der Nutzen der Therapie stellte sich erst später ein als die Kosten. Somit waren diese Geldwerte nicht direkt vergleichbar. Auf eine Diskontierung der Therapiekosten konnte verzichtet werden, da diese nur dann bereinigt werden müssen, wenn sie über mehrere Jahre

anfallen. Die Therapien der Patienten dauerten aber niemals länger als ein Jahr.

Für Nutzaussagen zu Zeiträumen, die länger sind als ein Jahr musste hingegen eine Diskontierung vorgenommen werden. Der nichtdiskontierte Nutzen berechnete sich dabei als Differenz der Kosten X Jahre vor der Therapie abzüglich der Kosten X Jahre nach der Therapie. Die Kosten vor der Therapie errechneten sich hierfür aus der Anzahl der Jahre multipliziert mit den ermittelten Gesamtkosten im Jahr vor der Therapie. Die Gesamtkosten im Jahr vor der Therapie wurden als Hälfte der Gesamtkosten in den letzten 2 Jahren vor der Therapie bestimmt (unter der Annahme, dass sich die Kosten in diesen zwei Jahren gleich verteilten).

Die Berechnung des diskontierten Nutzens (N_x) erfolgte auf der Grundlage des Nutzens nach einem Jahr (N) für bis zu 5 Jahren nach der Therapie. Die Diskontierung erfolgte mit folgender Formel:

$$N_x = N + \frac{N}{(1+r)} + \dots + \frac{N}{(1+r)^{x-1}}$$

Die Wahl der Diskontierungsrate (r) fand nach dem Vorschlag der Konsensgruppe „Gesundheitsökonomie“ (1996) statt und wurde auf 5% festgelegt.

4.2.6 Marginalanalyse

In der Marginalbetrachtung ist nach zusätzlichen kosten- oder nutzenverursachenden Faktoren zu suchen, die bei der Berechnung der „großen“ Kostenfaktoren mit Durchschnittswerten keine Rolle spielten. Diese hinzukommenden „kleinen“ Beträge sind dabei genau zu ermitteln; die Verwendung von Durchschnittswerten ist laut der Konsensgruppe „Gesundheitsökonomie“ (1996) nur dann erlaubt, wenn keine konkreten Marginalwerte zur Verfügung stehen. In der vorliegenden Arbeit konnten für die gefundenen zusätzlich anfallenden Kosten und Nutzen weder genaue Einzelwerte noch Durchschnittswerte ermittelt werden. Im Ergebnisteil werden die zusätzlichen Posten daher lediglich beschrieben. Die damit entstehende Ungenauigkeit wurde in der folgenden Sensitivitätsanalyse kalkuliert.

4.3 Rahmenbedingungen und Stichprobe

Die Rahmenbedingungen des Therapiesetting sind die selben wie bei der in Teil 3 dargestellten Kosten-Effektivitäts-Analyse (s. 3.3.1).

Die Kosten-Nutzen-Analyse wurde längsschnittlich durchgeführt, d.h. es wurden nur diejenigen Patienten aufgenommen, für die vor Therapie *und* zur 1-Jahres-Katamnese Kostendaten vorlagen (N=67)¹⁴.

4.3.1 Soziodemografische Variablen

Zur Stichprobenbeschreibung werden in folgender Tabelle 4.4 die Verteilung in den Variablen Alter, Geschlecht, Bildungsstand und ausgeübter Beruf dargestellt.

Die soziodemografischen Daten entsprechen im wesentlichen denen der Gesamtstichprobe der Patienten, die in der Christoph-Dornier-Stiftung behandelt wurden; einziger bedeutsamer Unterschied ist der erhöhte Anteil Arbeitsloser in dieser Stichprobe (vgl. Tab. 4.3). Da die Arbeitsunfähigkeit einen bedeutenden Kostenfaktor darstellt, muss diese Tatsache gesondert diskutiert werden.

¹⁴ In Abschnitt 4.6.1 werden die so erlangten Ergebnisse mit querschnittlich erhobenen Daten (N=213 prä vs. N=205 1-J-Kat.) verglichen. In der Diplomarbeit von Pieper (1999) wurde analog dem Vorgehen in der Kosten-Effektivitäts-Studie gezeigt, dass keine Selektionseffekte dahingehend gefunden werden können, dass sich die Population der Patienten, für die Nachuntersuchungsdaten vorliegen von der ohne Nachuntersuchungsdaten unterschieden.

Tabelle 4.4: Soziodemographische Stichprobendaten (N=67)

Alter	$\bar{x} = 35,9$	$s = 10,8$
Geschlecht	männlich:	37 %
	weiblich:	63 %
höchster Schulabschluß	keiner:	3,1 %
	Hauptschule:	21,5 %
	Realschule:	16,9 %
	Fachabitur/Abitur:	30,8 %
	abgeschlossenes Fach-/Hochschulstudium:	27,7 %
derzeitige Berufsgruppe	Arbeiter:	1,6 %
	Facharbeiter/Handwerker:	11,5 %
	Angestellter/Beamter einfacher/mittlerer Dienst:	16,4 %
	gehobener Dienst:	9,8 %
	Selbständiger:	9,8 %
	Schüler/Azubi/Student:	21,3 %
	Hausfrau:	8,2 %
	arbeitslos:	18 %
	berentet:	3,3 %

4.3.2 Störungsspezifische Variablen

In der folgenden Tabelle 4.5 wird ein Überblick über die folgenden störungsspezifischen Variablen gegeben: Primärdiagnose, Anzahl weiterer Diagnosen, Krankheitsdauer, Schweregrad der Primärdiagnose.

Tabelle 4.5: Störungsspezifische Stichprobendaten (N=67)

Verteilung der Primär Diagnosen (Angststörungen; N=67)	Agoraphobie (mit Panikstörung):	53,7 %
	Soziale Phobie:	23,9 %
	Zwangsstörung:	14,9 %
	sonst.:	7,5 %
	<hr/>	
Anzahl weiterer Diagnosen: 41	davon	
	weitere Angststörung: affektive Störung:	44 %
	sonst. psych. Störung:	10 %
<hr/>		
Krankheitsdauer	Mittelwert:	8,9 Jahre (s = 9,7)
	Median:	5 Jahre
<hr/>		
Schweregrad der Primär diagnose (Skala von 0 = "keine Störung bis" 8 = "sehr stark störend")	Mittelwert:	6,1 (s = 1,1)
	Median:	6

Der Schwerpunkt bei den Diagnosen liegt ebenso wie bei der Gesamtstichprobe bei der Agoraphobie mit Panikstörung (vgl. Tab. 3.4). Die Krankheitsdauer und der (von den Diagnostikern eingeschätzte) Schweregrad weisen ähnlich der Gesamtstichprobe darauf hin, dass es sich um eine klinisch relevante Stichprobe mit schweren Angststörungen handelt.

4.4 Ergebnisse

4.4.1 Der Nutzen aus der Perspektive volkswirtschaftlicher Einsparungen

4.4.1.1 *Kosten vor und nach Therapie*

Die folgenden Ergebnisse beruhen auf dem Vergleich der Krankheitskosten 2 Jahre vor vs. 1 (erhoben) bis 5 (hochgerechnet) Jahre nach Therapie. Aus methodischen Gründen wurden die Krankheitskosten, die länger als zwei Jahre zurücklagen, nicht berücksichtigt (Gedächtniseffekte, unklare Bestimmung des genauen Krankheitsbeginns; vgl. die vollständigen Auswertungen bei Pieper, 1999 und Runde, 1999). Eine Ausnahme bilden die stationären Aufenthalte, für die vor Therapie die gesamte Krankheitsdauer herangezogen wurde (da davon ausgegangen wurde, dass durch die Störung verursachte stationäre Aufenthalte nicht vergessen wurden und dass zum Zeitpunkt eines stationären Aufenthaltes die Störung bereits definitiv vorhanden gewesen sein muss).

Die folgende Tabelle 4.6 gibt Aufschluss über die Kosten in den letzten 2 Jahren vor Therapiebeginn. Es wird dort zunächst aufgeführt, bei wie vielen Patienten der jeweilige Kostenfaktor überhaupt anfiel. Für die Kosten-Nutzen-Analyse ist der Gesamtdurchschnitt der angefallenen Einheiten über alle Patienten relevant. Um eine bessere Vorstellung zu vermitteln, wie umfangreich die angefallenen Kosten waren, wenn der jeweilige Kostenfaktor einmal anfiel, wird der durchschnittliche Umfang in der 3. Spalte gesondert aufgeführt. So fielen z.B. durchschnittlich 29,4 Psychotherapiesitzungen an, sofern ein Patient eine Psychotherapie in Anspruch genommen hatte. Bezogen auf alle Patienten (also auch auf die ohne Psychotherapie) betrug der Durchschnittswert 8,86. Die Arbeitsunfähigkeit wird nur für die Hälfte der Patienten dargestellt, da die andere Hälfte nicht berufstätig war. Dieser Aspekt wird unter 4.4.2 im Abschnitt zu Marginalkosten noch einmal aufgegriffen.

Ein erwerbstätiger Patient der Untersuchungsstichprobe verursachte durchschnittlich in den 2 Jahren vor der Therapie ca. DM 9910,- Kosten, die der Volkswirtschaft zu Lasten fielen. Umgeschlagen auf die gesamte Stichprobe (incl. der Erwerbslosen) beträgt dieser Wert DM 7540,- .

Unter der Annahme der Gleichverteilung in den letzten beiden Jahren kann man also von durchschnittlichen DM 3770,- krankheitsbezogener Ausgaben pro Person im Jahr vor der Therapie ausgehen (Erwerbstätige: 4955,-).

Im Jahr nach der Therapie waren die krankheitsbezogenen Ausgaben deutlich gesunken. 25 Patienten (38%) verursachten überhaupt keine Kosten in dieser Zeit. Vor der Therapie hatten nur 6 Personen (9%) keinerlei Ausgaben hervorgerufen. Doch auch im Katamnesezeitraum schwankten die Kosten,

wenn auch nicht so stark wie vor der Therapie. Tabelle 4.7 informiert über Kostenfaktoren und Kosten im Jahr nach der Therapie. Im Jahr nach der Therapie fielen pro Person noch DM 665,- volkswirtschaftliche Kosten an (Erwerbstätige: DM 687.-).

Tabelle 4.6: Krankheitskosten je Patient (N=67) für die 2 Jahre vor der Therapie

Kostenfaktoren (KF)	% der Fälle, bei denen jeweiliger KF anfiel	Mittelwert der Anzahl der Einheiten des jeweiligen KF (wenn KF anfiel)	Mittelwert der Anzahl der Einheiten des jeweiligen KF (über alle Patienten)	DM (über alle Patienten)
Psychotherapie sitzungen	29,3	29,4	8,86	1063,20
Kontakte mit Allgemeinarzt aufgrund Störung	31,7	16,8	5,32	159,60
Kontakte mit Facharzt aufgrund Störung	54,0	15,3	8,25	495,00
Kontakte mit Notarzt aufgrund Störung	12,5	2,5	0,31	124,00
stationäre Aufenthalte (Tage)¹⁵ aufgrund Störung	24	89,3	21,6 (umgerechnet auf 2 Jahre: 4,8)	1443,50
AU-Tage (bei Erwerbstätigkeit) aufgrund Störung	55,8	35,44	19,78	6626,30 (für gesamte Stichprobe: 4252,70)

¹⁵ Bei den stationären Aufenthalten wurde die gesamte Krankheitsdauer herangezogen

Tabelle 4.7: Krankheitskosten je Patient (N=67) für das Jahr nach der Therapie

Kostenfaktoren (KF)	% der Fälle, bei denen jeweiliger KF anfiel	Mittelwert der Anzahl der Einheiten des jeweiligen KF (wenn KF anfiel)	Mittelwert der Anzahl der Einheiten des jeweiligen KF (gesamt über alle Patienten)	DM (über alle Patienten)
Psychotherapie sitzungen	7,5	7,8	0,58	69,60
Kontakte mit Allgemeinarzt	17,9	1,17	0,21	6,30
Kontakte mit Facharzt	7,5	2,2	0,16	9,60
Kontakte mit Notarzt	0	0	0	0
stationäre Auf- enthalte (Ta- ge) ¹⁶	4,5	40	1,79	538,30
AU-Tage (bei Erwerbstätig- keit)	2,3	13	0,19	63,60 (für gesamte Stichprobe: 40,90)

4.4.1.2 Nutzen als Reduktion von Krankheitskosten

Tabelle 4.8 zeigt eine Übersicht der Reduktion der Krankheitskosten vom Katamnese- verglichen mit dem Anamnesezeitraum. Die Einsparungen betragen insgesamt rund DM 3103.- im ersten Jahr (für die gesamte Stichprobe, auf die der Nutzen bei den Erwerbstätigen umgelegt wurde).

Es handelte sich bei der Verringerung der Kosten in den einzelnen Bereichen um überzufällige Schwankungen, d.h. man kann von einer wirklichen Einsparung ausgehen (Wilcoxon-Tests, da Voraussetzungen für t-test bei abhängigen Stichproben nicht erfüllt). Bei allen Kostenfaktoren-Vergleichen waren die positiven Ränge geringer als die negativen Ränge, d.h. es kam in den Einzelvergleichen grundsätzlich seltener vor, dass zum Zeitpunkt Post höhere Kosten anfielen als zum Zeitpunkt Prä.

Der Nutzen berechnete sich als Differenz aus den Kosten eines festgelegten Zeitraumes vor und nach der Therapie. Der absolute Nutzen betrug 1 Jahr nach der Therapie 3026,- DM. Diese Kosteneinsparung wird weiter unten den Therapiekosten gegenübergestellt und ein Kosten-Nutzen-Index ermittelt.

¹⁶ Bei den stationären Aufenthalten wurde die gesamte Krankheitsdauer herangezogen

Tabelle 4.8: Reduktion der Krankheitskosten bei N=67 Patienten nach Therapie (Standardisiert auf ein Jahr)

Kostenfaktoren (KF)	% Reduktion der Einheiten des jeweiligen Kostenfaktors	Reduktion DM (über alle Patienten standardisiert auf 1 Jahr)
Psychotherapiesitzungen	86,9	462,00
Kontakte mit Allgemein- arzt	92,1	73,50
Kontakte mit Facharzt	96,1	236,00
Kontakte mit Notarzt	100	62,00
stationäre Aufenthalte (Tage)¹⁷	25,4	183,50
AU-Tage (bei Erwerbstätigkeit)	98,1	2085,50

Keine der untersuchten Behandlungen dauerte länger als 9 Monate. Das heißt, alle Therapiekosten fielen innerhalb eines Jahres an. Es war also nicht nötig diese Kosten zu diskontieren. Der ermittelte Nutzen betrug 3103,- DM im Jahr nach der Therapie.

Für eine Aussage über einen längeren Zeitraum musste dieser Nutzen mit einer Diskontierungsrate von 5% pro Jahr relativiert werden. In der folgenden Tabelle 4.9 ist der hochgerechnete Nutzen über längere Zeiträume bis 5 Jahre aufgeführt.

Tabelle 4.9: Diskontierter Nutzen über größere Zeiträume (in DM)

Bezugszeitraum	Nutzen	diskontierter Nutzen
1 Jahr	3103,-	3103,-
2 Jahre	6205,-	6057,-
3 Jahre	9307,-	8871,-
4 Jahre	12410,-	11551,-
5 Jahre	15512,-	14103,-

¹⁷ Bei den stationären Aufenthalten wurde die gesamte Krankheitsdauer herangezogen

Auch wenn bislang ein Mangel an Langzeitkatamnesen in der Psychotherapieforschung zu verzeichnen ist, besteht doch Grund zu der Annahme, dass Therapieerfolge zumindest bis fünf Jahre nach Therapie stabil bleiben oder sich sogar weiter verbessern (z.B. Grawe et al. 1994; Grawe, 1999). Dies bezieht sich jedoch grundsätzlich nur auf den *Durchschnitt*, der dadurch zustande kommt, dass sich manche erfolgreiche Patienten weiter verbessern, einige ihren Erfolg halten und wiederum andere sich wieder verschlechtern bzw. auch bereits während der Therapie weniger Erfolg hatten.

4.4.1.3 Kostenverursachung: reduzierte Fälle und reduzierte Gesamtkosten

Die Reduktion der Inanspruchnahme von Gesundheitsleistungen sowie die der Arbeitsunfähigkeit ist in allen Bereichen deutlich festzustellen.

Zu unterscheiden ist die Zahl derer, die einen Kostenfaktor verursachen und der Umfang der verursachten Kosten. Die anfallenden Kosten wurden im Jahr nach der Therapie von bedeutend weniger Patienten produziert als vor der Therapie. 7,4% der Patienten verursachten 90% der Kosten im Jahr nach Therapie – gegenüber 29,7% zum Zeitpunkt Prä. Auch diese Reduktion der Varianz spricht für den relativ durchgängigen und wenig differentiellen Effekt der Therapie.

Zudem produzierten diejenigen, für die auch nach der Therapie noch Kosten anfielen, durchschnittlich weniger Kosten pro Kostenfaktor als vor Therapie. Einzige Ausnahme sind die stationären Aufenthalte: die Zahl derer, die nach Therapie stationär behandelt wurden, reduzierte sich zu vorher um 81%, die Gesamtkosten für diesen Posten jedoch nur um 25,4%. Die wenigen, die trotz störungsspezifischer Therapie noch einmal stationär behandelt wurden, hatten pro Fall also relativ große Kosten zu verbuchen. Dies könnte darauf zurückzuführen sein, dass es denjenigen, die nach der Therapie tatsächlich noch Behandlung in Anspruch nahmen, auch immer noch besonders schlecht ging.

4.4.2 Weitere Posten

Es können mehrere zusätzliche kosten- und nutzenverursachende Faktoren gefunden, aber leider nicht in Geldeinheiten umgerechnet werden. Für die Monetarisierung wird eigentlich eine Berechnung mit konkreten Werten gefordert, und nur ersatzweise dürfen gut begründete Durchschnittswerte verwendet werden. Da beides nicht zur Verfügung stand, sollen die Posten hier lediglich genannt werden.

1. Positiver Nutzen durch gestiegene Erwerbstätigkeit

Beim Vergleich der Erwerbstätigkeit vorher und nachher fällt auf, dass ein Jahr nach der Therapie mehr Personen einer Arbeit nachgingen als zum Beginn der Therapie¹⁸:

Tabelle 0.4: Veränderung der Erwerbstätigkeit von Prä nach Post

	vorher erwerbstätig	nicht erwerbstätig
nachher erwerbstätig	26	6
nachher nicht erwerbstätig	1	15

6 der 48 Patienten (12,5%) waren demnach vor der Therapie nicht erwerbstätig, hatten nach der Therapie aber (wieder) eine Arbeit aufgenommen, während lediglich ein vorher erwerbstätiger dies nachher nicht mehr war (diese Patientin hatte in der Zwischenzeit ein Studium aufgenommen und dafür ihre Selbständigkeit aufgegeben).

Ob diese Veränderung des beruflichen Status` durch die Therapie bedingt wurde, konnte allerdings nicht in dieser Untersuchung geklärt werden. Eine Prüfung des vorher-nachher-Unterschiedes mit dem McNemar-Unterschiedstest für abhängige Stichproben wurde nicht signifikant ($p = 0,13$). Aus diesen Gründen wurde der volkswirtschaftliche Nutzen, der mit einer gesteigerten Erwerbstätigenquote einhergeht, nicht quantifiziert und in die Kosten-Nutzen-Analyse einbezogen. Nichtsdestotrotz kann erwartet werden, dass hier bei sorgfältigerer ökonomischer Untersuchung noch ein nicht unbedeutender weiterer Nutzen zutage gefördert werden könnte.

2. Medikamentenkonsum und privat bezahlte Heilmittel

Ein weiterer nichtquantifizierter Posten betrifft die Medikamente, welche die Patienten aufgrund ihrer Beschwerden eingenommen haben. Auch für diesen (in der Literatur als eher gering bezeichneten) Kostenfaktor war es weder möglich, einen Durchschnittspreis pro Einheit zu ermitteln, noch den Umfang der verbrauchten Medikamente im Einzelnen valide zu ermitteln. Es wurde

¹⁸ Der Vergleich beruht auf den Angaben von 48 Personen (von den anderen 19 fehlten die Daten zum Erwerbstätigkeitsstatus zum Katamnesezeitpunkt).

daher nur grob analysiert, wie sich der Medikamentenkonsum (privat oder von der Kasse bezahlt) veränderte. Von 62 Patienten, bei denen die Angaben zum Medikamentenkonsum für Prä und Post vorlagen, hatten 40 während ihrer gesamten Krankheit schon einmal Medikamente dagegen eingenommen. 29 von ihnen konnten im Jahr nach der Therapie darauf verzichten. Der Effekt war hochsignifikant auf dem Niveau 0,00 (McNemar-Test für abhängige Stichproben). Ebenfalls gesunken ist der Umfang privat bezahlter Heilmittel oder Behandlungen (z.B. Akupunktur), was aber aufgrund der Heterogenität der Daten nicht in Geld quantifiziert werden konnte.

3. Berentung

Eine Berentung aufgrund der Angststörung, derentwegen die Therapie aufgesucht wurde, bedeutet einen enorm hohen volkswirtschaftlichen Verlust, da eine Arbeitskraft dem Markt komplett verloren geht. Je jünger die Person dabei ist, umso höher sind, bei dauerhafter Berentung, die Kosten für die Gesellschaft. In der Stichprobe war ein 30-jähriger Mann die letzten zwei Jahre vor der Therapie komplett arbeitsunfähig und deshalb zeitweilig berentet worden. Durch die Therapie konnte er laut eigener Angabe hinterher wieder einer Arbeit nachgehen. Der damit einhergehende immense finanzielle Nutzen wurde nicht mit in die Kosten-Nutzen-Analyse einbezogen, weil es auch hier wieder zu schwierig war, diesen genau zu quantifizieren. Außerdem handelte es sich um ein Einzelschicksal, dessen vermutlich extrem hoher positiver Nutzen den Gesamtdurchschnitt hätte verzerren können. Geht man einmal nur von der entgangenen Arbeitsproduktivität aus (und vernachlässigt also die Rentenzahlungen), dann könnten die Kosten für das Jahr vor der Therapie genauso ermittelt werden wie bei einem Arbeitsunfähigkeitstag eines Erwerbstätigen. Die Ausgaben beliefen sich demnach auf 87 100,- DM. Die Ausgaben für das Jahr nach der Therapie waren nicht so leicht zu berechnen, da nicht bekannt war, seit wann der Mann wieder arbeiten gehen konnte, nur dass er 12 Monate nach Therapieende wieder erwerbstätig war. Somit konnte kein konkreter finanzieller Nutzen bestimmt werden; der Kostenfaktor Berentung wurde deshalb aus der Analyse ausgeschlossen.

4. Unbezahlter Urlaub

Ein vierter Kostenfaktor, der zwar erhoben aber nicht in die Analyse einbezogen wurde, betrifft unbezahlte Urlaubstage, welche Patienten aufgrund ihrer Krankheit nahmen. Vier Personen gaben an, aufgrund der Erkrankung im unbezahlten Urlaub gewesen zu sein; und zwar 2, 4, 10 und 20 Tage. Diese Position wurde allerdings aus der Berechnung eines Kosten-Nutzen-Indexes aus-

geschlossen, da sie Post nur an einer kleinen Substichprobe erhoben wurde. Im Sinne einer konservativen Vorgehensweise wurden die Kosten auch deswegen nicht in die deskriptive Betrachtung der Gesamtkosten vor der Therapie einbezogen, weil auch hier die Monetarisierung schwer fiel. Wenn man, wie bei der Arbeitsunfähigkeit, mit dem entgangenen Bruttosozialprodukt rechnen würde, käme es in diesem Fall vielleicht eher zu einer Überschätzung der Kosten. Es ist z.B. denkbar, dass, wenn ein Arbeitgeber unbezahlten Urlaub genehmigt, in dieser Zeit ohnehin nur ein geringes Auftragsvolumen bestanden hat, so dass nicht wirklich von einem Verlust am Sozialprodukt zu sprechen ist.

5. Sonstige Posten

Von den bis jetzt genannten zusätzlichen kosten- und nutzenverursachenden Faktoren war bekannt, dass sie aufgetreten sind. Da Schätzungen in diesem Fall zu ungenau gewesen wären, mussten die Positionen aus der Kosten-Nutzen-Analyse ausgeschlossen werden; dies entspricht einer konservativen Vorgehensweise.

Weitere Faktoren, die nicht erhoben wurden, aber ebenfalls krankheitsbezogene Kosten bedeuten, sind denkbar. So wurde zwar die Arbeitsunfähigkeit erhoben, aber nicht eine, durch krankheitsbedingtes Unwohlsein hervorgerufene, qualitätsgeminderte Arbeitsleistung. Auch durch einen derart reduzierten Arbeitsausfall entstehen Verluste am Sozialprodukt. Ebenso wurde nicht erhoben, wie viel Zeit von der Arbeitszeit durch die Vorbehandlungen verloren ging.

In einer noch weitergehenden Perspektive könnten außerdem zusätzlich die volkswirtschaftlichen Kosten, die andere Personen für diesen Patienten verursachten (z.B. Arbeitsausfall wegen Betreuung des Patienten) erhoben werden.

Nicht einbezogen, weil mit unseren Mitteln nicht quantifizierbar, ist der Nutzen durch verringerte Arbeitsunfähigkeit bei nicht erwerbstätigen. Gemäß dem Opportunitätskostenprinzip (vgl. 1.2) ist auch unbezahlte Arbeit (z.B. Hausfrau) für eine Gesellschaft nicht umsonst, da sie vielleicht an einem anderen Ort ebenfalls nutzenbringend hätte eingesetzt werden können. Es entsprach auch hier einem sehr konservativen Vorgehen, die von nicht erwerbstätigen angegebenen Arbeitsunfähigkeitszeiten aus der Kosten-Nutzen-Analyse auszuschließen und lediglich den durch die Erwerbstätigen erzielten Nutzen auf die Gesamtstichprobe umzulegen.

Auch das Verhindern von Brüchen in der beruflichen Biografie oder erweiterte Karrieremöglichkeiten nach einer Therapie stellen bedeutsame volkswirtschaftliche Nutzenfaktoren dar, die in einzelnen Fällen vorlagen, aber nicht monetär bewertet werden konnten.

6. Weitere Kostenfaktoren der untersuchten Behandlung

Über den Betriebskostenansatz fließen die relevanten aufgewendeten Ressourcen durch die Behandlungseinrichtung in die Kostenseite ein. Nicht berücksichtigt werden konnten Nebenkosten (z.B. für Anfahrt, Übernachtungen, Kosten von Konfrontationsübungen etc.) und indirekte intangible Kosten für die Patienten und ggf. deren Umfeld (z.B. Zeitaufwand für die Behandlung). Diese Marginalkosten sind monetär schwer zu beziffern und hätten zudem dann auch für die Störungskostenfaktoren ermittelt werden müssen; der dadurch zu erwartende Erkenntnisgewinn hätte den dadurch entstehenden Aufwand nicht gerechtfertigt. Sie werden aber bei der Einbeziehung der Zahlungsbereitschaft (4.5) indirekt berücksichtigt, da die Patienten dort alle für sie kostenverursachenden Aspekte einer Behandlung einbeziehen sollen..

4.4.3 Kosten-Nutzen-Index

Unter Kosten werden die in Rechnung gestellten Kosten der Therapie verstanden. Der durchschnittliche Therapiepreis betrug DM 5371.- (vgl. 4.2.1).

Als Nutzen gingen die Einsparungen infolge reduzierter Arbeitsunfähigkeit und verminderter Inanspruchnahme von Behandlungen ein, die in den vorhergehenden Abschnitten berechnet wurden. Da die Gesamteinsparungen im ersten Jahr nach der Therapie DM 3103.- betragen, hat sich die Therapie unter dieser Perspektive und unter Vernachlässigung weiterer direkter und indirekter Nutzenfaktoren bereits nach zwei Jahren amortisiert.

Die gängige Form der Präsentation der Ergebnisse sind der *Kosten-Nutzen-Index* und der *Netto-Nutzen*, der stets mit angegeben werden soll, damit nicht nur das Verhältnis, sondern auch das Ausmaß bewertet werden kann. Der Kosten-Nutzen-Index berechnet sich aus der Division des Nutzens durch die Kosten, der Netto-Nutzen bedeutet die Subtraktion der Kosten vom Nutzen.

Der Kosten-Nutzen-Index bezogen auf das erste Jahre nach Therapieende beträgt hier $1 : 0,58$. Das bedeutet, dass für ein in die Therapie investierte DM im ersten Jahr nach Therapie DM 0,58 als Nutzen entstehen. Der Netto-Nutzen ist im ersten Jahr also noch negativ. Ein „Reingewinn“ pro behandeltem Patienten entsteht im Laufe des zweiten Jahres (Tab. 4.11), sofern man davon ausgeht, dass die Krankheitskosten konstant bleiben.

Tabelle 4.11: Kosten-Nutzen-Indices und Netto-Nutzen (DM) für verschiedene Zeiträume¹⁹

Bezugszeitraum	Kosten-Nutzen-Index	Netto-Nutzen
1 Jahr	1 : 0,58	- 2268.-
2 Jahre	1 : 1,13	686.-
3 Jahre	1 : 1,65	3500.-
4 Jahre	1 : 2,15	6180.-
5 Jahre	1 : 2,63	8732.-

4.4.4 Sensitivitätsanalyse

Die unsicheren Preise der Kosten- und Nutzenfaktoren sollen in der Sensitivitätsanalyse hinsichtlich ihres Einflusses auf das Ergebnis (d.h. auf das Verhältnis von Kosten und Nutzen) überprüft werden. Das bezieht sich auf die Kosten der ambulanten Behandlungen und die Arbeitsunfähigkeit. Die Kosten der stationären Aufenthalte sind exakten Preisangaben entnommen; eine Sensitivitätsanalyse muss daher dafür nicht durchgeführt werden.

Es werden prozentuale Werte – Hebung bzw. Senkung der bisher genutzten Einzelkosten von 50% bzw. 150 % – festgelegt, mit Hilfe derer ein Einfluss festgestellt werden soll (Tab. 4.12).

Tabelle 4.52: Werte für die Sensitivitätsanalyse (in DM)

	tatsächlich genutzter Wert	auf 50% gesenkter Wert	auf 150% erhöhter Wert
Notarzt	400	200	600
Allgemeinarzt	30	15	45
Psychotherapeut	120	60	180
Facharzt	60	30	90
AU	215	107,5	322,5

In Tabelle 4.13 ist der mit diesen veränderten Kosten für Behandlungen verbundene veränderte Wert für die Einsparung dargestellt. Unter KNI ist der Kosten-Nutzen-Index aufgeführt, der sich aus der jeweiligen Änderung ergeben würde.

¹⁹ Ab dem 2. Jahr hochgerechnet

Tabelle 4.13: Sensitivitätsanalyse: Veränderung des Kosten-Nutzen-Indizes durch Preisvariation im ersten Jahr

Kostenfaktor	Preis (DM)	Gesamtnutzen (DM)	KNI
Psychotherapie	120,-	3103,-	1 : 0,58
	60,-	2871,-	1 : 0,54
	180,-	3333,-	1 : 0,62
ambulante	s. Tab. 4.11	3103,-	1 : 0,58
Arzt-	für einzelne	2917,-	1 : 0,54
Kontakte (zusammengefasst)	Gruppen	3289,-	1 : 0,61
Arbeitsunfähigkeit	335,-	3103,-	1 : 0,58
	235,-	2060,-	1 : 0,39
	436,-	4145,-	1 : 0,77

Die Schwankungen in der Sensitivitätsanalyse sind gering. Keine der Preis-Variationen kehrt das Kosten-Nutzen-Verhältnis um. Dies war auch zu erwarten, da die Einsparungen in allen Faktoren ähnlich hoch waren (vgl. Tab. 4.7). Die Analyse ist somit als sicher zu betrachten. Zu beachten ist allerdings, dass die Veränderung des Wertes für die Arbeitsunfähigkeit, zwar nicht das Kosten-Nutzen-Verhältnis umdreht, aber jedoch eine recht beträchtliche Veränderung des Kosten-Nutzen-Index bewirkt. Der Grund dafür ist der hohe Anteil der Einsparungen durch die Arbeitsunfähigkeit an den Gesamteinsparungen.

Wie verändern sich nun die Kosten-Nutzen-Indizes für längere Zeiträume? Zur Beantwortung dieser Frage wurden alle ambulanten Gesundheitsleistungen zusammen und die Arbeitsunfähigkeit getrennt variiert und der diskontierte Nutzen für 3 und 5 Jahre nach der Therapie berechnet. Die Indizes sind folgender Tabelle zu entnehmen

Tabelle 4.14: Sensitivitätsanalyse für medizinische Leistungen und AU für 1, 3 und 5 Jahre

		Gesundheitsleistungen			Arbeitsunfähigkeit		
Preisanteil		100%	50%	150%	100%	50%	150%
diskontierter							
Nutzen	1 Jahr	3103,-	2686,-	3519,-	3103,-	2060,-	4145,-
	3 Jahre	8871,-	7680,-	10063,-	8871,-	5889,-	11853,-
	5 Jahre	14103,-	12209,-	15998,-	14103,-	9363,-	18844,-
KNI							
	1 Jahr	1 : 0,58	1 : 0,50	1 : 0,66	1 : 0,58	1 : 0,38	1 : 0,77
	3 Jahre	1 : 1,65	1 : 1,43	1 : 1,87	1 : 1,65	1 : 1,10	1 : 2,21
	5 Jahre	1 : 2,63	1 : 2,27	1 : 2,98	1 : 2,63	1 : 1,74	1 : 3,51

Auch die Kosten für die Verhaltenstherapie konnten möglicherweise fehlerbehaftet sein, da bei den Personen, bei denen keine exakten Werte vorlagen, mit dem Mittelwert des jeweiligen Institutes gerechnet wurde. Selbst bei einer relativ großen Variation um 30% blieb hier die Grundaussage der Analyse bestehen. Getestet wurde der Einfluss auf den Index, wenn die durchschnittlichen Therapiekosten von 5371,- DM auf 3760,- DM gesenkt bzw. auf 6982,- DM gehoben wurden.

Tabelle 4.6: Veränderung der Kosten-Nutzen-Indices bei Variation der Therapiekosten

Kosten-Nutzen-Index für	Therapiekosten		
	100% (DM 5371,-)	70% (DM 3760,-)	130% (DM 6982,-)
1 Jahr	1 : 0,58	1 : 0,83	1 : 0,45
2 Jahre	1 : 1,13	1 : 1,62	1 : 0,87
3 Jahre	1 : 1,65	1 : 2,38	1 : 1,28
4 Jahre	1 : 2,15	1 : 3,10	1 : 1,68
5 Jahre	1 : 2,63	1 : 3,78	1 : 2,04

Die selben Werte würden sich ergeben, wenn bei gleicher Stundenzahl der Preis einer Behandlungseinheit DM 91,- (70%; entspricht in etwa dem Regelsatz für Psychotherapie der Ersatzkassen nach Inkrafttreten des Psychotherapeutengesetzes) bzw. DM 169,- (130%; entspricht in etwa dem derzeitigen Erstattungssatz bei Privatkassen) betragen würde. Würde der Preis einer Behandlungseinheit unter DM 75,- sinken, so würde sich das Kosten-Nutzen-Verhältnis im ersten Jahr umdrehen, d.h. dass bereits im ersten Jahr ein Netto-Nutzen größer 0 zu verzeichnen gewesen wäre.

Es wäre aber nicht zulässig, diese Kosten-Nutzen-Analyse mit derart niedrigen Preisen durchzurechnen, da angenommen werden muss, dass sich ein solcher Preisverfall auf die Qualität der Therapie (und damit auch auf den Nutzen durch die erzielten Einsparungen) auswirken kann.

4.5. Die Perspektive des Patienten: Einbezug der Zahlungsbereitschaft (WTP)

Die Perspektive der Patienten wird in diesem Abschnitt anhand einer Pilotstudie mit relativ geringem Stichprobenumfang ($N = 30$) analysiert, die in erster Linie der Entwicklung eines geeigneten Fragebogenmaßes diene. Es handelt sich also um vorläufige Ergebnisse, die durch weiteren Einsatz des Instrumentes untermauert werden sollten.

4.5.1 Fragen zur Validität der Methode

Wie unter 2.4 erläutert kann die Frage danach, wie viel Patienten an Geld und Zeit aufwenden würden, um eine Verbesserung ihres Gesundheitszustandes oder anderer psychotherapierelevante Ziele zu erreichen, Aufschluss über Nutzenaspekte geben, die bislang nicht berücksichtigt wurden. Die betrifft vor allem intangible Faktoren (Lebensqualität für Patienten und deren Umfeld) und den indirekten, nur sehr schwer abschätzbaren tangiblen Faktor der in der Zukunft liegenden verbesserten beruflichen Möglichkeiten.

Die eigentliche Fragestellung – *Wie verändert die Hinzunahme des WTP-Ansatzes das Kosten-Nutzen-Verhältnis?* – wird ergänzt durch einige weitere Fragen, die die Validität und die Anwendbarkeit dieses noch nicht etablierten Konzeptes testen sollen. Zur Überprüfung der Validität soll der Zusammenhang zwischen Höhe der Zahlungsbereitschaft in Geld und dem Einkommen erfasst werden. Ebenso soll der Zusammenhang zwischen dem auf einer Skala angegebenen wahrgenommenen Wichtigkeit der Therapie für das Leben einerseits und der Höhe der Zahlungsbereitschaft in Zeit und Geld andererseits überprüft werden. Dabei bestehen folgende Hypothesen:

Hypothese 1: Es besteht ein positiver Zusammenhang zwischen der Höhe der Zahlungsbereitschaft (Geld) und dem Einkommen.

Hypothese 2: Es besteht ein positiver Zusammenhang zwischen der Höhe der Zahlungsbereitschaft (Geld) und dem Anteil an Freizeit, den die Patienten für die Therapie zu opfern bereit sind.

Hypothese 3: Es besteht ein positiver Zusammenhang zwischen Zahlungsbereitschaft und Therapieerfolg bzw. wahrgenommener Wichtigkeit der Therapie

Laut den NOAA - Richtlinien soll außerdem das Verständnis und die Akzeptanz der Methode bei den Probanden mit offenen Fragen überprüft werden (Arrow et al., 1993; vgl. 2.4.2.3).

4.5.2 Methoden der Datenerhebung: Auswahl der hypothetischen Situation und Umsetzung der NOAA-Richtlinien

Die Zahlungsbereitschaft wurde per Fragebogen erfasst. Die Erhebung fand ex post, also nach der Intervention statt. Da die Befragten die Therapie bereits durchgeführt haben, sind sie „Experten“ bezüglich des Themas. So ist mit einer höheren Validität der Antworten zu rechnen (vgl. Krischke und Petermann, 1997; s. 2.4). Eine akkurate Beschreibung des Programms (NOAA-Richtlinie) ist somit nicht mehr erforderlich.

Da dies das erste Mal ist, dass die Methode der Zahlungsbereitschaft Anwendung im Bereich Psychotherapie verwendet wird, liegen noch keine Erfahrungswerte über die Höhe der Beträge vor, die Betroffene zu zahlen bereit sind. Daher werden in dem Fragebogen offene Fragen ohne vorgegebene Antwortalternativen eingesetzt. Analog zu dem Vorschlag von Fellmann (1996) wird in dieser Studie neben der Frage nach dem höchsten Geldbetrag, den jemand für die Therapie ausgeben würde auch die Frage nach dem Freizeitbetrag gestellt, den man für die Teilnahme an der Therapie opfern würde. Zeit hat aber für jeden einen unterschiedlich großen Wert. Es ist sogar vorstellbar, dass manche „zu viel“ Zeit haben (z.B. Arbeitslose) und darunter leiden, nichts mit ihrer Zeit anfangen zu können. Daher wird die Zufriedenheit mit der Menge der Freizeit ebenfalls erhoben.

Die Frage nach der Zahlungsbereitschaft erfordert immer den Umgang mit einer hypothetischen Situation. Es stehen hier zwei mögliche Situationen zur Auswahl:

1. Der Patient soll sich vorstellen, wieder unter Angst zu leiden und eine Therapie zu benötigen. Die Frage wäre, wie viel Geld er für die Teilnahme an einer Therapie, wie er sie kennen gelernt hat und mit eben dem Effekt, den seine Therapie hatte, zahlen würde.

Dagegen spricht, dass die Frage die tatsächliche Wertschätzung der realen Therapie messen soll. Der Wert der Therapie ergibt sich aus der Wirkung auf das Wohlergehen des Patienten, also aus der qualitativen Verbesserung sowie dem zeitlichen Andauern dieser Verbesserung. Stellt dieser sich aber vor, ihm ginge es wieder sehr schlecht, müsste automatisch die Therapie für ihn an Wert verlieren, weil der Effekt nur kurz überdauert hätte. Er müsste damit rechnen, dass der therapeutische Effekt jeweils nur kurze Zeit anhält. Daher fiel die Entscheidung auf das zweite Situation:

2. Der Patient wird aufgefordert, sich in die Zeit vor der Therapie hineinzuversetzen. Er soll sein jetziges Wissen, das er über Prozess und Wirkungsgrad der Therapie hat, nutzen, um die folgenden Fragen nach zu investierendem Geld- und Zeitbetrag zu beantworten.

Wie unter 2.4 erläutert, wird eine Orientierung an den NOAA-Richtlinien angestrebt. Diese werden in Kasten 4.1 aufgeführt und auch hinsichtlich ihrer Tauglichkeit im Bereich Psychotherapie untersucht.

Kasten 4.1: Umsetzung der NOAA-Richtlinien (Arrow et al., 1993; vgl. 2.4.2.3)

- *Konservative Schätzung:* Diese Forderung bedeutet hier, dass der Median zur Darstellung der Summe der Zahlungsbereitschaft genutzt werden soll.
- *Akkurate Beschreibung des Programms:* Das Expertentum zum Thema Psychotherapie besteht bei den ehemaligen Psychotherapiepatienten bereits.
- *Unterschiedliche Befragungszeitpunkte zur Überprüfung der Unabhängigkeit der Antworten von zeitlichen Einflüssen:* Dies ist auf Psychotherapie nicht übertragbar. Der Wert einer Psychotherapie steigt in der Tat, je länger das Ergebnis überdauert bzw. sinkt, wenn sich der psychische Zustand nach einer gewissen Zeit wieder verschlechtert. Dies ist bei einem Umweltunfall anders, bei dem etwa die Beeinträchtigung der Lebensqualität durch zerstörte Natur nach einem halben Jahr noch genauso schmerzlich für die Betroffenen sein sollte wie nach sechs Wochen.
- *Offene Fragen zur Ermittlung der Gründe des Antwortverhaltens:* Nach Angabe des Betrages wird in unserem Fragebogen zusätzlich die Begründung für die Höhe des Betrages erfasst.
- *Einbezug von Interpretationshilfen der Antworten:* Als Interpretationshilfe der Antworten werden das monatliche Einkommen der Probanden, die Anzahl der von diesem Einkommen lebenden Haushaltsmitglieder, die Zufriedenheit mit dem Einkommen, die Zufriedenheit mit dem Ausmaß der Freizeit und Beurteilungen der Besserung des Befindens erfasst. Es wurde davon abgesehen, auch die zur Verfügung stehende Freizeit zu erfassen, da es kaum möglich scheint, die Zeit, die frei von Verpflichtungen wie Beruf, Einkauf, Beschäftigung mit Kindern o.ä. ist, zu errechnen. Daher konnte nur in einer zusätzlichen Frage erfasst werden, wie viel der Freizeit der angegebene Betrag in Kategorien wie „alles“, „viel“, „eher wenig“ darstellt. Auf eine Erfassung in Prozent der Freizeit wurde aufgrund der noch höheren Komplexität dieser Frage verzichtet.
- *Akzeptanz:* Im Fragebogen wird zusätzlich die Meinung der Versuchspersonen zu dem Fragebogen erhoben, um Einblick in die Akzeptanz und das Verständnis der Erhebungsmethode zu gewinnen.
- *Einbezug des Budgets:* Die Probanden werden daran erinnert, dass sie den angegebenen Betrag dann nicht mehr für andere Dinge ausgeben könnten.
- *Geschlossene Antworten:* In dieser Studie werden entgegen der Empfehlung offene Antworten genutzt, um erstmals einen Einblick in die Höhe der Zahlungsbereitschaft zu gewinnen. Zwar muss hier mit einer geringeren Rücklaufquote brauchbarer Antworten gerechnet werden (Johannesson et al., 1991), aber mit Baron (1995) wird argumentiert, dass die höhere Validität der geschlossenen Antworten nicht bewiesen wurde (vgl. Abschnitt 2.4.2.3).

Der vollständige Fragebogen findet sich im Anhang.

Die mittlere Zahlungsbereitschaft wird über den Mittelwert wie auch über den Median bestimmt, wobei der Median als konservativere Schätzung gilt (vgl. Ryan, 1996). Ebenso wird die mittlere Zahlungsbereitschaft als Prozentsatz des Haushaltseinkommens bestimmt (ebenfalls Median und Mittel), da nur so der eigentliche Wert für die Patienten wiedergegeben werden kann. Eine weitere Möglichkeit, den Wert, den das Geld für die Befragten hat, zu bestimmen, ist es, das Einkommen durch die Anzahl der Personen zu dividieren, die von diesem Einkommen leben. Für die angegebene Freizeit ist Median und Mittel bzw. für den Betrag der Freizeit („alles“, „viel“, „wenig“,...) der Median zu bestimmen. Analog zur vorhandenen Literatur wird die Korrelation zwischen Höhe der Zahlungsbereitschaft und Einkommen (bzw. dem durch die Anzahl der Haushaltsmitglieder dividierten Einkommen) ermittelt.

4.5.3 Stichprobe

Befragt wurden Patienten der CDS Dresden mit der Diagnose einer Angststörung nach DSM III-R bzw. DSM IV. Der Fragebogen wurde im Rahmen von regulären 6-Wochen-Nachuntersuchungen und 1-Jahres-Katamnesen im Zeitraum von Februar bis April 1999 mitgeschickt. Das Einverständnis der Patienten wurde schriftlich auf dem Fragebogen eingeholt.

In Tabelle 4.16 wird die Stichprobe hinsichtlich der Merkmale Geschlecht, Einkommen, Bildungsstand, und Anzahl der Haushaltsmitglieder. Es fällt auf, dass der Bildungsstand der Stichprobe Zahlungsbereitschaft höher ist, als der der Gesamtstichprobe in der Christoph-Dornier-Stiftung. Möglicherweise ist das darauf zurückzuführen, dass die Teilnahme an der Untersuchung „freiwilliger“ war als die reguläre Katamnese und hier eine Selbstselektion stattgefunden hat.

Der Therapieerfolg der Stichprobe wurde hier nicht so differenziert wie bei der Kosten-Effektivitätsstudie ausgewertet. Es wurde lediglich das siebenstufige globale Maß herangezogen („Im Vergleich zum Zeitpunkt vor Therapiebeginn geht es mir...“ 1: ...sehr viel besser bis 7: ... sehr viel schlechter). Der Therapieerfolg war im Median mit „viel besser“ angegeben; diesbezüglich gibt es keinen Unterschied zwischen der untersuchten Stichprobe und der Gesamtstichprobe der Patienten der Christoph-Dornier-Stiftung (Mann-Whitney U-Test: $p = 0.695$).

Tabelle 4.16: Stichprobendaten WTP-Studie (N = 30)

Alter:	$\bar{x} = 36,6$	$s = 9,8$
Geschlecht:	männlich:	31 %
	weiblich:	69 %
höchster Schulabschluß:	Hauptschule:	3,8 %
	Realschule:	15,4 %
	Fachabitur/Abitur:	34,6 %
	abgeschlossenes Fach-/Hochschulstudium:	46,2 %
Einkommen:	Median: 2000-3000 DM	(entspricht dem Median des durchschnittlichen Dresdner Patienten)
Anzahl der Haushaltsmitglieder:	2,2 Personen	

4.5.4 Ergebnisse

4.5.4.1 WTP-Beträge und deren Verteilung

Von den 30 ausgefüllten Fragebögen waren vier nicht zur Ermittlung der Zahlungsbereitschaft brauchbar, da anstelle konkreter Zahlen Antworten wie „sehr viel“ oder „alles, was ich habe“ angegeben waren. Im folgenden sind die Ergebnisse aus den 26 anderen Fragebögen dargestellt.

Die Befragten waren durchschnittlich bereit, DM 5004,17 für die Therapie auszugeben. Im Median ausgewertet entsprach dies einer Zahlungsbereitschaft von DM 3500,- (Tab. 4.17). Der angegebene Betrag betrug 130% (Median) des monatlichen Gesamt-Haushalts-Nettoeinkommens bzw. 161,4 % im Mittelwert.

Tabelle 4.18 gibt die Verteilung der WTP (Geld) an, Tabelle 4.19 die Verteilung des Anteils am Haushalts-Nettoeinkommen. Ein Viertel der Befragten würde insgesamt bis 50% des monatlichen Haushaltseinkommens für die Therapie bezahlen, die Hälfte der Befragten über ein halbes bis zu zwei Monatseinkommen und ein weiteres Viertel über 2 Monatseinkommen.

Der angegebene Freizeitbetrag wurde auf einen Zwölfstundentag umgerechnet, das heißt „ein Tag“ wurde mit zwölf Stunden eingegeben. Die Befragten gaben durchschnittlich an, zwei Tage (24 h) pro Woche für die Dauer eines halben Jahres für die Therapie opfern zu wollen (Median). Die entsprechende Verteilung findet sich in Tabelle 4.20.

Tabelle 4.7: Angegebene Beträge der Zahlungsbereitschaft in Geld und Zeit bei N=26 Patienten nach Therapie

	WTP (in DM)	WTP (als Prozentsatz des Haushalts- Nettoeinkommens)	Zeit in Stunden, die wö- chentlich für Therapie aufgewendet werden würde	Betrag der Zeit
Mittelwert	5004,17	161,4	30,67	-
Median	3500	130,56	24	4 = „viel“

Tabelle 4.18: Verteilung der Zahlungsbereitschaft in DM bei N=26 Patienten nach Therapie

WTP in DM	Häufigkeit	Prozent	kumulierte Prozente
bis 2000	10	38,5	38,5
bis 4000	5	19,2	57,7
bis 6000	5	19,2	76,9
bis 8000	1	3,8	80,7
bis 10000	3	11,5	92,3
über 10000	2	7,7	100,0

Tabelle 4.19: Verteilung der Zahlungsbereitschaft in % des Haushaltseinkommen bei N=26 Patienten nach Therapie

WTP in % des Haushalts- Nettoeinkommens	Häufigkeit	Prozente	kumulierte Prozente
bis 50%	6	23,1	23,1
bis 100%	5	19,2	42,3
bis 150 %	4	15,4	57,7
bis 200%	5	19,2	76,9
bis 300%	2	7,7	84,6
über 300%	4	15,4	100

Tabelle 4.20: Verteilung des absoluten zu investierenden Betrages der Freizeit pro Woche (für die Dauer eines halben Jahres) bei N=26 Patienten nach Therapie

WTP in Stunden	Häufigkeit	Prozente	Kumulierte Prozente
bis 12 h	5	19,2	19,2
bis 24 h	10	38,4	57,7
bis 36 h	5	19,2	75,0
bis 48 h	1	3,8	80,8
bis 60 h	5	19,2	100,0

Der Anteil an der Freizeit hat eine deutlich geringere Varianz als der Absolutbetrag. Er wurde von den Probanden recht konsistent als „viel“ bzw. „alles“

von der zur Verfügung stehenden Freizeit eingeschätzt. Der angegebene Betrag der Freizeit wird bei nur einer Person mit „eher wenig“ bezeichnet, etwa 80% der Befragten würden viel ihrer Freizeit bzw. ihre gesamte Freizeit opfern (Tab. 4.21).

Tabelle 4.8: Verteilung des zu investierenden Betrages der Freizeit pro Woche (für die Dauer eines halben Jahres) nach Kategorien

WTP in Anteil an Freizeit	Häufigkeit	Prozent
eher wenig	1	3,8
die Hälfte	4	15,4
viel	15	57,7
alles	6	23,1

4.5.4.2 Begründung der Antworten und Bewertung des Fragebogens

Die Frage nach der Begründung der Antworten wird gestellt, um Missverständnisse aufzudecken. Antwortet zum Beispiel ein Proband darauf, er würde den genannten Betrag zahlen wollen, weil er glaube, dass der Therapeut Geld brauche oder dass es eine gute Sache sei, für eine Behandlungseinrichtung zu spenden, hat er die Frage nicht verstanden. Die Antwort kann dann nicht gewertet werden.

Die Begründungen legten alle ein Verständnis der Frage nahe. Die Höhe des Betrages wird mit der durch die Therapie gewonnene bessere Lebensqualität begründet („Die Therapie hat mir einen neuen Lebensbeginn ermöglicht.“). Manche Probanden verweisen darauf, dass ein höherer Betrag nicht möglich ist, da nicht mehr Geld zur Verfügung steht. („Ich würde alles, was möglich ist zahlen, aber mein Einkommen ist zu klein“). Eine interessante Begründung war: „Die Therapie ist die Voraussetzung für meine Erwerbstätigkeit“. Die betreffende Patientin scheint eine „private Kosten-Nutzen-Analyse“ durchgeführt zu haben, um den Betrag, den sie zahlen würde zu schätzen.

Eine Patientin empfand die Frage nach der Höhe der Summe, die sie zu zahlen bereit wäre, generell als problematisch und störend, ohne nähere Begründungen anzugeben. Keiner der Probanden gab an, Verständnisprobleme mit dem Fragebogen gehabt zu haben oder es schwierig, gefunden zu haben, den Fragebogen auszufüllen. Hierbei ist zu beachten, dass die Probanden eine hohe Schulbildung hatten.

Zur Frage nach dem Nutzen des Fragebogens gab es verschiedene Ansichten. Einige Probanden waren der Meinung, dass diese Art der Therapiebewertung zur Urteilsbildung beitragen könne und wichtig sei. Andere hatten daran Zweifel; zwei Personen waren der Meinung, dass der Erfolg viel besser an

bildhaften und konkreten Veränderungen dargestellt werden könne und nicht in Zahlen. Eine Probandin empfand die Tatsache, dass bei unterschiedlichen Leuten stark unterschiedliche finanzielle Prioritäten bestehen, als problematisch. Andererseits sei der Geldbetrag ein Mittel, um auszudrücken, wie viel einem an der Therapie liegt, wenn sie einem tatsächlich geholfen hat.

4.5.4.3 Validierung über Zusammenhangsmaße

Zum Zwecke der Validierung werden in den bisherigen Studien zur Zahlungsbereitschaft die Zusammenhänge zwischen Höhe der Summe der Zahlungsbereitschaft und Höhe des Einkommens festgestellt (Tab. 4.22). Hier soll auch der Zusammenhang zwischen der auf einer Skala angegebenen Wichtigkeit, die die Therapie für das Leben des Einzelnen hat, und der Höhe des Zeit- und Geldbetrag zum Zwecke der Validierung überprüft werden. Ebenfalls soll auch die interne Konsistenz der Antworten bezogen auf die Frage nach Zeit und Geld überprüft werden. Die Summe der Zahlungsbereitschaft wird sowohl als Absolutsumme als auch als Prozentsatz des Haushaltseinkommens bzw. als Prozentsatz des Haushaltseinkommens, dividiert durch die Anzahl der davon lebenden Haushaltsmitglieder, mit den Maßen für Therapie-Erfolg in Beziehung gesetzt.

Tabelle 4.22: Zusammenhang zwischen Summe der Zahlungsbereitschaft, Einkommen, Therapieerfolgsmaßen und zeitlicher Zahlungsbereitschaft (Ergebnisse der Korrelationsanalyse)

	Zeit	Anteil an Freizeit	Therapie-Erfolg	Einkommen	Zufriedenheit mit Einkommen	Wichtigkeit der Therapie fürs Leben
Summe	.04	.60**	.11	.54**	.50**	.00
	Pearson	Spearman	Spearman	Pearson	Spearman	Spearman
Summe als Betrag des Haushaltseinkommens	.09	.21	.12	-.21	.11	.09
	Pearson	Spearman	Spearman	Pearson	Spearman	Spearman
Summe als Betrag des Haushaltseinkommens durch Anzahl der Haushaltsmitglieder	.13	.04	.15	.13	.19	.28
	Pearson	Spearman	Spearman	Pearson	Spearman	Spearman

Ein hoch signifikanter Zusammenhang kann zwischen der Höhe der Summe der Zahlungsbereitschaft und der Höhe des Einkommens sowie zwischen der Höhe der Summe und der Zufriedenheit mit dem Einkommen festgestellt werden. Dies entspricht der Hypothese 1 aus 4.5.1.

Ebenfalls signifikant ist die Korrelation zwischen den Zahlungsbereitschaften für Geld und Freizeitanteil, was die Hypothese 2 aus 4.5.1 bestätigt.

Die Korrelationsanalyse zeigt keinen Zusammenhang zwischen Zahlungsbereitschaft (Geld) und dem Maß für Therapieerfolg oder der Frage nach der Wichtigkeit der Therapie fürs Leben auf. Auch wenn die Summe bezüglich Haushaltseinkommen und die davon lebenden Haushaltsmitglieder korrigiert wird, findet sich kein Zusammenhang. Hypothese 3 aus 4.5.1 konnte hiermit also nicht bestätigt werden.

Allerdings fand sich ein Zusammenhang zwischen Therapie-Erfolg und dem Betrag an Freizeit. In der Frage nach der Wichtigkeit der Therapie für das Leben wird die Tendenz zu diesem Zusammenhang jedoch nicht signifikant (Tab. 4.23).

Die niedrige Korrelation zwischen angegebenen Beträgen und dem globalen Therapieerfolgsmaß lässt sich möglicherweise auf die niedrige Varianz dieser Werte innerhalb der Stichprobe zurückführen (77% antworteten mit „sehr viel besser“ oder „viel besser“). Die Varianz der Variable „Wichtigkeit der Therapie für das Leben“ ist noch geringer: alle Probanden antworteten entweder mit „sehr wichtig“ oder mit „eher wichtig“, also nur mit den zwei höchsten Werten.

Tabelle 4.9: Zusammenhang zwischen zu opfernder Freizeit, Zufriedenheit mit Freizeit und Maßen für Therapie-Erfolg und Wichtigkeit der Therapie (Ergebnisse der Korrelationsanalyse; Korrelationskoeffizienten nach Spearman)

	Therapie-Erfolg	Wichtigkeit der Therapie	Zufriedenheit mit Freizeit
Zeit	.10	.43	.138
Anteil an Freizeit	.36*	.40	-.422

4.5.4.4 Verrechnung der WTP in der Kosten-Nutzen-Analyse

Ziel der Ermittlung der subjektiven Zahlungsbereitschaft ist, auch intangible und indirekte Nutzen-Aspekte monetär zu erfassen und somit zusätzlich zur volkswirtschaftlichen die Patientenperspektive einfließen zu lassen. Da für Kosten-Nutzen-Analysen eine möglichst umfassende Bewertung gefordert wird, und da sich der bei der Zahlungsbereitschaft erfragte Betrag nicht mit den Nutzenfaktoren der volkswirtschaftlichen Kosten-Nutzen-Analyse dieser Arbeit überschneidet²⁰, erscheint es gerechtfertigt, den WTP-Wert vollständig dem in 4.4.1 ermittelten tangiblen Nutzen zuzuschlagen. Ein solches Vorgehen wurde bereits von Fellmann (1996) vorgeschlagen.

In Tabelle 4.24 finden sich die in 4.4.3 ermittelten Kosten-Nutzen-Werte unter Einbezug der Zahlungsbereitschaft (gewählt wurde der – konservative – Median von DM 3500.-), die als einmalig auftretender Nutzen veranschlagt wurde. Es wird ersichtlich, dass die Kosten-Nutzen-Analyse vor unter Hinzunahme dieser Perspektive deutlich günstiger ausfällt.

Tabelle 4.24: Kosten-Nutzen-Indices und Netto-Nutzen (DM) für verschiedene Zeiträume bei Hinzunahme der WTP auf der Nutzen-Seite

Bezugszeitraum	Kosten-Nutzen-Index (ohne WTP)	Kosten-Nutzen-Index (mit WTP)	Netto-Nutzen (ohne WTP)	Netto-Nutzen (mit WTP)
1 Jahr	1 : 0,58	1 : 1,23	- 2268.-	1232.-
2 Jahre	1 : 1,13	1 : 1,78	686.-	4186.-
3 Jahre	1 : 1,65	1 : 2,30	3500.-	7000.-
4 Jahre	1 : 2,15	1 : 2,80	6180.-	9680.-
5 Jahre	1 : 2,63	1 : 3,28	8732.-	12232.-

²⁰ Überschneidungen gäbe es lediglich für Selbständige und Unternehmer, die aufgrund von durch die Störung verminderter Arbeitsleistung direkt selbst monetär von einer durch die Therapie gebesserten Arbeitsfähigkeit profitierten und dies in ihre WTP-Angaben einfließen ließen. Dies war in dieser kleinen Stichprobe jedoch nicht der Fall.

4.6 Vereinfachung in Durchführung und Auswertung von Kosten-Nutzen-Analysen im Psychotherapiebereich

Als Nebenaspekt dieser Arbeit soll nun noch zwei forschungspraktischen Aspekten nachgegangen werden, die Kosten-Nutzen-Analysen im Psychotherapiebereich vereinfachen sollen. Dies soll zu einer häufigeren – weil dringend benötigten – Durchführung empirischer Kosten-Nutzen-Analysen in Behandlungseinrichtungen oder in Psychotherapiestudien in der Zukunft beitragen.

4.6.1 Querschnitt statt Längsschnitt?

Die Durchführung einer Kosten-Nutzen-Analyse im Längsschnitt ist relativ aufwendig, da von allen Probanden sowohl vor wie nach Therapie Kostendaten vorliegen müssen. Um zu prüfen, ob querschnittliche Daten vergleichbar sind, also ob es ausreicht, lediglich die Störungskosten des durchschnittlichen Patienten vor Therapie mit denen des durchschnittlichen Patienten nach Therapie gegenüberzustellen, wurde eine entsprechende Querschnittsuntersuchung angestellt.

Die Prä-Stichprobe bestand aus N=213 Probanden, die sich zwischen Anfang 1997 und Ende 1998 bei einem der vier Institute der Christoph-Dornier-Stiftung vorstellten und mindestens eine Hauptdiagnose einer Angststörung nach DSM-III-R bzw. DSM-IV erfüllten. Die Post-Stichprobe setzte sich zusammen aus N = 205 Patienten der vier Institute, für die Kostendaten aus Nachuntersuchungen vorlagen. Die Patienten, die in der längsschnittlichen Kosten-Nutzen-Analyse untersucht worden waren, waren in diese Gruppen nicht mit aufgenommen worden.

Die beiden Querschnitts-Stichproben wurden auf eventuelle Unterschiede untersucht und hinsichtlich ihrer Ausprägung in folgenden Variablen zum Prä-Zeitpunkt verglichen: soziodemografische Variablen Geschlecht, Familienstand, Alter und Schulbildung und die krankheitsbezogenen Variablen Krankheitsdauer, Primärdiagnose und Schweregrad der Beeinträchtigung (gemessen anhand des Gesamtwertes in der SCL-90-R; vgl. Kosten-Effektivitäts-Studie). Es ergaben sich diesbezüglich keine signifikanten Unterschiede; es darf also angenommen werden, dass Prä- und Post-Stichprobe aus der selben Population entstammten. Ebenfalls keine Unterschiede ergaben sich zur Längsschnitt-Stichprobe (4.3.1 und 4.3.2) und zur Gesamtstichprobe (vgl. Kosten-Effektivitäts-Studie; 3.3.3), weshalb die Stichprobendaten hier nicht nochmals gesondert aufgeführt werden. Die Methoden zur Kosten- und Nutzenbestimmung entsprachen denen der Längsschnittuntersuchung.

In Tabelle 4.26 sind die Kosten-Nutzen-Indices und die Nettonutzen beider Untersuchungen einander gegenübergestellt.

Tabelle 4.25: Kosten-Nutzen-Indices und Netto-Nutzen (DM) für Längsschnitt (N=67) und Querschnittsuntersuchung (N=213 vor und 205 nach Therapie); Zeiträume größer ein Jahr: diskontierte Hochrechnung

Bezugszeitraum	Kosten-Nutzen-Index	Kosten-Nutzen-Index	Netto-Nutzen	Netto-Nutzen
	Längsschnitt, N = 67	Querschnitt, N1 = 213, N2 = 205	Längsschnitt, N = 67	Querschnitt, N1 = 213, N2 = 205
1 Jahr	1 : 0,58	1 : 0,99	- 2268.-	- 15,-
2 Jahre	1 : 1,13	1 : 1,95	686.-	5085,-
3 Jahre	1 : 1,65	1 : 2,85	3500.-	9943,-
4 Jahre	1 : 2,15	1 : 3,71	6180.-	14570,-
5 Jahre	1 : 2,63	1 : 4,53	8732.-	24346,-

Dass die Querschnitts-Ergebnisse besser ausfallen, liegt daran, dass in der Querschnitts-Prä-Stichprobe ein bedeutend höheres Ausmaß an Arbeitsunfähigkeit zu verzeichnen war als in der Längsschnittstichprobe zum Prä-Zeitpunkt. Zum Postzeitpunkt unterschieden sich Längs- und Querschnittstichprobe in keinem Kostenfaktor signifikant. Die Einsparungen waren deshalb im Bereich Arbeitsunfähigkeit im Querschnitt bedeutend höher. Die Aufschlüsselung von Unterschieden und Gemeinsamkeiten getrennt nach Kostenfaktoren ist den Diplomarbeiten von Pieper (1999) und Runde (1999) zu entnehmen.

Dennoch differieren die Werte nicht dramatisch. Die grundsätzliche Aussage – Psychotherapie spart Kosten in allen untersuchten Kostenfaktoren und amortisiert sich bereits nach ein bis zwei Jahren – sind in beiden Studien gleich. Man kann also schlußfolgern, dass eine Querschnittsuntersuchung zur Kosten-Nutzen-Analyse ausreicht.

4.6.2 Ein Programm zur Vereinfachung von Kosten-Nutzen-Berechnungen: Der „Cost-Benefit-Calculator“

Die Berechnung von Kosten-Nutzen-Indices und diskontierten Nutzenwerten ist eine umständliche Angelegenheit, deren Aufwand mit der Zahl der berücksichtigten Kostenfaktoren und berechneten unterschiedlichen Zeiträume stetig steigt. Insbesondere im Rahmen der Sensitivitätsanalyse müssen ganze Serien

solcher Berechnungen durchgeführt werden. Deshalb wurde im Rahmen dieser Studie ein Tool entwickelt, mit dem man Kosten-Nutzen-Index und diskontierten Nutzen nach Eingabe der Variablen in eine Maske einfach per Mausklick erhält.

Somit kann man in kurzer Zeit für alle möglichen Variationen von Preisen und Einheiten der berücksichtigten Kostenfaktoren die Ergebnisse bestimmen. Die Oberfläche des Programms ist einer HTML- bzw. Java-Seite nachempfunden.

Der in Abbildung 4.1 dargestellte Cost-Benefit-Calculator befindet sich auf beiliegendem Datenträger²¹.

²¹Der Autor dankt Markus Pieper für die Programmierung.

Abb. 4.1: Formularmaske des Cost-Benefit-Calculators mit Beispielberechnung

5 Zusammenfassungen und Diskussion

In diesem Teil werden die Ergebnisse der Arbeit hinsichtlich eher technisch-methodischer Aspekte (z.B. zu Design und Störhypothesen) diskutiert. Die Diskussion übergreifender Aspekte und ein Ausblick finden sich im Abschnitt 6.

5.1 Kosten-Effektivitäts-Analyse: Zusammenfassung

Eine Stichprobe von 493 Patienten mit Angststörungen, die eine ambulante störungsspezifische kognitive Verhaltenstherapie absolviert haben, wurden vor Therapie, 6 Wochen nach Therapie und zur 1-Jahres-Katamnese untersucht. Das Spektrum der Erhebungen umfasste störungsspezifische und störungsübergreifende Symptombelastung und Beeinträchtigungen in verschiedenen Lebensbereichen, sowie eine globale Therapieerfolgseinschätzung von Patienten und Therapeuten. Das Behandlungssetting der Studie ist charakterisiert durch klinisch relevante Rahmenbedingungen und Praxisnähe.

Im ersten Themenkomplex dieser Arbeit wurden die Ergebnisse dieser Effektivitätsstudie im Rahmen einer Kosten-Effektivitäts-Analyse mit den Behandlungskosten in Beziehung gesetzt. Damit wurde eine lang geforderte Transparenz bezüglich dem Verhältnis von Therapieerfolg und den dafür aufgewendeten Ressourcen geschaffen. Die ermittelten Kosten-Effektivitäts-Indices stehen nun bereit zum Vergleich mit den – in weiteren Studien oder durch Re-Analysen bereits durchgeführter Effektivitätsstudien oder anhand von Qualitätssicherungs-Daten zu ermittelnden – Ergebnissen anderer Behandlungsformen oder Behandlungssettings.

Zu den Ergebnissen:

- *Kostenseite:* Anhand der durchschnittlichen Zahl der Behandlungseinheiten (VT à 50 Minuten) und nach Kenntnis der Tarife wurden 6587,- DM ($s=3001,-$ DM) als durchschnittliche Therapiekosten errechnet.
- *Effektstärken:* Die Effektstärken reichten in Abhängigkeit von der Meßmethode von 0.9 bis 1.9. Die Verbesserungen waren durchgängig signifikant und blieben bis zur 1-Jahres-Katamnese stabil.
- *Besserungsraten und Kosten-Effektivitäts-Indices:* Die Raten klinisch bedeutsamer Besserungen betragen 63% - 79% in Abhängigkeit von Meß- und Auswertungsmethode. Die entsprechenden Kosten-Effektivitäts-

Indices betragen DM 8338.- bis DM 10456.- pro erfolgreiche Behandlung. Die KE-Indices geben an, was *eine erfolgreiche* Behandlung durchschnittlich kostet (also unter Einbezug der nicht erfolgreichen Behandlungen in die Gesamtkosten).

Es wurden auswertungs-, störungs- und patientenspezifische Unterschiede bei der Kosten-Effektivität gefunden. Dies verweist auf die Notwendigkeit, bei Veröffentlichung und Vergleich von Kosten-Effektivitäts-Indices auf Objektivität bzw. Nachvollziehbarkeit zu achten und bei der Interpretation zu berücksichtigen, dass der Forschungsstand zu Kosten-Effektivitäts-Analysen im psychotherapeutischen und psychiatrischen Bereich (im Gegensatz zur klassischen Outcome-Forschung) noch relativ dünn ist.

5.2 Sind klinisch-psychologische Behandlungen modernen Zugschnitts effektiv?

In diesem Abschnitt wird zum Erfolg der untersuchten Therapieform – noch ohne Beachtung von Kosten – Stellung bezogen.

5.2.1 Allgemeine Bewertung der Effektgrößen und der Besserungsraten

Die vorgefundenen Effekte sind als groß zu bewerten. Im Allgemeinen werden Effektstärken ab 0.5 als „mittel“ und ab 0.8 als „groß“ bezeichnet. Rosenthal (1982, 1990) verweist zudem auf die häufige Unterschätzung von „moderaten“ Effektstärken. Er zeigt, dass eine Effektstärke von 0,85, wie sie von Smith, Glass und Miller (1980) ermittelt wurde, ein äußerst bedeutungsvolles Ergebnis ist. Effektstärken unbehandelter Kontrollgruppen bewegen sich in der Regel um 0²² (Grawe et al., 1994, Ruhmland & Margraf, 2000).

Außerdem ist es als günstig zu bewerten, dass sich die Varianzen von Prä nach Post durchgehend verringert haben (eine Vergrößerung der Varianzen hätte auf starke differentielle Effekte der Behandlung hingedeutet).

Besonders hervorgehoben werden sollte auch die *Breite der Veränderungen* (nicht zuletzt, um der „Symptomverschiebungs-Hypothese“ entgegenzutreten). Gemäß den Empfehlungen aus der bisherigen Outcome-Forschung wurde multimodal und multimethodal vorgegangen. Sowohl die Beeinträchtigung in verschiedenen Lebensbereichen als auch die Symptombelastung wurde berücksichtigt. Hierbei kamen international anerkannte, reliable und validierte störungsspezifische und störungsübergreifende Fragebogenmaße zum Einsatz, so dass die Ergebnisse leicht mit denen anderer Studien verglichen werden können. Die Auswahl der Instrumente erfolgte auch aufgrund Überlegungen zur inhaltlichen Reinheit (möglichst geringe Überschneidung der erfassten Bereiche).

Zusätzlich wurde mit einem globalen Item der Therapieerfolg aus Patienten- und Therapeutenperspektive eingeschätzt. Bei dieser globalen Veränderungsmessung wurden die höchsten Effektstärken gefunden. Dies mag zum einen auch mit einem möglichen *bias* in Richtung positiv aufgrund der Skala zusammenhängen. Möglich wäre auch, dass dort am meisten Information der globalen Befindlichkeit quasi additiv einfließt und die Effektstärke damit gegenüber den Veränderungen in Teilbereichen in die Höhe treibt. Auch auf-

²² Die mittleren Effekte von Placebo-Kontrollgruppen schwanken allerdings um 0.4; zur Problematik von „Placebo“ als Behandlungsbedingung in der Psychotherapieforschung s. Grawe (1998, S.26ff)

grund methodischer Besonderheiten der Effektstärkeberechnung bei direkter Veränderungsmessung sollte das globale Maß gesondert betrachtet werden.

Neben der gruppenstatistischen Signifikanz der Veränderungen und der Größe der Effektstärken sollte auch die klinische Signifikanz und damit die individuellen Verbesserungen auf Patientenebene untersucht werden. Für die verschiedenen Erhebungsmodalitäten wurden dabei Kriterien klinisch signifikanter Veränderungen definiert, so dass *Besserungsraten* ermittelt werden konnten²³. Dabei wurde berücksichtigt, dass sich manche Patienten in einzelnen Bereichen auch schon vor der Therapie im unauffälligen Bereich bewegten und sich dort also nicht weiter verbessern konnten. Zur Analyse klinisch signifikanter Veränderungen eignet sich die Darstellung in (Kosten-) Effektivitätsmatrizen, aus denen sich die jeweiligen Veränderungen nicht nur global (Randsummen), sondern in den einzelnen Zellen hinsichtlich der „Wanderungen“ zwischen den Kategorien von prä nach post differenziert ableiten lassen. Die *Besserungsraten* zwischen 63% und 79% (in Abhängigkeit der Messmodalität) erachten wir ebenfalls als ermutigend (auch wenn leider in der Regel keine Vergleichswerte klinischer Signifikanz berichtet werden und so die Vergleichbarkeit zu anderen Studien nur bedingt gegeben ist).

Die *Besserungsraten* aufgrund der Fragebogendaten sind relativ am niedrigsten. Eine mögliche Erklärung wäre, dass die in den Fragebögen erfassten (Symptom-) Bereiche nur einen Teil der wahrgenommenen Verbesserung erklären und sich manche Patienten subjektiv global „sehr stark“ (d.h. klinisch signifikant) gebessert haben, obwohl sie in dem einen oder anderen Fragebogenbereich immer noch im auffälligen Bereich liegen (d.h. dort nicht klinisch signifikant gebessert sind).

Die Berücksichtigung klinischer Signifikanz erlaubt auch Aussagen zu individuellen klinisch signifikanten *Verschlechterungen* (die bei der rein statistischen Betrachtungsweise bzw. bei Effektstärken im Durchschnitt untergehen). Die damit ermittelten Werte (globales Maß: 2,6%, Beeinträchtigung: 6%, störungsspezifische Symptomatik: 4%, störungsübergreifende Symptomatik: 4%) sind niedrig und weisen auf die weitgehende Unbedenklichkeit der untersuchten Behandlungsform hin (vgl. Jacobi, 2001).

²³ Eine weitere Methode zur Erweiterung statistischer Signifikanz ist der *Reliable Change Index* (Christensen & Mendoza, 1986; Jacobson et al, 1984), bei dem der Standardfehler der Messungen und damit die Reliabilität der Messinstrumente einbezogen wird. Entsprechende Analysen unserer Stichprobe finden sich bei Sekot (1999), wurden aber aus Gründen der Übersichtlichkeit bzw. nicht ausreichendem zusätzlichem Erkenntnisgewinn ebenso wie die Analysen zahlreicher weiterer Fragebogenmaße nicht in die vorliegende Arbeit mit eingeschlossen.

5.2.2 Berücksichtigung einer „subklinischen“ Kategorie

Entscheidend bei der Ermittlung von klinisch signifikanten Besserungsraten ist die Festlegung von Kriterien, was als „gesund“ oder „krank“ erachtet wird (z.B. cut-off-Werte in Fragebögen). Dabei haben wir zusätzlich „subklinische“ Kategorien eingeführt, also Bereiche, in denen Belastung und Beeinträchtigung bereits mäßig auffällig sind, aber noch nicht klar „gestörte“ Ausmaße angenommen haben. Diese Idee entsprang der Beobachtung, dass Patienten ihre Befindlichkeit häufig als stark gebessert erleben, obwohl sie sich in Einzelbereichen noch von der gesunden Population unterscheiden. Außerdem können so auch Verschlechterungen differenzierter analysiert werden.

Der Einbezug einer subklinischen Kategorie erhöht die Sensitivität für Veränderungen. Dies führt zu deutlich höheren klinisch bedeutsamen Besserungsraten (aber nur zu leicht erhöhten Verschlechterungsraten) im Vergleich zur dichotomen Betrachtung. Hier wird argumentiert, dass die dichotome Methode zu „streng“ ist, also einige Patienten mit individuell als bedeutsam erlebte Verbesserungen als Non-Responder klassifiziert werden.

Der Vergleich der beiden Methoden zeigt auch, dass die Ergebnisse unter Hinzunahme der subklinischen Kategorie stabiler zwischen den Meßmethoden sind als bei dichotomer Betrachtungsweise. So beträgt die Besserungsrate im Beck-Angst-Inventar bei dichotomer Betrachtungsweise nur 38% (60% klinisch relevant geblieben, 2% von „gesund“ nach „krank“ verschlechtert) und fällt damit gegenüber den anderen Maßen klar aus dem Rahmen. Mit subklinischer Kategorie beträgt die Besserungsrate 56% (bei 38% klinisch relevant geblieben und 6% entweder von „unauffällig“ nach „subklinisch“ oder von „subklinisch“ nach „krank“ verschlechtert) und entspricht damit eher den Besserungsraten der anderen Symptommaße.

Im Rahmen der Diskussion um klinische Signifikanz vorgeschlagene Kriterien (z.B. „gesund“ gleich zwei Standardabweichungen außerhalb der dysfunktionalen Population und vice versa) finden ihre Begrenzung meist darin, dass sich dann funktionaler und dysfunktionaler Bereich so stark überlappen, dass kaum noch Veränderungen gefunden werden. Solch strenge Kriterien sind somit nicht geeignet, den Therapieerfolg unserer derzeitigen Behandlungsmethoden zu bewerten; die Entsprechung bei der statistischen Signifikanz wäre eine zu geringe Teststärke.

Letztlich gibt es noch keine Konventionen, wie Besserungskriterien festgelegt werden. Die hier gewählte Vorgehensweise ist: „subklinisch“ definiert als „mäßige“ Beeinträchtigung bzw. als der Bereich außerhalb der jeweiligen halben Standardabweichung um die Mitte eindeutig unauffälliger und eindeutig dysfunktionaler Population. Dies erscheint uns plausibel, transparent und praktikabel für die Evaluation von Therapieerfolg zu sein. Um Vergleichbarkeit zu Ergebnissen anderer Studien zu gewährleisten, in denen eine dichotome

Sichtweise verwendet wurde, wurden die entsprechenden Ergebnisse allerdings für beide Methoden berichtet.

5.2.3 Anregungen für die Therapieerfolgsmessung

Bei der Vielzahl möglicher Meßebenen und Meßmethoden drängt sich die Frage auf, inwieweit sich die (Teil-) Ergebnisse integrieren bzw. aggregieren lassen. Da ein Konsens darüber besteht, dass Therapieerfolg mehrere Dimensionen bzw. viele Facetten hat, ist ein übergreifender Gesamtindex wenig aussagekräftig – auch wenn es für das Argumentieren mit Gesundheitspolitikern und Ökonomen sicher von Vorteil ist, die Effektivität einer Maßnahme möglichst prägnant darzustellen. Für die Ermittlung und Darstellung von Therapieerfolg im Rahmen von Effektivitätsstudien (und auch bei der Qualitätssicherung) sollte ein Kompromiß gefunden werden, der der Anforderung „so differenziert wie nötig und so einfach wie möglich“ genügt. Es ist zudem unrealistisch, dass eine breite Einigung darüber erzielt wird, welche Modalitäten in einen solchen Gesamtindex wie gewichtet einfließen sollten.

Vertretbar erscheint die Zusammenfassung innerhalb einer jeweiligen Modalität und die getrennte Ergebnispräsentation über die Modalitäten hinweg. In der vorliegenden Studie wurden zusammenfassend die Besserungsraten und Effektstärken der Beeinträchtigungsbereiche, der störungsspezifischen und der störungsübergreifenden Symptomatik gemittelt und dann für jede Modalität getrennte Kosten-Effektivitäts-Indices dargestellt. Diese können mit anderen Befunden verglichen werden, auch wenn in anderen Studien „Beeinträchtigung“ oder „störungsspezifische Symptombelastung“ ggf. anders ermittelt wurde. Für eine Einigung auf die Auswahl geeigneter Maße und zur Standardisierung von Diagnostik und Outcome-Evaluation gibt es bereits einige Ansätze (z.B. Fydrich et al., 1996; Lambert & Hill, 1994; Strupp et al., 1997).

In Kasten 5.1 sind weitere, in dieser Studie nicht verwendete Maße aufgeführt, die ebenfalls wichtige Facetten therapeutischer Veränderungen erfassen. Der Einsatz von zu vielen Maßen in einer Studie gefährdet allerdings die Interpretierbarkeit, so dass jeweils eine gezielte und reduzierte Auswahl getroffen werden sollte, die den Fragestellungen der jeweiligen Studie am besten entspricht.

Kasten 5.1: Weitere sinnvolle Wege in Therapieerfolgsmessung und Begleitdiagnostik

In der vorliegenden Studie wurden gemäß den Forderungen nach multimodaler und multimethodaler Erhebung (vgl. Schulte, 1993) drei Perspektiven der Erfolgsmessung gewählt: globale Veränderungsbeurteilung, Veränderung bezüglich der Beeinträchtigungen durch die Störung und die Symptomatik (störungsspezifisch und –unspezifisch). Folgende Perspektiven konnten – leider – nicht oder nur in geringem Umfang in der vorliegenden Studie realisiert werden, sollten aber bei zukünftigen Studien und Auswertungen in Betracht gezogen werden:

- Mehr externe Maße

Die hier verwendeten Daten spiegeln vor allem die subjektive Patientensicht wider (einzige berücksichtigte Fremdbeurteilung: globales Therapeutenurteil). Es wäre lohnend, dies in weiteren Studien mit mehr externen Indikatoren zu vergleichen (z.B. Fremdbeurteilung durch Angehörige oder Hausarzt, Einbezug von Krankheitskosten)

- Einbezug weiterer (Kontroll-) Dimensionen

An hier nicht berücksichtigten, aber gleichwohl lohnenden Bereichen für Eingangs-, Begleit- und Outcome-Diagnostik seien genannt: Achse-II-Störungen und Syndrome bzw. Einbezug von Persönlichkeitsvariablen, Achse-III-Befunde (körperlicher Zustand), Achse-IV-Befunde (aktuelle und überdauernde Stressoren), Konflikt-Messung und Ressourcen-Diagnostik (vgl. Grosse Holtforth & Regli, 1999). Diese Variablen eignen sich auch bei Fragen zur Indikation innerhalb der Therapie und als Kontrollvariablen, wenn Überlegungen zu Prozeß und Wirkungsweise angestellt werden sollen. Ebenfalls häufig empfohlen wird die Erfassung individueller Zielerreichung (z.B. Goal Attainment Scaling), wobei es fraglich ist, ob in einer großen Stichprobe und in einem Setting wie dem vorliegenden diesbezügliche statistische Auswertungen sinnvoll sind (mangelnde Vergleichbarkeit aufgrund von Schwierigkeiten der Standardisierung von Zielbereichen und deren Bewertung). Weiterhin wird die Forderung nach systematischer Erfassung möglicher Nebenwirkungen oder Therapieschäden aufgestellt (vgl. Jacobi, 2001).

- Einbezug von nicht unmittelbar therapielevanten Bewertungen

Vor allem unter Qualitätssicherungsaspekten bietet es sich bei der Untersuchung der Umsetzung einer empirisch fundierten Behandlungsform in den Praxisalltag an, auch die Patientenzufriedenheit im Sinne einer *Kundenzufriedenheit* zur Therapie im Allgemeinen zu erfassen. Die Erfolgserwartung des Patienten zu Therapiebeginn stellt einen wichtigen Prädiktor für den letztendlichen Outcome dar; diese Erfolgserwartung ist sicherlich auch abhängig von der wahrgenommenen Professionalität des Behandlungssettings (vgl. Grawe, 1998). Auch die Zufriedenheit der Behandler mit den Rahmenbedingungen ihres Tuns kann in einem Ansatz des *Total Quality Managements* einbezogen werden (vgl. Fiegenbaum et al., 1997).

5.2.4 Validität und Design:

Wegen des „schwachen“ prä-post-Designs (vgl. Cook & Campbell, 1979) sollten die gefundenen Ergebnisse hinsichtlich ihrer (internen) Validität hinterfragt werden.

5.2.4.1 „Die prä-post-Veränderungen könnten auch ohne diese Therapie zustande gekommen sein“

Die untersuchte Behandlungsform hat in zahlreichen kontrollierten Studien ihre Wirksamkeit unter Beweis gestellt. Bei unbehandelten Kontrollgruppen erreichen die Effektstärken selten 0.2. Ferner wurde bei epidemiologischen Langzeituntersuchungen regelmäßig niedrige Spontanremissionsraten gefunden. Es gibt zudem Hinweise darauf, daß sich das Inanspruchnahmeverhalten und die Arbeitsunfähigkeit bei störungsunspezifischer Behandlung nicht über die Zeit hinweg verringert.

Die in dieser Replikation unter Praxis-Bedingungen gefundenen Effekte sind mit den Effekten vergleichbar, die die entsprechenden Behandlungsgruppen in den kontrollierten Wirksamkeitsstudien erreichen (Effektstärken zwischen 0.9 und 1.9 in Abhängigkeit von der Messmodalität). Gegebenenfalls könnte man von den vorgefundenen Effekten noch die (niedrigen) Effekte von unbehandelten Kontrollgruppen aus Wirksamkeitsstudien abziehen, um so auf konservative Weise lediglich eine inkrementelle Effektstärke abzuschätzen.

Die prä-post-Veränderungen sind durchgehend statistisch hochsignifikant. Die unter der Perspektive klinischer Signifikanz ermittelten Besserungsraten erscheinen ebenfalls als deutlich (zwischen 63% und 79% in Abhängigkeit von der Messmodalität).

So lassen es zunächst die *Einordnung in Vorbefunde* und die *Größe* der gefundenen Effekte plausibel erscheinen, daß die Effekte durch die Therapie erklärt werden können. Zudem ist die Stichprobe relativ groß (N=493).

Auch die Tatsache, daß die Effekte, die von prä nach post gefunden wurden, anschließend bis zur 1-Jahres-Katamnese stabil bleiben, spricht deutlich für den Einfluß des Treatments und gegen Zufallseinflüsse oder Spontanremission.

Schließlich kann noch ein weiterer explorativer Befund zur Attribution der Patienten angeführt werden, der in die Richtung weist, daß die Therapie ursächlich für die gefundenen Besserungen verantwortlich ist. Im Zusammenhang mit der Kosten-Nutzen-Analyse wurden einige Patienten nach ihrer Einschätzung der Wichtigkeit der Therapie für ihre Veränderungen befragt. Die Bewertung der Wichtigkeit der Therapie für stattgefundene Lebensveränderungen wurde durchgehend als hoch bewertet; allerdings konnten die Ergeb-

nisse dieser (Neben-) Erhebung aufgrund kleiner Stichprobengröße und unklarer Selektion dieser Gruppe nicht statistisch abgesichert werden.

Die Summe dieser Argumente bei gleichzeitig fehlender Theorie für systematische Alternativ-Ursachen für die vorgefundenen Effekte lässt darauf schließen, daß es auch ohne Absicherung durch ein experimentelles Design sehr wahrscheinlich ist, dass die Effekte auf die Intervention zurückzuführen sind.

5.2.4.2 „Die Behandlung ist aufgrund niedriger Standardisierung nicht repräsentativ, so daß die Ergebnisse nicht auf andere Behandler oder Behandlungseinrichtungen generalisiert werden können.“

Gegen diese Hypothese spricht, dass die Behandler angehalten waren, sich bei der Therapie auf empirisch fundierte störungsspezifische Störungsmodelle und Behandlungstechniken zu beziehen. Dabei stehen vor allem die Vermittlung eines individuellen Krankheitsmodells, die Modifikation aufrechterhaltender Störungsfaktoren und die Motivation zu selbstgesteuerter Angstbewältigung im Mittelpunkt. Auf etablierte Techniken wie etwa Hausaufgaben zur Selbstbeobachtung, Konfrontation mit angstausslösenden Reizen, Disputation irrationaler Annahmen etc. kann in diesem Rahmen je nach Bedarf zurückgegriffen werden. Diesbezügliche Literatur und Weiterbildungsmöglichkeiten sind mittlerweile überall zugänglich. Alle Behandler durchliefen eine moderne kognitiv-verhaltenstherapeutische Ausbildung und wurden regelmäßig von anerkannten Supervisorinnen und Supervisoren supervidiert. Trotz fehlender Manualisierung kann also von einer einheitlichen und replizierbaren Therapiestrategie gesprochen werden. Ob sie allerdings auch auf die breite Versorgungsrealität übertragbar ist, bleibt fraglich, da die qualitätssichernden Maßnahmen in dieser Studie untersuchten Behandlungseinrichtungen sicher als überdurchschnittlich bezeichnet werden müssen.

5.2.4.3 „Die Ergebnisse dieser Studie können nicht auf alle Personen mit Angststörungen übertragen werden, da eine (Selbst-) Selektion bezüglich soziodemografischer Variablen und Motivation stattfand.“

In der Tat sind dies Ergebnisse, die an einer Stichprobe gewonnen wurde, die zur Behandlung bereit war, sie in vielen Fällen selbst aufgesucht und auch „in relevanter Dosis“ erhalten hatte. Frühe Abbrecher (unter 10 Behandlungseinheiten) wurden nicht weiterverfolgt und nachuntersucht. Auch aufgrund des

überdurchschnittlich hohen Bildungsniveaus kann die Repräsentativität für alle Angstpatienten angezweifelt werden²⁴.

Aus unserer Sicht stellt diese Einschränkung kein grundsätzliches Problem dar. Denn die Grundgesamtheit, aus der diese (relevant durch mindestens eine Angststörung beeinträchtigte, therapiesuchende, motivierte, gebildete) Stichprobe entstammt, ist so groß, dass allein aus pragmatischen Gründen empirische Aussagen über die Kosten-Effektivität der Behandlung bei dieser Gruppe unter dem Versorgungsaspekt relevant sind und ggf. in Allokationsentscheidungen einbezogen werden können. Der Bedarf an dieser Behandlungsform allein in dieser umgrenzten Gruppe dürfte das Angebot dieser Behandlung durch qualifizierte Psychotherapeutinnen und Psychotherapeuten bei weitem übersteigen.

Es bestand also kein Anspruch darauf, universelle Aussagen zu machen. Gleichwohl ist es durchaus möglich, dass auch der „ungebildete, extern zur Behandlung gedrängte, interaktionsgestörte, schlecht motivierte Patient mit rein organischem Störungsverständnis und sekundärem Krankheitsgewinn“ von dieser Art der Psychotherapie profitiert; hierzu sind allerdings weitere Forschungsbemühungen nötig (unter gewisse Modifikationen und Ergänzungen der Behandlungsstrategie).

5.2.4.4 „Es gibt eine systematische Verzerrung dahingehend, dass „schwierige“ und nicht erfolgreiche Patienten nicht nachuntersucht wurden.“

Diesem wichtigen Punkt wurde unter 3.5 ein eigener Abschnitt gewidmet, in dem eine Gruppe von nicht nachuntersuchten Patienten mit unserer Stichprobe hinsichtlich soziodemografischer und störungsspezifischer Variablen verglichen wurden. Es fanden sich keine Unterschiede oder Anhaltspunkte, die diese Hypothese untermauern. Hierbei muss angewerkt werden, dass diese Aussage nicht trivial ist. Die Teststärke hätte ausgereicht, um Differenzen zwischen den Drop-outs bzw. missings und der Treatmentgruppe aufzudecken.

Zudem gleichen sich die Ergebnisse mit denen aus früheren Auswertungen von Therapien in den untersuchten Behandlungseinrichtungen der Christoph-Dornier-Stiftung, in denen im Rahmen eines Forschungsprojekts nahezu 100% der behandelten Patienten nachuntersucht worden waren.

Die Rahmenbedingungen der Datenerhebung legen die Sichtweise nahe, daß „drop-outs“ bzw. missings bei den Nachuntersuchungen unsystematischer Natur sind. Auch die Größe der Stichprobe macht es eher unwahrscheinlich,

²⁴ Das Validitätsproblem vieler kontrollierter Studien, in denen Komorbidität ausgeschlossen wurde (was angesichts epidemiologischer Ergebnisse ebenfalls die Generalisierbarkeit auf alle Angstpatienten stark beschränkt) bestand allerdings nicht.

daß systematische Selektionseffekte die Ergebnisse entscheidend ins Positive verzerrt haben könnten.

5.2.5 Forderungen an eine verbesserte standardmäßige zukünftige Forschung

Auch wenn die Ergebnisse im vorangegangenen Abschnitt für hinreichend valide erachtet wurde, wird empfohlen, in weiteren Studien wenn möglich mit Vergleichsgruppen zu arbeiten. Ebenso sind zusätzliche längere Katamnesezeiträume wünschenswert, um die Langzeitstabilität nicht nur aus Gründen der Plausibilität zu proklamieren (Vorbefunde aus kontrollierten Wirksamkeitsstudien und Epidemiologie, 1-Jahres-Stabilität nach hochsignifikanten Besserungen prä-post), sondern sie konkret empirisch abzusichern. Da in den untersuchten Behandlungseinrichtungen auch 5-Jahres-Katamnesen durchgeführt werden, könnte dies in diesem Setting schon bald stattfinden.

Das Design würde auch durch eine zusätzliche prä-Messung im Sinne einer baseline – etwa bei bestehender Warteliste ein halbes Jahr vor Behandlungsbeginn – gestärkt werden; hierbei müsste jedoch überprüft werden, ob die Symptomatik möglicherweise durch die Hoffnung und Erleichterung durch die Aussicht auf die Therapie beeinflusst wird.

Hinweise zur möglichen Verschränkung von Therapieforschung und Qualitätssicherung im ambulanten Bereich geben Jacobi, Poldrack & Hoyer (in Druck). Insbesondere im Hinblick auf die Misserfolgsvorschung sind verstärkte Bemühungen notwendig, was die Dokumentation und Nachuntersuchung von *Therapie-Abbrechern* betrifft (Jacobi, 2001).

Um eine qualitativ hochwertige Forschung – gerade unter Praxisbedingungen – zu gewährleisten, müssen notwendigerweise die materiellen Rahmenbedingungen stimmen, sprich: Implementation und Aufrechterhaltung der Datenerhebung sollte professionell, nachhaltig und angemessen vergütet z.B. von externen Forschungsstellen übernommen werden. Die Datenqualität wird unnötig reduziert, wenn das Datenmanagement Praktikern überlassen wird, die die Daten idealistischerweise unbezahlt oder unterbezahlt erheben, oder wenn in Versorgungsambulanzen diese Arbeit durch häufig wechselnde Praktikanten, Promovenden, Diplomanden etc. erledigt wird.

5.2.6 Fazit Therapieerfolg: Klinisch-psychologische Behandlungsmethoden sind wirksam und effektiv, kommen aber in der Versorgungsrealität häufig nicht zum Zuge

Im Rahmen dieser Studie konnte gezeigt werden, dass klinisch-psychologische Behandlungsmethoden modernen Zuschnitts (hier: störungsspezifische kognitive Verhaltenstherapie) in ihrem Indikationsgebiet (hier: Angststörungen) bei der Mehrzahl der Patienten effektiv sind. Die gefundenen Effekte werden als groß erachtet und entsprechen den in Meta-Analysen gefundenen. Zudem sind die Verbesserungen breit, d.h. finden sich nicht nur bei der störungsspezifischen Symptombelastung, sondern auch in den anderen übergreifenden untersuchten Modalitäten. Außerdem bleiben die Effekte im Schnitt bis zur 1-Jahres-Katamnese stabil – ein ebenfalls bedeutsamer Befund. Schließlich ist auch noch zu betonen, dass die Behandlungen nicht „im Labor“, sondern unter klinischen Alltagsbedingungen (z.B. Probleme mit Kostenerstattung, Komorbidität, viele verschiedene Therapeuten etc.) stattfanden.

Unserer Einschätzung nach können die Ergebnisse – vorsichtig – auf folgende Bereiche generalisiert werden:

- *Patientenseite:* Angststörungen verschiedener Diagnosen aller Schweregrade (incl. Komorbidität, ohne PTSD), Bereitschaft zur kognitiven Verhaltenstherapie (bei z.T. aktivem Suchprozess). Die Schichtzugehörigkeit bzw. das Bildungsniveau tendierte hier zwar zu einer Selektion in Richtung eines hohen Bildungsstandes; es fand sich aber kein Zusammenhang zwischen Bildungsstand (oder anderer soziodemografischer Variablen) auf die Kosten-Effektivität, so dass empfohlen wird, diesbezüglich vorläufig keine Einschränkungen der Generalisierbarkeit anzunehmen.
- *Behandlungsseite:* störungsspezifische kognitive Verhaltenstherapie, wie sie in zahlreichen Quellen und Manualen beschrieben und in Supervisionen moderner Verhaltenstherapie-Weiterbildung vermittelt wird; ambulante Behandlung; flexible und selbstkorrigierende Therapie mit einem Umfang von durchschnittlich etwa 40-50 Behandlungseinheiten (à 50 Minuten) mit der Möglichkeit, im Rahmen von Konfrontationsübungen mehrere Stunden an aufeinanderfolgenden Tagen zu absolvieren.

Ob die Ergebnisse auch für Kurzzeit-Verhaltenstherapie (insgesamt bis zu 30 Behandlungseinheiten) in einem Praxissetting gelten können, das aufgrund von Unterbezahlung (wie sie seit Einführung des Psychotherapeutengesetzes in Deutschland grassiert) keine Flexibilität für begleitete Konfrontationsübungen ermöglicht, darf angezweifelt werden und sollte in weiteren Studien überprüft werden. Es muss also betont werden, dass die hier untersuchten Behandlungen wahrscheinlich sachgerechter und qualitätsgesicherter durchgeführt wurden, als dies bei vielen „Verhaltenstherapien“ im

Versorgungsalltag niedergelassener Psychotherapeutinnen und Psychotherapeuten der Fall ist.

5.3 Spezielle Aspekte der Kosten-Effektivitäts-Analyse

5.3.1 Therapiekosten und aufgewendete Ressourcen

Es wurde in dieser Studie der Betriebskostenansatz zur Kostenermittlung eingesetzt, also festgelegt, dass „die Therapie das kostet, was sie kostet“. Dabei kann davon ausgegangen werden, dass in diesem Marktpreis die aufgewendeten Ressourcen der Behandlungseinrichtung enthalten sind²⁵.

Unter den Tisch fallen damit indirekte Kosten (z.B. Zeit des Patienten für Anfahrt und Therapie). Auch intangible Kosten wie Ängste und Stimmungstiefs, die etwa im Zusammenhang mit Konfrontationsübungen oder bei Therapiegesprächen mit aversiven Inhalten entstehen, werden im Betriebskostenansatz nicht berücksichtigt.

Dieser Informationsverlust durch die Beschränkung auf die direkten tangiblen Therapiekosten wird aus zwei Gründen in Kauf genommen. Erstens fallen die indirekte Kosten insofern nur wenig ins Gewicht, da die Patienten die Therapie in ihrer Freizeit absolvierten (eine Krankschreibung für die Zeit einer rein psychologischen Therapie wäre allein schon aus rechtlichen Gründen nicht möglich gewesen). Es wäre schwierig geworden, diese aufgewendete Freizeit monetär zu bewerten (d.h. in DM anzugeben, was diese Zeit wert gewesen wäre, wenn sie von den Patienten anders genutzt worden wäre). Der Zusatzaufwand für eine umfassende Kostenberechnung scheint problematisch, zudem die monetäre Quantifizierung intangibler Kosten ohnehin noch unzureichend erforscht ist.

Zweitens hätte nach unserer Einschätzung damit auch die Vermittelbarkeit der Ergebnisse gelitten, weil der Therapiekosten anschaulich ist und indirekte und intangible Kosten mit anderen ambulanten ärztlichen und psychotherapeutischen Leistungen vergleichbar sind, so dass sie nicht eigens einbezogen werden müssen. Dies wäre anders, wenn die Therapiekosten der hier untersuchten Behandlung mit denen einer stationären Behandlung verglichen werden würden (u.a. da dort für die Zeit des Aufenthaltes Arbeitsunfähigkeit gegeben ist).

²⁵ Dies wäre z.B. dann nicht der Fall, wenn eine Behandlung staatlich unterstützt wird und etwa in einer öffentlichen Beratungsstelle für kleine Unkostenbeteiligungen angeboten werden kann, oder wenn im Rahmen eines psychotherapeutischen Weiterbildungsanges von Ausbildungskandidaten Therapien umsonst durchgeführt werden.

5.3.2 Die Kosten-Effektivität als Erweiterung klassischer Outcome-Forschung

Unter Einbezug der Kostenseite – also des *Inputs* für die Behandlung – wird die Bewertung einer Behandlungsform um eine Dimension erweitert. Die Kosten-Effektivität stellt somit eine transparentere, differenziertere Entscheidungsgrundlage für Ressourcenallokationen im Gesundheitswesen bereit als bisherige eindimensionale Ansätze (Betrachtung *entweder* der Kostenseite *oder* der Outcome-Seite). Beim Vergleich mehrerer Behandlungsalternativen sollten in Abhängigkeit von der Perspektive u.U. auch indirekte und intangible Kosten einbezogen werden.

Nützlich ist der Einbezug von Effektivitäts-Untergrenzen und Kosten-Obergrenzen, um den sinnvollen Bereich abzustecken, in dem die Auswahl eines Behandlungsverfahrens auf dem direkten Vergleich von Kosten-Effektivitäts-Indices basieren kann. Denn sowohl die Zuweisung eines Patienten zu einer billigen, aber unzureichend effektiven Behandlung A als auch die Wahl einer hocheffektiven, aber extrem teuren Behandlung B ist ökonomisch nicht vernünftig, auch wenn A und B möglicherweise ebenso gute Kosten-Effektivitäts-Indices aufweisen würden wie Verfahren mittlerer Effektivität und Preisklasse.

Da der Forschungsstand zu Kosten-Effektivitäts-Studien im Psychotherapiebereich bislang relativ dünn ist, wird es noch eine Weile dauern, bis die Kosten-Effektivität wirklich systematisch zur Bewertung von Therapiealternativen genutzt werden kann.

5.3.3 Optimierung der Kosten-Effektivität einer Behandlung

Die vorliegende Studie kann auch als Pionierarbeit betrachtet werden, diese Methode zu Evaluations- und Qualitätssicherungszwecken einzusetzen. Denn auch innerhalb einer Behandlungseinrichtung finden ständig (zumindest implizite) Ressourcenallokationen statt, die durch Kosten-Minimierung und Effektivitäts-Maximierung (via richtiger Indikation zu Art und Umfang von Interventionen) optimiert werden können. Dazu soll im Folgenden auf das vorgefundene Fehlen eines Zusammenhangs von Therapiedosis und Therapieerfolg eingegangen werden.

Während die verschiedenen Outcome-Modalitäten und -Methoden miteinander zusammenhängen (durchgehend signifikante Korrelationen zwischen den Therapieerfolgsmaßen), erwies sich die Kostenseite im Rahmen der Behandlung der untersuchten Behandlungseinrichtungen als unabhängig vom Therapieerfolg: bis auf eine signifikante (mit 0.12 niedrige) Korrelation zwi-

schen Therapieumfang und der Differenz der Beeinträchtigung im Freizeitbereich finden sich keine Zusammenhänge.

Auch bei der deskriptiven Gegenüberstellung (z.B. in Kosten-Effektivitäts-Matrizen) von Kosten und Erfolg wurden keine systematischen Zusammenhänge gefunden: die Verteilung von „billig“ und „teuer“ war für erfolgreiche und nicht erfolgreiche Therapien gleich. Dies kann in erster Linie durch settingspezifische Einflüsse erklärt werden, da die Behandlungsdauer nicht standardisiert, sondern selbstkorrigierend war.

Der Einfluss der Therapiedosis auf den Therapieerfolg kann mit dem Auflösungsgrad dieser Studie nicht genauer untersucht werden, da dazu die einzelnen Therapien inhaltlich folgenden sieben Varianten zugeordnet werden müssten:

kurze Therapie, erfolgreich

⇒ *erfolgreiche minimale Intervention*

kurze Therapie, nicht erfolgreich

- ⇒ *Abbruch wegen ausbleibendem Eingangserfolg* bzw. enttäuschter Erwartungen
- ⇒ *Abbruch aufgrund externer Einflüsse* (z.B. Begrenzung der Therapiedauer aus Kostengründen)

lange Therapie, erfolgreich

- ⇒ *erfolgreiche Langzeitbehandlung* (die bei kleinerem Umfang nicht erfolgreich gewesen wäre)
- ⇒ *erfolgreiche nicht-minimale Intervention* (bei der „großzügig“ mit weiterer zur Verfügung stehender Behandlungszeit umgegangen wurde, obwohl sie auch bei kleinerem Umfang ausreichend erfolgreich gewesen wäre)

lange Therapie, nicht erfolgreich

- ⇒ *Abbruch wegen ausbleibendem Erfolg*
- ⇒ *Abbruch bei ausbleibendem Erfolg aufgrund externer Einflüsse* (z.B. Begrenzung der Therapiedauer aus Kostengründen)

Es gibt unter Gesichtspunkten der Kosten-Effektivität theoretisch sowohl die Möglichkeit einer *zu hohen* Therapiedosis („Verschwendung“) als auch die einer *zu geringen* Therapiedosis (therapeutisch nicht indizierter Abbruch trotz realistischer Erfolgchancen); als Analogon könnte man hier die Unterdosie-

rung oder ein zu frühes Absetzen eines Serotonin-Wiederaufnahme-Hemmers anführen.

Dies ist von therapeutischer Seite aus im Rahmen einer selbstkorrigierenden Therapieplanung, bei der hinsichtlich des Therapieumfangs eine gewisse Flexibilität herrscht, zu berücksichtigen. Auch wenn dies auf Grundlage unserer Datenbasis nicht empirisch weiter verfolgt werden kann, so können doch ausgehend von dieser Systematik Empfehlungen für die Steigerung der Kosten-Effektivität einer Behandlung gegeben werden (Kasten 5.1).

5.3.4 Variabilität des Kosten-Effektivitäts-Index in Abhängigkeit von Auswertungsmethode und (Sub-) Stichproben

Da der Einsatz verschiedener Messmodalitäten und Messmethoden konzeptuell damit begründet wird, dass für „den“ Therapieerfolg kein einheitliches allumfassendes Maß gefunden werden kann, liegt es nahe, für jede Modalität einen eigenen Index anzugeben. Dies scheint bei drei Modalitäten (global, Beeinträchtigung, Symptombelastung) noch überschaubar zu sein; mögliche Erklärungen für die gefundenen Unterschiede wurden bereits in den vergangenen Abschnitten genannt.

Auch der *Zähler* des Kosten-Effektivitäts-Index (Therapiekosten) variiert in dieser Untersuchung. In einem der vier Behandlungsinstitute waren die Behandlungen signifikant länger, in einem anderen wurde ein niedrigerer Stundensatz in Rechnung gestellt. Auch hier bietet sich ein transparentes Vorgehen an, so dass die Kosten-Effektivität bei Zugrundelegung anderer Therapiekosten erneut ermittelt werden kann. So könnten z.B. die Kosten steigen (und damit die Kosten-Effektivität sinken), wenn der Punktwert ambulanter Verhaltenstherapie auf 10 Pfennig (entspricht DM 145.- pro Behandlungseinheit) erhöht würde. Dieser Wert wurde laut eines Urteiles des Bundessozialgerichts vom November 1999 für die Jahre vor 1999 als im Vergleich zu anderen medizinischen Abrechnungsleistungen für „gerecht“ erachtet. Umgekehrt ist aber der Punktwert nach 1999 auf z.T. unter 5 Pfennig gesunken. Hier wäre es aber fragwürdig, die Kosten-Effektivität mit gleichem *Nenner* neu zu berechnen (bzw. „schönzurechnen“), da nicht ausgeschlossen werden kann, dass ein derartiger existenzbedrohender Preisverfall sich auf die Qualität der Therapien auswirkt (z.B. Verzicht auf begleitete Konfrontationsübungen wegen Organisationsproblemen).

Die Variabilität der Kosten-Effektivität bezüglich störungsspezifischer Variablen entsprach den Erwartungen: die Behandlung war generell wirksam, aber in unterschiedlichem Ausmaß und bei unterschiedlichen Kosten. Ein eingangs höherer Schweregrad war gegenüber dem Durchschnitt zwar mit größeren Besserungen, aber mit noch längerer Behandlungsdauer (und damit ungünstigerer Kosten-Effektivität) verbunden.

Kasten 5.1: Empfehlungen zur Steigerung der Kosten-Effektivität einer Behandlung

Die folgenden Richtlinien können für die Optimierung der Kosten-Effektivität einer definierten Behandlung herangezogen werden.

1. Anstreben erfolgreicher minimaler Intervention

Wenn erfolgreiche Therapien frühzeitig beendet werden (bzw. bereits zu Therapiebeginn als Minimalintervention konzipiert wurden), wird die Kostenseite minimiert²⁶

Materielle Hintergrundeinflüsse („Wenn die Therapie schon bewilligt wurde, soll das zur Verfügung stehende Budget auch ausgeschöpft werden, denn sonst ist die Arbeit nicht rentabel!“) oder der verständliche Wunsch von Patientenseite nach dem „Luxus“ weiterer therapeutischer Begleitung bei hohem Budget können mögliche minimale Interventionen vereiteln.

2. Frühzeitiger Abbruch der Therapie bei niedrigen Erfolgschancen

Wenn die Therapien frühzeitig identifiziert werden können, die am Ende doch erfolglos bleiben, werden so ebenfalls die durchschnittlichen Behandlungskosten minimiert. Zudem werden intangible Kosten einer nicht erfolgreichen Behandlung (langanhaltende Frustration, Verhinderung eines Therapieerfolgs im Rahmen einer möglicherweise passenderen Alternativbehandlung) eingespart.

Schwierigkeiten entstehen hier, wenn gute Prädiktoren für den Therapiemißerfolg fehlen – was in der Regel der Fall sein dürfte. Zudem können auch bei Vorliegen solcher Prädiktoren menschlich durchaus nachvollziehbarer naiver Optimismus sowie o.g. materielle Hintergrundeinflüsse einen frühzeitigen therapeutisch gerechtfertigten Abbruch verhindern.

3. Richtige Indikation zu einer hohen Therapiedosis

Wenn aussichtsreiche Therapien innerhalb ihres (indizierten) Langzeitumfangs erfolgreich abgeschlossen werden, wird so die Besserungsrate maximiert, was die erhöhten Kosten zumindest ausgleicht. Eine aufgrund zu frühen Abbruchs erfolglose Behandlung nützt weder der Behandlungseinrichtung (deren Besserungsrate sich verschlechtert) noch dem Kostenträger, der u.U. in der Folge für andere, neue Behandlungen aufkommen muss, obwohl zur klinisch signifikanten Besserung in der ersten Behandlung nur noch wenig gefehlt hätte. Zudem würde der Patient unnötig demoralisiert.

Indizierte Langzeitbehandlungen werden verhindert, wenn sie den Rahmen nicht hinterfragbarer Kosten-Obergrenzen sprengen.

²⁶ Ganz allgemein entspricht es im übrigen einem verhaltenstherapeutischen Selbstverständnis, minimale Interventionen anzustreben (vgl. Margraf & Lieb, 1995); denkbar ist sogar eine Schmälerung des Therapieerfolgs bei zu viel Intervention (z.B. aufgrund Verwirrung und Interferenz, Hemmung von Selbstwirksamkeit und Selbstmanagement oder durch Abwertung des Erfolges aufgrund subjektiv als zu hoch bewerteter Kosten)

Die Behandlung unterschiedlicher Störungen unterschieden sich in bezug auf die Kosten. Zwangstherapien waren am umfangreichsten, was den Vorbefunden entspricht (vgl. z.B. Reinecker, 1991). Die Effektivität der Sozialphobiebehandlung war am wenigsten erfolgreich. Dies mag mit einer erhöhten Assoziation mit Persönlichkeitsstörungen zusammenhängen (Fydrich & Schowalter, 1998); auch möglich ist eine kritischere Haltung gegenüber dem eigenen Befinden und Befindlichkeitsverbesserungen aufgrund erhöhtem Perfektivismus in dieser Störungsgruppe. Eine dritte Erklärungsmöglichkeit besteht darin, dass bei der Sozialphobiebehandlung das hochwirksame Therapieelement der massierten langandauernden Konfrontation mit Reaktionsverhinderung nicht so einfach umgesetzt werden kann wie etwa bei der Behandlung einer Höhenphobie.

Unterschiede dieser Art verweisen auf die ethische Forderung, Möglichkeiten der Langzeittherapie bei schweren Fällen bereitzuhalten und nicht alle Störungen über einen Kamm zu scheren. So darf etwa der Behandlung von Störungen, die als besonders „schwierig“ erachtet werden (z.B. Alkoholabhängigkeit, „Borderline“-Spektrum) die Finanzierung nicht deshalb verweigert werden, weil sie eine schlechtere Kosten-Effektivität aufweisen als die hier untersuchte ambulante Angstbehandlung.

5.4 Kosten-Nutzen-Analyse: Zusammenfassung

In der Kosten-Nutzen-Studie sollte geprüft werden, ob und in welchem Ausmaß psychologische Psychotherapie einen finanziellen Nutzen produzieren kann. Dass die untersuchte Behandlungsform – störungsspezifische kognitive Verhaltenstherapie – effektiv ist, wurde bereits in der Kosten-Effektivitäts-Studie der vorliegenden Arbeit gezeigt. Als Nutzen wurde die Einsparung tangibler Krankheitskosten nach der Therapie verstanden (Reduktion der Inanspruchnahme von Gesundheitsleistungen und Arbeitsunfähigkeit im Jahr nach Therapie). Dieser Nutzen wurde den Kosten der Behandlung (Betriebskostenansatz) gegenübergestellt.

Die Ermittlung der Krankheitskosten vor und nach Therapie erfolgte an einer Stichprobe von 67 Angstpatienten (Komorbidität zugelassen), die in vier Ambulanzen der Christoph-Dornier-Stiftung eine ambulante Psychotherapie absolviert haben.

Im Jahr nach Therapie reduzierten sich Kostenfaktoren beträchtlich (stationäre Kosten um 25%, restliche Kostenfaktoren 87% - 100%). Die dadurch erzielten Einsparungen betragen durchschnittlich DM 3026.- Das Verhältnis von Therapiekosten und Einsparungen beträgt damit im ersten Jahr 1 : 0.58. Unter der Annahme, dass die Auswirkungen der Therapie auch über die 1-Jahres-Katamnese hinaus stabil bleiben, kann man abschätzen, dass sich die Therapie innerhalb des zweiten Jahres amortisiert und das Kosten-Nutzen-Verhältnis nach fünf Jahren (diskontiert) 1 : 2.63 beträgt. Die Sensitivitätsanalyse ergab, dass das Kosten-Nutzen-Verhältnis auch bei einer Variation der geschätzten oder aus Statistiken entnommenen Preisen für die Kostenfaktoren stabil bleibt. Bei der Berechnung half der eigens für derartige Routinen programmierte *cost benefit calculator*.

Diese Ergebnisse wurden mit denen einer Querschnittsstudie verglichen, in der die Krankheitskosten von N = 213 Patienten vor Behandlung mit denen von N = 205 Patienten zur 1-Jahres-Katamnese in Beziehung gesetzt wurden. Das Kosten-Nutzen-Verhältnis war vergleichbar (etwas günstiger in der umfangreicheren Querschnittsstudie), so dass es prinzipiell sinnvoll erscheint, zukünftig auch mit einfacher zu realisierenden Querschnitts-Designs zu arbeiten.

In einer weiteren Fragestellung wurde auch der intangible Nutzen einbezogen (Größen wie Lebensqualität und Zeitaufwand, denen kein fester Marktpreis zugeordnet werden kann). In einer Pilotstudie zum Ansatz der Zahlungsbereitschaft (*willingness to pay, WTP*) wurden alle verfügbaren Patienten der 1-Jahres-Katamnesen zwischen 1/99 und 5/99 dazu befragt, wie viel Ihnen die Effekte, die bei Ihnen durch die Therapie erzielt wurden, in Geld (und Zeit) wert wären. Der subjektive Behandlungsnutzen, ausgedrückt durch den Median der monetären Zahlungsbereitschaft betrug DM 3500.- (130% des Netto-

einkommens). Darin sind die tangiblen Kosten der Behandlung bereits verrechnet. Da die WTP-Methode aber im Psychotherapiebereich noch nicht etabliert ist und noch kein Konsens zu deren Validität besteht, sollte eine Vermengung tangiblen und intangiblen Therapienutzens nur unter Vorbehalt vorgenommen werden.

5.5 Spezielle Aspekte der Kosten-Nutzen-Analyse

5.5.1 Validität bezüglich Design und Datenerhebung

Die Aussagen der Diskussion der Kosten-Effektivitäts-Analyse zur Generalisierbarkeit bzw. zu Repräsentativität von Behandlung, Setting und Patientpopulation gelten auch für die Kosten-Nutzen-Studie. Uns erscheint es plausibel, dass die Therapieerfolge mit einer Reduktion von Weiterbehandlung und Arbeitsunfähigkeit einhergehen, bzw. dass die gefundenen Einsparungen gut zu diesen Effekten passen.

Außerdem gibt es Hinweise darauf, dass sich die Krankheitskosten im Laufe der Zeit ohne störungsspezifische Behandlung nicht ändern. Sommer (1998) untersuchte an einer Teilstichprobe den Verlauf der Krankheitskosten über die Zeit und fand keine bedeutsamen Unterschiede der Krankheitskosten im ersten, zweiten und dritten Drittel der gesamten Krankheitsdauer vor Therapie. Es scheint also unwahrscheinlich, dass sich die Kosten plötzlich innerhalb eines Jahres spontan so drastisch reduzieren könnten, wie dies in der Kosten-Nutzen-Studie gefunden wurde.

Schließlich muss noch darauf hingewiesen werden, dass das Design keine gesicherten Aussagen über die Langzeitstabilität der gefundenen Einsparungen erlaubt und die langfristigen Kosten-Nutzen-Aussagen lediglich Schätzungen darstellen. Es ist allerdings ebenfalls unwahrscheinlich, dass die deutlichen Effekte nach einem Jahr plötzlich wieder verloren gehen sollten. Dafür spricht auch die Tatsache, dass die Krankheitskosten zwischen Therapieende und Katamnese auf ein Jahr standardisiert auch für diejenigen gleich waren, deren Nachuntersuchungen aus organisatorischen Gründen verspätet durchgeführt wurden. Untermauert werden können diese Überlegungen aber letztendlich nur durch konkrete Nachuntersuchungsdaten aus längeren Katamnesen (z.B. zukünftig auszuwertender 5-Jahres-Katamnesen der untersuchten Behandlungseinrichtungen).

Weitere spezielle Aspekte der Kosten Nutzen-Analyse sind:

5.5.1.1 „Es bleiben Kostenfaktoren und Kostenperspektiven bei der Ermittlung von Kosten und Nutzen unberücksichtigt.“

Dass es Sinn macht, bei der Berechnung der Kostenseite auf den Betriebskostenansatz zurückzugreifen, wurde bereits dargelegt. Zusätzlich kann hier angeführt werden, daß die Berücksichtigung zusätzlicher indirekter und intangibler

Kosten der Therapie verlangt hätte, diese auch für die bei der Erhebung der Krankheitskosten ermittelten Behandlungen zu bestimmen. Dieser große zusätzliche Aufwand hätte in keinem vertretbaren Verhältnis zu einem zusätzlichen Erkenntnisgewinn gestanden.

Im Rahmen der Marginalanalyse wurden allerdings einige für die Nutzenberechnung relevante tangible Kostenfaktoren gefunden, die aus methodischen Gründen nicht in die Kosten-Nutzen-Analyse einbezogen werden konnten. So wurde etwa nach Therapie ein höherer Anteil Erwerbstätiger als vor Therapie gefunden werden, was sich jedoch nicht monetär quantifizieren ließ (betrachtet werden konnten lediglich die durchschnittliche Zahl der Arbeitsunfähigkeitstage bei Erwerbstätigen und nicht der Status der Erwerbstätigkeit an sich). Auch die Kostenfaktoren Berentung, unbezahlter Urlaub aufgrund der Störung und Verminderung der Arbeitsqualität konnten nicht sauber monetär bewertet werden und wurden deshalb aus den Kosten-Nutzen-Berechnungen ausgeschlossen, obwohl anzunehmen ist, daß der entstandene Nutzen durch die in Einzelfällen gefundene Reduktion nach Therapie beträchtlich gewesen wäre.

Ebenfalls nicht einbezogen, weil mit unseren Mitteln nicht quantifizierbar, ist der Nutzen durch verringerte Arbeitsunfähigkeit bei nicht Erwerbstätigen. Gemäß dem Opportunitätskostenprinzip ist auch unbezahlte Arbeit (z.B. als Hausfrau) für eine Gesellschaft nicht umsonst, da sie vielleicht an einem anderen Ort ebenfalls nutzenbringend hätte eingesetzt werden können. Es entsprach auch hier einem sehr konservativen Vorgehen, die von nicht erwerbstätigen angegebenen Arbeitsunfähigkeitszeiten aus der Kosten-Nutzen-Analyse auszuschließen und lediglich den durch die Erwerbstätigen erzielten Nutzen auf die Gesamtstichprobe umzulegen. Würde man die Gegenposition vertreten und die Kosten für Arbeitsunfähigkeit bei nicht Erwerbstätigen denen bei Erwerbstätigen gleichsetzen, so würde sich das Kosten-Nutzen-Verhältnis der Behandlung bereits im ersten Jahr drastisch verbessern (vgl. Runde, 1999). Wir hielten es aber vorläufig für die „sauberere“ Variante, konservativ vorzugehen und die (intangiblen) Kosten der Arbeitsunfähigkeit von Hausfrauen, Schülern und Studenten, Rentnern und Arbeitslosen nicht zu verrechnen.

Ein in der Literatur häufiger erwähnte Posten, der nicht in der Index-Berechnung enthalten war, war der Medikamentenkonsum. Dieser wurde ebenfalls wegen Berechnungsschwierigkeiten zu den nicht monetarisierbaren Marginalfaktoren gezählt. Sein Einfluss wird allerdings z.B. von Zielke (1993, S.409) als derartig klein eingeschätzt, daß er unter alleiniger Betrachtung von Kostengesichtspunkten zu vernachlässigen sei.

Es gibt also in der Tat eine Reihe von im Zuge der Behandlung reduzierten Kostenfaktoren, die nicht in die Kosten-Nutzen-Berechnung mit einbezogen werden konnten. Wenn es für diese Faktoren etablierte Berechnungsgrundlagen gäbe, würde sich das Kosten-Nutzen-Verhältnis weiter verbessern.

5.5.1.2 „Die Daten zu Krankheitskosten beruhen auf falsch erinnerten Werten.“

Es ist möglich, daß die Daten, auf denen die Untersuchung basiert, von den Patienten falsch erinnert wurden. Dabei kann davon ausgegangen werden, daß Arztbesuche, Krankenhaustage und Fehltage am Arbeitsplatz aufgrund von Vergessen in ihrer Häufigkeit unterschätzt wurden. Es ist hingegen wahrscheinlich, dass Patienten Kosten berichteten, die nicht existierten.

Die Validität der Instrumente zur Erhebung von Krankheitskosten wurde anhand von Interviews geprüft, bei denen die Patienten noch einmal sehr detailliert zu den relevanten Kostenfaktoren befragt wurden (Sommer, 1998, Pieper, 1999). Auch wenn keine signifikanten Differenzen zwischen Interviews und Fragebögen gefunden werden konnten, können Vergessensfehler nicht völlig ausgeräumt werden. Derartige Fehler hätten eher ein schlechteres Kosten-Nutzen-Verhältnisses zur Folge, weil für den Zeitraum vor der Therapie viel mehr Möglichkeit bestand, Behandlungen zu vergessen als nach der Therapie (zum einen wegen des kürzeren Zeitraumes und des klaren Zeit-Ankers bei der Katamnesebefragung, zum anderen wegen der durchschnittlich stärkeren Inanspruchnahme vor Therapie).

Optimal zur Bestimmung der Krankheitskosten wären objektive Krankenkassendaten gewesen. Dies ist derzeit in Deutschland allerdings aus datenschutzrechtlichen Gründen nur in Zusammenarbeit mit Privat- und Betriebskrankenkassen möglich (vgl. Zielke, 1993), welchen nur ein geringer Anteil der untersuchten Personen angehörten. Vielleicht können zukünftige Forschungsergebnisse, z.B. aus der im letzten Jahr angelaufenen großen „DKV-Studie“ der Stuttgarter Forschungsstelle für Psychotherapie (<http://www.psyres-stuttgart.de>), Aufschluss darüber geben, inwieweit sich erinnerte Behandlungen und Arbeitsunfähigkeit von objektiven Krankenkassendaten unterscheiden.

5.5.1.3 „Die Kostenberechnung aufgrund von geschätzten Preisen für Kostenfaktoren ist unsicher.“

Die Ermittlung der Preise für die einzelnen Kostenfaktoren wurde ebenfalls im Rahmen dieser Arbeit dargestellt. Zum Teil handelte es sich nur um Durchschnitts- und Schätzwerte. Sie können also nicht exakt die Kosten der einzelnen Patienten widerspiegeln; außerdem können auch die Statistiken, denen sie entnommen wurden, bereits fehlerhaft gewesen sein.

Dennoch fällt dieser Kritikpunkt nicht ins Gewicht. Mit Hilfe der Sensitivitätsanalyse konnte gezeigt werden, daß selbst bei starker Variation der einzelnen Preise die Hauptaussage der Analyse nicht verändert wurde. Erfüllt man also die Forderung von Gesundheitsökonomern nach einer solchen Überprü-

fung der Fehlerrobustheit, kommt es nicht so sehr darauf an, wirklich exakte Werte zu verwenden (zumal es diese eben noch nicht gibt).

5.5.1.4 „Der subjektive Nutzen wurde vernachlässigt“

In der vorliegenden Untersuchung wurde nicht der gesamte Therapienutzen monetär bewertet, sondern nur der sogenannte tangible, also direkt berechenbare Nutzen. Der intangible Nutzen und damit die Verbesserung des subjektiven Befindens wurde vernachlässigt.

Dies birgt zwar einen Informationsmangel in sich, für die gewählte Betrachtungs-perspektive war die Berechnung des intangiblen Nutzens allerdings nicht notwendig. Es ging zunächst nur um die Frage, ob sich die Psychotherapie aus finanzieller Sicht für die Gesellschaft lohnt. Die Perspektive des Patienten ist dennoch wichtig und in Kostenträgerentscheidungen mit einzubeziehen. Bei der Entscheidung für eine Behandlungsalternative muß es immer zunächst um die Betrachtung der Effektivität gehen, bevor Kosteninformationen hinzugezogen werden.

Der subjektive Nutzen wurde im Rahmen dieser Arbeit allerdings in der WTP-Pilotstudie ermittelt und kann den tangiblen Kosten-Nutzen-Ergebnissen an die Seite gestellt werden.

5.5.2 Vereinfachung von Datenerhebung und Berechnung

Die Ergebnisse der längsschnittlichen Kosten-Nutzen-Analyse wurden mit denen einer Querschnittsstudie verglichen, in der die Krankheitskosten von $N = 213$ Patienten vor Behandlung mit denen von anderen $N = 205$ Patienten zur 1-Jahres-Katamnese in Beziehung gesetzt wurden (Pieper, 1999; Runde, 1999). Das Kosten-Nutzen-Verhältnis war vergleichbar (etwas günstiger in der umfangreicheren Querschnittsstudie), so dass dazu ermutigt wird, aus forschungsökonomischen Gründen zukünftig auch mit einfacher zu realisierenden Querschnitts-Designs zu arbeiten.

Die Berechnung des finanziellen Nutzens aus der Reduktion von Kostenfaktoren vorher-nacher (oder zwischen zwei Behandlungsbedingungen) unter Berücksichtigung unterschiedlicher (ggf. diskontierter) Zeiträume sowie die Variation von Preisen in der Sensitivitätsanalyse sind aufwendig. Das im Rahmen dieser Studie entwickelte Programm des *cost benefit calculators* erleichtert derartige Routinen (auf beiliegendem Datenträger Datei cbc.htm).

5.5.3 Der Einbezug der Zahlungsbereitschaft (WTP) zur Ermittlung intangiblen Nutzens

5.5.3.1 Neues Fragebogeninstrument

Entsprechend der Richtlinien, die zum Ansatz der Zahlungsbereitschaft in anderen Gebieten entwickelt wurden, wurde ein auf den Psychotherapiebereich zugeschnittener Fragebogen entwickelt. Die Kontentvalidität wurde anhand von Korrelationen geprüft (z.B. ob höhere Einkommen oder kleine Zahl an Haushaltsmitgliedern zu höherer Zahlungsbereitschaft führen). Akzeptanz und Verständnis des Fragebogens scheinen gegeben; dies sollte jedoch nochmals an einer weniger gebildeten Population als der in der vorliegenden Pilotstudie geprüft werden. Da Fragen zur Zahlungsbereitschaft auch in wenigen Minuten per Interview erhoben werden können, wird empfohlen, den (ggf. auf das jeweilige Setting hin modifizierten) Fragebogen in weiteren Studien als Interviewleitfaden zu benutzen, um Missverständnissen in den dort verwendeten Gedankenspielen vorzubeugen.

Ein Zusammenhang zwischen globaler Therapie-Erfolgs-Einschätzung und der Zahlungsbereitschaft konnte nicht festgestellt werden. Das ist wahrscheinlich auf die geringe Varianz dieser Werte zurückzuführen, die fast alle bei „sehr viel besserem“ und „viel besserem“ Befinden nach der Therapie liegen. Dasselbe gilt für die Variable „Wichtigkeit der Therapie für das Leben“, die als Entsprechung der Frage nach der Zahlungsbereitschaft den Wert der Therapie auf einer Skala erfassen sollte.

Eine Innovation der WTP-Methode stammt von Möser (1998). Hierbei wird der intangible Nutzen ermittelt als die Abnahme der WTP für eine Weiterbehandlung in verschiedenen Therapiestadien. Wurde beispielsweise vor der Therapie ein Betrag von DM 5000.- als akzeptabler Therapiepreis genannt und nach 30 Behandlungseinheiten lediglich DM 300.- für weitere Sitzungen, so wird gefolgert, dass der intangible Nutzen der Maßnahme bis zur 30. Sitzung DM 4700.- beträgt (dass also für weitere Besserungen nicht mehr so viel aufgewendet werden würde). Für Varianten dieser Art müsste das Fragebogeninstrument selbstverständlich ebenfalls nochmals überarbeitet werden.

5.5.3.2 Inhaltlicher Eindruck der Pilotstudien-Ergebnisse

Mit 130% des Haushalts-Nettoeinkommens (Median; Durchschnitt: 168%) und einer deutlichen Bereitschaft, viel Zeit für Anfahrt, Therapie und zugehörigen „Hausaufgaben“ aufzuwenden, erscheint die Zahlungsbereitschaft für störungsspezifische Psychotherapie für Patienten, die in der 1-Jahres-Katamnese dazu befragt wurden, hoch zu sein. Die Ergebnisse lassen sich aber

derzeit noch nicht gut vergleichen, da die wenigen anderen WTP-Studien im Gesundheitswesen nicht im Psychotherapiebereich angesiedelt sind.

5.5.3.3 Anwendung der Methode Zahlungsbereitschaft in der ökonomischen Evaluation: Ausblick

Eine Entscheidung, die nur die Ergebnisse von Kosten-Nutzen-Analyse einbezieht, muss gerade im Gesundheitswesen als verkürzt betrachtet werden. Denn das Ziel einer medizinischen Maßnahme ist die Verbesserung der Lebensqualität des einzelnen unter einer Krankheit leidenden Patienten, nicht die Einsparung von Kosten für Volkswirtschaft und Krankenkassen. Daher ist es unerlässlich, Instrumentarien, die den Einfluss auf die Lebensqualität und den subjektiven Wert von Maßnahmen für die Betroffenen abbilden, einzusetzen und deren Ergebnisse in die Entscheidung einzubeziehen.

Die neue Qualität der Methode Zahlungsbereitschaft liegt in der Möglichkeit des Einbezugs in finanzielle Analysen. Wenn die Methode besser gearbeitet und an verschiedenen Patientengruppen größeren Umfangs überprüft ist, wäre es vorstellbar, die Zahlungsbereitschaft als Netto-Nutzen der Psychotherapie aus Perspektive der Patienten anzusehen.

Dennoch ist die Methode auch kritisch zu betrachten. Bei der Angabe der Zahlungsbereitschaft spielen vermutlich nicht nur Therapie-Outcome-Variablen eine Rolle, sondern auch Einstellungen dem Geld gegenüber (beispielsweise die Angst, die es jemandem verursachen würde, Schulden für eine Therapie zu machen; das Ausmaß, in dem man sich etwas gönnt, das dann auf Kosten der Familienmitglieder geht etc.).

Auch die Erfassung der tatsächlichen finanziellen Situation ist mit Unschärfe behaftet, da sie sich nicht nur aus monatlichen Einkünften, sondern auch aus Grundbesitz und Erspartem etc. zusammensetzt. Sicherlich sollten sich diese Einflüsse bei einer genügend großen Stichprobe ausmitteln. Die Repräsentativität der Stichprobe gegenüber der Grundgesamtheit ist daher unbedingt zu überprüfen.

In der vorliegenden Stichprobe ist der Überschuss an Personen aus den neuen Bundesländern als eine Verzerrung gegenüber der Grundgesamtheit anzusehen, die das Ergebnis systematisch beeinflussen könnte: Es wäre denkbar, dass in den alten und neuen Ländern unterschiedliche Einstellungen dem Geldausgeben gegenüber herrschen, oder dass die Ansicht im Osten weiter verbreitet ist, dass Gesundheitsleistungen gefälligst zum Nulltarif dazusein hätten; ganz abgesehen von unterschiedlichen Besitzverhältnissen.

Letztendlich ist noch nicht klar, wie die Summe der WTP von den Probanden gebildet wird: Wird der qualitative Wert der Therapie betrachtet? Wird errechnet wie viel Einsparungen die Therapie an anderer Stelle bringt (beispielsweise durch Wiederaufnahme der Erwerbstätigkeit)? Werden die subjektiven Kosten wie Angst, Probleme mit dem Partner etc. in die Bildung der

Summe einbezogen? Wird berücksichtigt, dass es noch andere Alternativen der Behandlung gibt, wenn überlegt wird, wie viel man für die Therapie zahlen würde? Oder misst die Zahlungsbereitschaft hauptsächlich, inwiefern ein Patient die Besserung seines Zustandes auf die Therapie oder auf äußere Umstände attribuiert?

Wegen der wichtigen Information, die die Methode im Zusammenhang einer ökonomische Evaluation bieten kann, ist es notwendig, weiter an der Verbesserung der Methode zu arbeiten und die oben genannten Unsicherheiten zu untersuchen. Ziel sollte sein, einen allgemein anwendbaren Fragebogen oder Interviewleitfaden zu entwickeln, damit unterschiedliche Formulierungen etc. die Vergleichbarkeit der Werte nicht erschweren. Es wäre dabei möglich, WTP-Instrumente für verschiedene Ansätze im Gesundheitswesen (ex post/vs. ex ante; Befragung von Betroffenen versus der Allgemeinbevölkerung) zu entwickeln. Auch muss ein Konsens darüber gefunden werden, auf welche Weise und in welchem Bezug die Zahlungsbereitschaft den Kosten einer Intervention gegenüber gestellt werden, also in eine allgemeine Kosten-Nutzen-Analyse einbezogen werden kann.

5.6 Einordnung der Ergebnisse in den Forschungsstand zu Kosten–Nutzen–Analysen in der Klinischen Psychologie

Beim Vergleich der Ergebnisse mit anderen Untersuchungen zur Kosten-Nutzen-Relation von Verhaltenstherapie bei Angststörungen müssen an dieser Stelle ausdrücklich die Grenzen eines solchen Vorgehens betont werden: Es ist aufgrund fehlender Standards derzeit nicht möglich, Kosten-Nutzen-Verhältnisse einfach miteinander zu vergleichen.

Hierfür gibt es verschiedene Gründe. Zum einen können als Vergleichsstudien nur solche Arbeiten ausgewählt werden, die im deutschen Sprach- und vor allem Wirtschaftsraum angelegt wurden, da die gesellschaftlichen, strukturellen und finanziellen Bedingungen in anderen Ländern ganz andere Ergebnisse zulassen. Zum anderen sollte sich die Vergleichsstudie ebenfalls mit einer Therapieform speziell von Angststörungen beschäftigt haben. Eine derartige Studie ist aus Deutschland nicht bekannt.

Weiterhin müssen die verwendeten Preise und Berechnungsmethoden der Studien annähernd übereinstimmen, will man die Ergebnisse miteinander vergleichen. Wie stark die Ergebnisse einer Kosten-Nutzen-Analyse mit beinahe identischem Ausgangsmaterial differieren können, zeigt der Vergleich der Ergebnisse von Längs- und Querschnittstichprobe im Rahmen dieses Forschungsprojekts. Runde (1999) kommt in ihrer Arbeit zu einem Verhältnis von 1 : 2,66 (nach 2 Jahren). In der hier berichteten Untersuchung wurde für denselben Zeitraum eine Relation von 1 : 1,03 bestimmt. Dieser Unterschied wurde bereits an früherer Stelle mit der verschiedenen Vorgehensweise bei der Bestimmung der Arbeitsunfähigkeitskosten erklärt. Runde hat alle berichteten AU-Tage (auch die von eigentlich Nicht-Erwerbstätigen) als solche gewertet und mit einem nach unten bereinigten Wert von 187,30 pro Tag berechnet. In der vorliegenden Studie wurde wie bei Zielke (1993) nur bei denjenigen Personen die Arbeitsunfähigkeit berechnet, die auch angaben, erwerbstätig zu sein; für sie schlug ein Fehltag mit volkswirtschaftlichen Kosten von 335,- DM zu Buche. Der Grund für diese konservative Vorgehensweise war der Zweifel daran, die Kosten für die Arbeitsunfähigkeit für nicht Erwerbstätige monetär richtig bewerten zu können. Würden aber trotz dieser Bedenken die Kosten auf jene Weise berechnet werden, ergäbe sich ein anderes Bild. Das Kosten-Nutzen-Verhältnis wäre mit 1 : 1,97 für 2 Jahre bereits fast doppelt so hoch wie vorher ermittelt und deutlich näher am Ergebnis von Runde (1999). Für den Vergleich von Studienergebnissen mit unterschiedlichen Berechnungsmethoden und Preisgefüge ist der im Rahmen dieser Arbeit entwickelte Cost-Benefit-Calculator hilfreich. Aufgrund der dargestellten Probleme beim Vergleich von Ergebnissen derartiger Untersuchungen soll die Einordnung der vorliegenden Arbeit lediglich auf der Ebene des methodischen Vorgehens

durchgeführt werden. Zwei Studien im deutschsprachigen Raum werden dafür als geeignet angesehen.

Die Kosten-Nutzen-Analyse von Breyer et al. (1997) wurde für die ambulante Psychoanalyse allgemein psychischer Störungen durchgeführt. Sie hatten für Arztkontakte niedrigere Preise ermittelt. Dieser Unterschied mag noch nicht so sehr ins Gewicht fallen, aber der Durchschnittssatz für einen Krankenhaustag lag doch deutlich über den hier verwendeten Werten. Kosten für eine vorherige Psychotherapie oder Notarztkontakte wurden gar nicht berechnet. Die Arbeitsunfähigkeitskosten wurden analog der hier gewählten Vorgehensweise über den Verlust am Bruttosozialprodukt durch krankgeschriebene Erwerbstätige ermittelt. Eine fand Sensitivitätsanalyse nicht statt.

Die Daten bei Breyer et al. (1997) wurden retrospektiv (ohne Diskontierung) bis zu fünf Jahren nach der Therapie ermittelt und auf ein Jahr heruntergerechnet. Dieses Vorgehen birgt gegenüber dem in der vorliegenden Arbeit verwendeten Verfahren der letzten zwei Jahre die Gefahr, dass für einen so langen Zeitraum starke Erinnerungsfehler auftreten können. Zudem waren dort die Selektionseffekte aufgrund der freiwilligen retrospektiven Befragung nicht kontrolliert, wohingegen bei unserem prospektiven Design lediglich die Gefahr der Verzerrung durch Dropouts bei den Nachuntersuchungen bestand (und als gering erachtet wurde). Die genannten Unterschiede weisen darauf hin, dass das bei Breyer et al. (1997) berechnete Ergebnis nicht mit dem in dieser Arbeit ermittelten verglichen werden kann; die Untersuchung war in mehrfacher Hinsicht als methodisch nicht so günstig, wie die vorliegende Studie einzuschätzen.

In der Untersuchung von Zielke (1993) zur stationären Verhaltenstherapie bei psychosomatischen Patienten wurde nicht mit Durchschnittswerten, sondern mit konkreten, individuell verursachten Kosten der Patienten (für zwei Jahre nach Therapie) gerechnet. Die entsprechenden Daten erhielt der Autor von kooperierenden Krankenkassen. Man könnte nun also einen Vergleich mit dessen Ergebnis als Vergleich mit einem „wahren“ Wert betrachten und daran die Güte der eigenen Untersuchung messen. Leider geht auch dies an dieser Stelle nicht. In der genannten Studie wurden nämlich sämtliche ambulanten und stationären Kosten der Patienten im untersuchten Zeitraum erfasst (auch die, die nicht wegen der psychischen Erkrankung verursacht wurden). Somit ergaben sich viel höhere Werte, auf deren Basis kein Vergleich möglich ist; außerdem handelte es sich um ein breiteres Spektrum psychosomatischer Störungen, deren Inanspruchnahmeverhalten wahrscheinlich höher liegt als das von Angstpatienten.

Bei der Ermittlung der Arbeitsunfähigkeitskosten wurde bei Zielke (1993) sehr genau vorgegangen. Auch dort wurden nur die Tage der tatsächlich erwerbstätigen Patienten gewertet, und zwar differenziert in Produktivitätsausfall und Lohnfortzahlung (ab dem 43. Krankenschreibungstag Krankengeld) sowie dem Steuerausfall. Damit wurde die volkswirtschaftliche Perspektive noch

stärker betont, als das in der vorliegenden Arbeit gelungen sein mag. Als einen hier nicht einbezogenen Kostenfaktor war in der genannten Studie zusätzlich der Medikamentenkonsum enthalten (der allerdings regelmäßig in Kosten-Nutzen-Analysen nicht besonderes besonders bedeutsam ist; vgl. "Marginalanalyse"). Eine Sensitivitätsanalyse und eine Diskontierung der Ergebnisse fehlte allerdings auch bei dieser Studie.

Für zukünftige Studien bleibt also zu wünschen, dass einerseits einer nötigen Standardisierung und der Orientierung an gesundheitsökonomischen Richtlinien Rechnung getragen wird. Zudem sollte wenn möglich wie bei Zielke (1993) die Datenqualität durch Kooperation mit Kostenträger maximiert werden.

6 Allgemeine Diskussion zu Kosten-Studien und Ausblick

In der vorliegenden Arbeit wurde gezeigt, dass klinisch-psychologische Behandlungen unter ambulanten Praxisbedingungen nicht nur effektiv sind, sondern auch kostengünstig: die Krankheitskosten reduzieren sich bereits im ersten Jahr nach der Therapie deutlich. Für den stationären Bereich liegt bereits ein entsprechender Nachweis vor (Zielke, 1993). Dies hat Implikationen für den Stellenwert klinisch-psychologischer Behandlungsmethoden in der Gesundheitsversorgung, denn Wirksamkeit und Wirtschaftlichkeit sind die zentralen Forderungen an therapeutische Maßnahmen schlechthin. Entsprechende Nachweise können also dazu beitragen (ebenso wie adäquate Qualitätssicherung), dem gerade seit Bestehen des Psychotherapeutengesetzes gestiegenen Legitimationsdruck gegenüber der Psychotherapie wirksam zu begegnen und im Wettbewerb mit konkurrierenden Anbietern gut dazustehen (vgl. Jacobi et al., im Druck).

Damit die angestrebte Funktion solcher Analysen erfüllt werden kann, nämlich eine Entscheidungsgrundlage für Kostenträger und Patienten zu sein, müssen vergleichbare Studien zu anderen Therapiealternativen und weiteren Störungsbildern folgen. Es ist also nötig, die Forschung auf diesem Gebiet und die Bereitstellung von Methoden für derartige Analysen voranzutreiben.

Diese Untersuchung zu Kosten und Nutzen störungsspezifischer kognitiv-behavioraler Angstbehandlung und deren Ergebnis wird insgesamt als potenter Ausgangspunkt für weitere Kostenstudien angesehen, so dass Vergleiche möglich werden. Hierfür wurde besonders auf Transparenz und Orientierung an gesundheitsökonomischen Standards geachtet, um die bislang mangelnde Vergleichbarkeit zu erhöhen. Die Kosten-Nutzen-Analyse könnte noch verbessert werden durch objektive Daten, wie sie Zielke (1993) in Zusammenarbeit mit Krankenkassen verwendet hat. Weiterhin wäre eine Miteinbeziehung weiterer volkswirtschaftlicher Kostenfaktoren für eine noch genauere Betrachtung wünschenswert.

Außerdem sollten in Kosten-Nutzen-Studien ermittelte Ergebnisse volkswirtschaftlich „hochgerechnet“ werden („Unbehandelte Angststörungen verursachen in Deutschland jährlich Kosten von...“) – dies ist aber eine Wissenschaft für sich und kann in der vorliegenden Arbeit nicht geleistet werden.

Darüber hinaus sind mehrere weitere Aspekte von Kosten-Studien für die Klinische Psychologie relevant.

So könnten Kosten-Aspekte z.B. dazu beitragen, das Konstrukt der „Klinischen Signifikanz“ (vgl. Abschnitt 2.2) voranzutreiben: Kostenreduktion könnte (ebenso wie z.B. Zielerreichung) als Außenkriterium für die anhand

von Fragebogenwerten ermittelten klinisch signifikanten Veränderungen dienen (Kosfelder, Jacobi & Grosse Holtforth, in Vorb.).

Ein weiterer Aspekt betrifft die Ethik bei Kosten-Nutzen-Analysen in der Psychotherapieforschung. Das Argument, Kosten-Nutzen-Analysen von Psychotherapie als unangebracht oder sogar unethisch abzutun, ist bekannt. Es wird damit begründet, dass eine Entscheidung über Psychotherapien, die durchgeführt werden, um kranken bzw. leidenden Menschen zu mehr Lebensqualität zu verhelfen, nicht aufgrund von finanziellen Überlegungen gefällt werden sollte. Dieses Argument ist insofern verkürzt, als die Kosten-Analyse ja nicht Entscheidung über Interventionen *ist*, sondern lediglich eines von vielen Instrumenten der *Entscheidungsfindung* sein sollte. Im Zeitalter begrenzter finanzieller Ressourcen im Gesundheitswesen muss der finanzielle Aspekt berücksichtigt werden, da gemäß dem Opportunitätskostenprinzip stets im Auge behalten werden muss, dass an einen Ort bereitgestellte Ressourcen an einem anderen Ort fehlen werden.

Ethische Überlegungen kann man auch zu den Vorgehens- und Berechnungswegen in der Kosten-Nutzen-Analyse anstellen, z.B. bei der Berechnung der Arbeitsunfähigkeit. Es ist ethisch fragwürdig, zwischen Erwerbstätigen, Hausfrauen etc. zu differenzieren. Eine Studie, die zu der Empfehlung kommt, hoch und niedrig verdienenden bzw. Erwerbslosen und Geldverdienern unterschiedliche Gesundheitsleistungen bereitzustellen, ist durchaus ethisch anfechtbar. Gegen eine solche Aufsplittung nach etwa Geschlecht oder Erwerbstätigkeit spricht neben einem Gleichheitsgrundsatz auch ein Argument von Kruppe & Hundhausen (1984). Hier werden die Kosten für die Krankheit von Hausfrauen oder ehrenamtlich Tätigen auf die gleiche Weise wie die der Erwerbstätigen berechnet, mit dem Argument, dass beispielsweise in Familien Hausfrauen und -männer die Erwerbstätigkeit ihres Partners in diesem Umfang ermöglichen. So sind sie, obwohl nicht erwerbstätig an der volkswirtschaftlichen Produktivität beteiligt.

Eine weitere ethische Überlegung betrifft unterschiedlichen Kosten-Nutzen-Indices, die sich bei unterschiedlichen Störungen ergeben: So sollten die Ergebnisse der Kosten-Nutzen-Analyse nicht dahingehend umgesetzt werden, dass nur Störungen, deren Behandlung Einsparungen andernorts verspricht, als behandlungswürdig betrachtet werden. Bei sexuellen Funktionsstörungen beispielsweise ist ein derartiges Kosten-Nutzen-Verhältnis nicht zu erwarten, da sexuelle Störungen in den seltensten Fällen zu Arbeitsausfall führen. Auch stationäre Aufenthalte und ambulante Behandlungen werden nicht so häufig wahrgenommen wie bei Angststörungen. Bei einer Heilung einer sexuellen Störung wäre daher nicht mit so großen volkswirtschaftlichen Einsparungen zu rechnen. Es ist ethisch problematisch, wenn eine Störung, die mit einer Beeinträchtigung der Lebensqualität einhergeht, aber nicht ein so gutes Kosten-Nutzen-Verhältnis aufzeigen, für die Krankenkassen nicht als finanzierenswert gelten würde.

Im Zusammenhang mit der Frage der Bedeutung von Kosten-Studien für die Klinische Psychologie soll nun noch auf einige weitere Aspekte der psychosozialen Versorgung hingewiesen werden, vor deren Hintergrund die vorliegende Kosten-Studien eingeordnet werden müssen.

Die Indikationsfrage wurde bereits im Zusammenhang mit der Optimierung der Kosten-Effektivität angesprochen. Es erscheint auf den ersten Blick ökonomisch sinnvoll, Patienten einer Behandlung wie der in dieser Studie untersuchten zuzuweisen, denn die Kosten scheinen im Vergleich zu anderen Gesundheitsleistungen überschaubar zu sein (bei gleichzeitiger hoher Effektivität der Behandlung). Die Möglichkeiten einer optimal kosten-effektiven Zuweisung sind aber allein aufgrund der Rahmenbedingungen psychologischer Psychotherapie begrenzt, da empirisch fundierte effektive Behandlungen, deren Wirksamkeit auch bei kurzer Therapiedauer nachgewiesen wurde, häufig gar nicht den Weg in die klinische Praxis finden. So fanden etwa Margraf & Poldrack (2000) in einer epidemiologischen Untersuchung, dass nur ein verschwindend geringer Bruchteil von Angstpatienten trotz jahrelanger Krankheitsdauer Verhaltenstherapie erhalten hatte, obwohl diese als Methode der Wahl gilt. Ähnliches fand Wittchen (2000) in Bezug auf die Versorgungslage Jugendlicher und junger Erwachsener. Die Verbreitung (*dissemination*) von Forschungsergebnissen und die praktikergerechte Aufbereitung dürfen als eigenständiger Aspekt der (Weiter-) Entwicklung einer Behandlungsform nicht vernachlässigt werden (vgl. Kosfelder et al. 2000; Jacobi & Poldrack, 2000).

Auch die fortlaufende Qualitätssicherung einmal implementierter empirisch fundierter Behandlungen muss gewährleistet sein (z.B. über Zertifizierung und kontrollierter Lehre der Behandlungsform, Supervision der Behandlungsintegrität, Erfolgskontrolle).

Deshalb wird hier dafür plädiert, einen Rahmen für eine kosten-effektive psychotherapeutische Versorgung zu schaffen, indem erstens die Mittel für weitere Forschung zum Transport empirisch bereits gut abgesicherter Verfahren in den klinischen Alltag bereitgestellt werden. Zweitens müssten Anreize dafür geschaffen werden, dass alle Behandler die Kosten-Effektivität ihrer Therapien berücksichtigen und *state of the art* arbeiten können (vgl. Jacobi et al., in Druck). Denn wer derzeit Therapien frühzeitig beendet oder abbricht, außerordentliche Verlängerungsanträge aushandelt, Therapien flexibel gestaltet und dabei wenn indiziert auch seine Praxisräume oder seine Station verlässt, Supervision und/oder Intervision pflegt und sich als *scientist-practitioner* regelmäßig auf dem laufenden hält, verhält sich ökonomisch ungeschickt oder ist außerordentlich idealistisch. Oder er/sie ist Weiterbildungskandidat.

Auch zur Frage der Rationalität von Entscheidungen im Gesundheitssektor soll hier Stellung bezogen werden. Es fällt schwer, Kosten-Nutzen-Aussagen wie die zur untersuchten Psychotherapie dieser Studie in größerem Rahmen zu interpretieren. Denn aus gesundheitsökonomischer Perspektive lässt sich

durchaus nicht behaupten, dass „einsparende“ Effekte wie die Reduktion von Krankheitskosten und Arbeitsunfähigkeit durch Psychotherapie einen wirklichen finanziellen Nutzen erbringen. Töpfe und Budgetierungen sind verteilt und werden in jedem Fall aufgebraucht („Nullsummen-Spiel“).

Allokationsentscheidungen in unserem Gesundheitssystem, einem „System ohne Steuerung“ (SPIEGEL-Titel vom 1.11.99), sind kaum durchschaubar, weil darin so vieles interagiert: betriebswirtschaftliche Interessen der Krankenkassen, gesetzlicher und volkswirtschaftlicher Auftrag, parteipolitische Gesundheitspolitik, ausgeprägtester Lobbyismus verschiedener Erbringer von Gesundheitsleistungen und Hersteller von Heilprodukten, enormer Fortschritt und kurze Halbwertszeit des medizinischen Wissensbestandes, Medienpräsenz bestimmter Krankheiten und Behandlungen, Einzelschicksale vs. Allgemeinwohl, „Nulltarif-Illusion“ und Patienten-Interessen, mächtige Qualitätssicherungs-Apparate, Wissenschaftliche Beiräte etc. – allein diese Vielzahl der Einflüsse macht es einem schwer zu glauben, dass Entscheidungen im Gesundheitsbereich rational getroffen werden bzw. überhaupt geschaffen werden können.

Ein einfaches Beispiel zeigt in Kasten 5.2 wie eine aus der Perspektive der „Volksgesundheit“ sehr kosteneffektive Maßnahme für den Kostenträger nach hinten losgehen kann.

Kasten 5.2: Effizienz für wen?

Folgendes Beispiel zeigt, wie die Etablierung eines äußerst kostengünstigen Programms die Kosten derjenigen in die Höhe trieb, die sie anboten.

Eine Ortskrankenkasse sah nach der gesetzlichen Erleichterung des Kassenwechsels ihre Mitgliederzahl schwinden. Durch das kostenlose Angebot einer Präventionsmaßnahme gegen Rückenschmerzen sollten Mitglieder gehalten und neue geworben werden. Das implementierte Programm konnte kontrollierte Studien vorweisen, die belegten, dass die Teilnahme die zukünftige Inanspruchnahme von Gesundheitsleistungen senkt.

Was zunächst, bis die Maßnahme ganz schnell wieder aus dem Angebot gestrichen wurde, nicht bedacht worden war, war die Perspektive der Einsparung. Die Kosten reduzierten sich natürlich nur relativ pro Teilnehmer. Die Gruppe derer, die solche Angebote in Anspruch nimmt, hat aber insgesamt im Schnitt überdurchschnittlich hohe Krankheitskosten aufzuweisen. Die neu geworbenen oder gehaltenen Mitglieder waren „schlechte Risiken“, weshalb sich die präventive Maßnahme als Kostentreiber herausstellte. Ironisch gesprochen: lieber eine Snowboard-Rallye anbieten, auch wenn sich jeder zehnte Teilnehmer dabei die Beine bricht – solange dabei „gute Risiken“ erhalten bleiben.

Hinzu kommt schlicht die Größe der Summen, um die es im Gesundheitssektor geht. Ständig neue Angebote treiben die Kosten in die Höhe, die Leistungsausgaben gesetzlicher Krankenversicherungen verdoppelten sich in den letzten 10 Jahren von 130 auf über 250 Milliarden DM. Bei anhaltendem

Boom entwickelte sich der Gesundheitssektor – bei weitgehender Abwesenheit echter Marktmechanismen – zu einem Wirtschaftsfaktor ersten Ranges.

Allein die Vielzahl und die Stärke der Interessen und Motive schaffen eine Komplexität, bei der man sich fragt, ob es sich überhaupt lohnt, ohne administrativen Auftrag als Wissenschaftler Kostenstudien anzustellen. Wichtiger denn je scheint in diesem Zusammenhang auch die effiziente Verbreitung und politische Verwertung von guten Ergebnissen zu sein.

Von einigen Krankenkassen wurde von Patienten unserer Stichprobe geradezu gefordert, erst eine gewisse Anzahl an anderen Behandlungen (z.B. tiefenpsychologisch fundiert, stationär) durchzuführen, bevor die störungsspezifische Therapie unserer Behandlungseinrichtung finanziell erstattet wird. Dabei wurde sich auf Richtlinien oder Gutachten berufen, die mittlerweile überholt erscheinen. Ein „Einkaufsmodell“, bei dem ein Netz an Behandlungseinrichtungen und Praxen unter Aspekten der Qualitätssicherung und der Kosten-Effektivität aufgebaut und flexibel finanziert wird, wäre möglicherweise eine Alternative. Transparenz der Kosten für Kostenträger und Patient, wie dies bei den privaten Krankenversicherungen gehandhabt wird, würde eine vernünftigeren und gerechteren Ressourcenverteilung begünstigen (übrigens auch im stationären Bereich).

Im Psychotherapiebereich sind zudem noch mehr als in den meisten anderen Fachgruppen bessere Rahmenbedingungen vonnöten, um effektiv und auch kostengünstig zu behandeln. Die Psychotherapie hat dafür schlechte Startbedingungen. Abgesehen davon, dass sie machtpolitisch kaum Gewicht hat, haben Gesetzeslage und nachgeordnete Psychotherapierichtlinien auch nach Inkrafttreten des seitdem oft gescholtenen Psychotherapeutengesetzes ihre Tücken. Eher als andere (ärztliche) Leistungen soll die Psychotherapie (auf eigene Kosten) Wirksamkeitsnachweise und fortlaufende Qualitätssicherung bereitstellen. Man verlangt von ihr, dass sie selbstverständlich endgültig heilt, und Rückfälle und Mehrfachbehandlung werden als Beweis ihrer Wirkungslosigkeit gewertet. Es ist eine gesundheitspolitisch wichtige Aufgabe für die Klinische Psychologie, derartige Fehlannahmen zurechtzurücken.

Es wäre lohnend, eine Recherche zu einer geeigneten Auswahl an nicht akut lebenserhaltender medizinischer Regelleistungen anzustellen, die innerhalb eines Jahres ebensoviel oder mehr kosten wie die in dieser Arbeit untersuchte Psychotherapie. Der Autor wettet, dass dabei viele diagnostische Maßnahmen und Therapien gefunden werden, die eine bedeutend niedrigere Kosten-Effektivität aufweisen und weniger direkten und indirekten Nutzen erzielen als die Psychotherapie, aber deren Status als selbstverständliche Regelleistung kein Sozialdemokrat, kein Christdemokrat und kein Grüner etc. je in Frage stellen würde – zumindest so lange das System noch hält.

Abschließend noch eine Anmerkung, die auch eine persönliche Seite hat. Im Laufe dieser Arbeit stand im Kollegenkreis mehrfach die Frage im Raum: „Kosten-Nutzen-Analyse – ist das denn wirklich noch Psychologie?“ Dem steht die Überzeugung gegenüber, dass unter gegebenen Bedingungen die Psychologie zunehmend auch ihren gesellschaftlichen Nutzen herausstellen muss. Dass dies mit klinisch-psychologischer Forschung durchaus gut vereinbar ist, entspricht nicht nur der Überzeugung des Autors; er hofft, mit der vorliegenden Arbeit auch einen Beitrag dazu geleistet zu haben.

Literatur

- Andrews, G., Sanderson, K. & Beard, J. (1998). Burden of disease: Methods of calculating disability from mental disorder. *British Journal of Psychiatry*, 173, 123-131.
- Andrews G. & Henderson S. (Eds.). *Unmet need in psychiatry*. Cambridge: Cambridge University Press.
- Arrow, K., Solow, R., Portney, P. R., Leamer, E.E., Radner, R. & Schuman, H. (1993). Report of the NOAA Panel of Contingent Valuation. *Federal Register*, 58, 4601-4614.
- Baltensperger, C. (1996). *Psychotherapie: Kostspieliger Luxus oder gesellschaftlicher Nutzen?*. Unveröff. Diss., Universität Bern.
- Baltensperger, C. & Grawe, K. (2001). Psychotherapie unter gesundheitsökonomischem Aspekt. *Zeitschrift für Klinische Psychologie*, 30 (1), 10-21.
- Barlow, D. H. (1988). *Anxiety and its disorders*. New York: Guilford.
- Barlow, D. H., Craske, M. G., Cerny, J. A. & Klosko, J. S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy*, 20, 261-282.
- Baumann, U. & Perrez, M. (1998). *Lehrbuch Klinische Psychologie – Psychotherapie* (2. vollständig überarbeitete Aufl.). Bern: Huber.
- Bellach, B. M., H. Knopf & W. Thefeld (1998). The German Health Survey. 1997/98. *Das Gesundheitswesen*, 60 (2), 59-68.
- Binder, J. & Angst, J. (1981). Soziale Konsequenzen psychischer Störungen in der Bevölkerung. Eine Feldstudie an jungen Erwachsenen. *Archiv für Psychiatrie und Nervenkrankheiten*, 229 (4), 355-370.
- Blue, I. & Harpham, T. (1994). The World Bank World Development Report 1993: Investing in Health: Reveals the burden of common mental disorders, but ignores its implications. *British Journal of Psychiatry*, 165 (1), 9-12.
- Bourdon, K. H., Boyd, J. H., Rae, D. S., Burns, B. J., Thompson, J. W. & Locke, B. Z. (1988). Gender differences in phobias - results of the ECA community survey. *Journal of Anxiety Disorders*, 2, 227-241
- Breyer, F., Heinzl, R. & Klein, T. (1997). Kosten und Nutzen ambulanter Psychoanalyse in Deutschland. *Gesundheitsökonomie und Qualitätsmanagement*, 2, 59-73.
- Bühringer, G., Hahlweg, K. (1986). Kosten-Nutzen Aspekte psychologischer Behandlung. *Psychologische Rundschau*, 37, 1-19.
- Butollo, W. & Höfling, S. (1984). *Behandlung chronischer Ängste und Phobien*. Stuttgart: Enke-Verlag.
- Candilis, P. J. & Pollack, M. H. (1997). The hidden costs of untreated anxiety disorders. *Harvard Review of Psychiatry*, 5, 40-42.

- Christensen, L. & Mendoza, J. L. (1986). A method of assessing change in a single subject: An Alteration of the RC Index. *Behavior Therapy*, 17, 305-308.
- Cienskowsky, U. (2000). Meta-Analyse. In F. Jacobi & A. Poldrack (Hrsg.), *Klinisch-Psychologische Forschung. Ein Praxishandbuch* (S. 109-120). Göttingen: Hogrefe.
- Clark, D. M. (1994). Cognitive Therapy for panic disorder. In: B. E. Wolfe & J. D. Maser (Eds.), *Treatment of panic disorder: A consensus development conference* (pp. 121-132). Washington: American Psychiatric Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.
- Cook, T. D. & Campbell, D. T. (1979). *Quasi-Experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin Company.
- Crits-Christoph, P. (1992). The efficacy of brief dynamic psychotherapy: A meta-analysis. *American Journal of Psychiatry*, 149, 151-158.
- Cummings, N. A. (1977). The anatomy of psychotherapy under national health insurance. *American Psychologist*, 32, 711-718.
- Davidson, J. R. T. (1996). Quality of life and cost factors in panic disorders. *Bulletin of the Menninger Clinic*, 60 (2) (suppl. A), A5-A11.
- Dilcher, K. (2000). Forschung im stationären Setting. In F. Jacobi & A. Poldrack (Hrsg.), *Klinisch-Psychologische Forschung. Ein Praxishandbuch* (S. 89-108). Göttingen: Hogrefe.
- Donaldson, C., Shackley, P. & Abdalla, M. (1997). Using willingness-to-pay to value close Substitues. *Economic Evaluation*, 6, 145-159.
- Drummond, M. F., Stoddart, G. L. & Torrance, G. W. (1987). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press.
- DuPont, R. L., Rice, D. P., Miller, L. S., Shiraki, S. S., Rowland, C. R. & Harwood, H. J. (1996). Economic costs of anxiety disorders. *Anxiety*, 2, 167-172.
- Edlund, M. J. & Swann, A. C. (1987). The economic and social costs of panic disorder. *Hospital and Community Psychiatry*, 38, 1277-1279.
- Ehlers, A. & Clark, D. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, 38 (4), 319-345.
- Ehlers, A., Margraf, J. & Chambless, D. L. (1993). *Fragebogen zu körperbezogenen Ängsten, Kognitionen und Vermeidung (AKV) mit den Skalen BSQ, ACQ und MI*. Weinheim: Beltz.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Consulting Psychology*, 16, 319-324.
- Eysenck, H. J. (1959). Learning theory and behaviour therapy. *Journal of Mental Science*, 195, 61-75.
- Eysenck, H. J. (1993). Grawe and the effectiveness of psychotherapy: Some comments. *Psychologische Rundschau*, 44 (3), 177-180.

- Fellmann, J. (1996). *Ökonomische Evaluationsmethoden im Gesundheitswesen. Theorie und Anwendung der Kosten-Nutzen-Indizes*. Unveröff. Diplomarbeit. Technische Universität Dresden.
- Frank, M. & Fiegenbaum, W. (1993). Therapieerfolgsmessung in der psychotherapeutischen Praxis. *Zeitschrift für Klinische Psychologie*, 23 (4), 268-275.
- Franke, G. (1995). *Die Symptom-Checkliste von Derogatis - Deutsche Version*. Göttingen: Hogrefe.
- Fiedler, P. (1997). Therapieplanung in der modernen Verhaltenstherapie: Von der allgemeinen zur phänomen- und störungsspezifischen Behandlung. *Verhaltenstherapie und Verhaltensmedizin*, 18, 7-39.
- Fiegenbaum, W., Tuschen, B. & Florin, I. (1997). Qualitätssicherung in der Psychotherapie. *Zeitschrift für Klinische Psychologie. Forschung und Praxis*, 26 (2), 138-149.
- Follette, W. T. & Cummings, N. A. (1967). Psychiatric services and medical utilization in a prepaid health plan setting. *Medical Care*, 5, 25-35.
- Galler, H.-P. (1990). The Willingness to Pay Approach: Caveats to biased Application. In Laaser, U., Roccella, E. J., Rosenfeld, J. B. & Wenzel H. (Eds.), *Costs and benefits in Health Care and Prevention* (pp. 43-61). Berlin: Springer-Verlag.
- Gathmann, P. (1990). *Pathologie des psychosomatischen Reaktionsmusters: Diagnose, Klinik, Therapie*. Berlin: Springer.
- Gebel, U. (1998). *Kosten-Effektivitäts-Analyse von ambulanter Psychotherapie*. Unveröff. Diplomarbeit, Technische Universität Dresden.
- Ginsberg, G. & Marks, I. (1977). Costs and benefits of behavioural psychotherapy: A pilot study of neurotics treated by nursetherapists. *Psychological Medicine*, 7, 685-700.
- Goldberg, I. D., Krantz, G. & Locke, B. Z. (1970). Effects of a short-term outpatient psychiatric therapy benefit on the utilization of medical services in a prepaid group practice medical program. *Medical Care*, 8, 419-428.
- Goldfried, M. R. & Wolfe, B. E. (1998). Toward a more clinically valid approach to therapy research. *Journal of Consulting and Clinical Psychology*, 66 (1), 143-150.
- Grawe, K. (1992). Psychotherapieforschung zu Beginn der neunziger Jahre. *Psychologische Rundschau*, 43, 189-196.
- Grawe, K. (1998). *Psychologische Therapie*. Göttingen: Hogrefe.
- Grawe, K. (1999). *Psychotherapie nach dem Wandel: Was können wir von einer professionellen Psychotherapie realistisch erwarten?* Plenarvortrag gehalten auf dem 1. Workshopkongress für Klinische Psychologie und Psychotherapie, Bad Dürkheim, 13.-15. Mai 1999.
- Grawe, K., Donati, R. & Bernauer, F. (1994). *Psychotherapie im Wandel - von der Konfession zur Profession*. Göttingen: Hogrefe.

- Grosse Holtforth, M. & Grawe, K. (2000). Fragebogen zur Analyse Motivationaler Schemata (FAMOS). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 29 (3), 170-179.
- Grosse Holtforth, M. & Regli, D. (2000). Möglichkeiten empirischer Psychotherapieforschung in einer psychotherapeutischen Ambulanz. In: F. Jacobi & A. Poldrack (Hrsg.), *Klinisch-Psychologische Forschung. Ein Praxis-Handbuch* (S.71-88). Göttingen: Hogrefe.
- Hahlweg, K. (1995). Zur Förderung und Verbreitung psychologischer Verfahren. Ein APA-Bericht. Editorial. *Zeitschrift für Klinische Psychologie*, 24, 275-284.
- Hartmann, A. & Herzog, T. (1995). Varianten der Effektstärkenberechnung in Meta-Analysen: Kommt es zu variablen Ergebnissen?. *Zeitschrift für Klinische Psychologie*, 24 (4), 337-343.
- Hartmann, A., Herzog, T. & Drinkmann, A. (1992). Psychotherapy of bulimia nervosa: What is effective? A meta-analysis. *Journal of Psychosomatic Research*, 36 (2), 159-167.
- Hautzinger, M., Bailer, M., Worall, H. & Keller, F. (1993). *Das Beck Depressionsinventar*. Bern: Huber.
- Havlicek, L. L. & Peterson, N. L. (1974). Robustness of the t-test: A guide for researchers on effect of violations of assumptions. *Psychological Reports*, 34 (3), 1095-1114.
- Heuft, G., Senf, W., Wagener, R. & Pintelon, C. (1996). Individuelle Therapieziele: Zur Ergebnisdokumentation stationärer Psychotherapie aus Patienten- und Therapeutesicht. *Zeitschrift für Klinische Psychologie, Psychopathologie und Psychotherapie*, 44 (2), 186-199.
- Höfling, S. (1988). Wirksamkeit und Nutzen psychotherapeutischer Behandlung. In Berufsverband Deutscher Psychologen e.V. (Hrsg.). *Schriftenreihe Heft 8*. Bonn: Deutscher Psychologen Verlag.
- Hofmann, S. & Barlow, D. H. (2000). The Costs of Anxiety Disorders: Implications for Psychosocial Interventions. In N. E. Miller & K. M. Magruder (Eds.), *Psychotherapy: Costs and Benefits; A Guide for Practitioners, Researchers and Policymakers* (pp. 224-234). New York: John Wiley & Sons.
- Hollon, S. D. & Flick, A. N. (1988). On the meaning and methods of clinical significance. *Behavioral Assessment*, 10, 197-206.
- Howard K. I., Orlinsky, D. & Lueger, R. (1994). Clinically relevant outcome research in individual psychotherapy. *British Journal of Psychiatry*, 165, 4-8.
- Jablensky, A. & Hugler, H. (1982). Möglichkeiten und Grenzen psychiatrischer und epidemiologischer Surveys für geographisch definierte Populationen in Europa. *Fortschritte für Neurologie, Psychiatrie*, 50, 215-239.
- Jacobi, F. (2001). Misserfolgsvorschung in der Verhaltenstherapie. In R. Dohrenbusch & F. Kaspers (Hrsg.), *Fortschritte der Klinischen Psychologie und Verhaltensmedizin* (S. 323-346). Lengerich: Pabst.

- Jacobi, F., Jahn, B. & Wittchen, H.-U. (2001). *Häufigkeit und Behandlungsstatus psychischer Störungen in Deutschland*. Posterbeitrag auf dem 2. Workshopkongress für Klinische Psychologie und Psychotherapie, 24.-26.Mai 2001 in Bern.
- Jacobi, F. & Kosfelder, J. (2000). Psychotherapieforschung. In F. Jacobi & A. Poldrack (Hrsg.), *Klinisch-Psychologische Forschung. Ein Praxishandbuch* (S. 245-266). Göttingen: Hogrefe.
- Jacobi, F. & Poldrack, A. (Hrsg.) (2000). *Klinisch Psychologische Forschung. Ein Praxishandbuch*. Göttingen: Hogrefe.
- Jacobi, F., Poldrack, A. & Hoyer, J. (2001). Qualitätssicherung in der ambulanten Verhaltenstherapie – Probleme, Strategien und Optionen. In J. Sturm & R.T. Vogel (Hrsg.), *Neue Entwicklungen in der Psychotherapie: Aufbruch in die Vernetzung* (S.161-185). Lengerich: Papst.
- Jacobi, F. & Schneider, S. (2000). Die Angst vor der Angst. Angststörungen: Ursachen und Behandlungsmöglichkeiten. In S. Barnow, H. J. Freyberger, W. Fischer & M. Linden (Hrsg.), *Von Angst bis Zwang. Ein ABC der psychischen Störungen* (S. 97-122). Bern: Huber.
- Jacobi, F. & Margraf, J. (in Druck). Kosten-Studien zu psychologischer Angstbehandlung. In W. Michaelis (Hrsg.), *Ausgabenbegrenzung im Gesundheitswesen*. Landsberg: ecomed.
- Jacobson, N. S., Follette, W. C. & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336-352.
- 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.
- Jacobson, N. S., Wilson, L. & Tupper, C. (1988). The clinical significance of treatment gains resulting from exposure-based interventions for agoraphobia: A reanalysis of outcome data. *Behavior Therapy*, 19, 539-552.
- Jenkins, R. (1997). Nations for mental health. *Social Psychiatry and Psychiatric Epidemiology*, 32 (6), 309-311.
- Johannesson, M., Jörnsson, B. & Borquist, L. (1991). Willingness to pay for antihypertensive therapy – results of a swedish pilot study. *Journal of Health Economics*, 10, 461-474.
- Kazdin, A. E. & Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology*, 57, 138-147.
- Kendall, P. C. (1998). Empirically supported psychological therapies. *Journal of Consulting and Clinical Psychology*. 66 (1), 3-6.
- Kessler, R. C., McGonagle, K. A., Zhao, A., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H.-U. & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. *Archives of General Psychiatry*, 51, 8-19.

- Kiesler, D. J. (1966). Some myths of psychotherapy research and the search for a paradigm. *Psychological Bulletin*, 1966, 65, 110-136.
- Klerman, G. L., Weissman, M. M., Oullette, R., Johnson, J. & Greenwald, S. (1991). Panic attacks in the community. Social morbidity and health care utilization. *Journal of the American Medical Association*, 265, 742-746.
- Koester, W., Schneider, R., Hachmann, E. & Mai, N. (1982). Ergebnisse des stationären verhaltenstherapeutischen Programms zur Behandlung von Alkohol- und Medikamentenabhängigen: Katamnesen nach einem Jahr. In R. Schneider (Hrsg.), *Stationäre Behandlung von Alkoholabhängigen* (S.215-243). München: Röttger
- Konsensgruppe „Gesundheitsökonomie,, (1996). Empfehlungen zur gesundheitsökonomischen Evaluation. *Zeitschrift für Allgemeine Medizin*, 72, 485-490.
- Kosfelder, J., Jacobi, F. & Grosse Holtforth, M. (in Vorb.). Wie valide ist das Konstrukt „Klinische Signifikanz“?; vgl. Jacobi, F. & Kosfelder, J. (2001). *Überlegungen zum Konstrukt „Klinische Signifikanz“*. Vortrag im Forschungskolloquium Klinische Psychologie und Psychotherapie Technische Universität Dresden, 24.1.2001.
- Kosfelder, J., Poldrack, A. & Jacobi, F. (2000). Klinisch-Psychologische Forschung: Themen und Kunden. In F. Jacobi & A. Poldrack (Hrsg.), *Klinisch-Psychologische Forschung. Ein Praxishandbuch* (S. 9-25). Göttingen: Hogrefe.
- Krischke, N. & Petermann, F. (1997). Gesundheitsökonomische Parameter in der Medizinischen Rehabilitation. In F. Petermann (Hrsg.) *Rehabilitation* (S. 62-75). Göttingen: Hogrefe.
- Kruppe, R. & Hundhausen, G. (1984). *Volkswirtschaftliche Bewertung von Personenschäden im Straßenverkehr*. Bergisch Gladbach: Bundesanstalt für Straßenwesen.
- Lambert, M. J. & Hill, C. E. (1994). Assessing psychotherapy outcome and process. In A. E. Bergin & L. S. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 72-113). New York: Wiley.
- Lee, S. J., Neumann, P. J., Churchill, W. H., Cannon, M. E. & Weinstein, M. C. (1997). Patients willingness to pay for autologous blood donation. *Health Policy*, 40, 1-12.
- Leon, A. C., Portera, L. & Weissman, M. M. (1995). The social costs of anxiety disorders. *British Journal of Psychiatry*, 116 (Suppl. 27), 19-22.
- Linden, M. (1989). Die Phase IV in der Therapie-Evaluation. *Der Nervenarzt*, 60 (8), 453-461
- Linden, M., Förster, R., Oel, M. & Schlötelborg, R. (1993). Verhaltenstherapie in der kassenärztlichen Versorgung: Eine versorgungsepidemiologische Untersuchung *Verhaltenstherapie*, 3, 101-111.

- Linster, H.-W. & Wetzell, H. (Hrsg.). (1980). *Veränderung und Entwicklung der Person - Grenzen und Möglichkeiten psychologischer Therapie*. Hamburg: Hoffmann und Campe.
- Lipsey, M. W. & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *American Psychologist*, 48, 1181-1209.
- Longobardi, P. G. (1981). The impact of a brief psychological intervention on medical care utilization in an army health setting. *Medical Care*, 19, 665-671.
- Loomis, J. B. (1990). Comparative Reliability of the Dichotomous Choice and Open-Ended Contingent Valuation Techniques. *Journal of Environmental Economics and Management*, 18, 78-85.
- Luborsky, L., Singer, B. & Luborsky, L. (1975). Comparative studies of psychotherapy. *Archives of General Psychiatry*, 32, 995-1008.
- Maercker, A. (Hrsg.). (1997). *Therapie der Posttraumatischen Belastungsstörungen*. Berlin: Springer-Verlag.
- Magee, W. J., Eaton, W. W., Wittchen, H.-U., McGonagle, K. & Kessler, R. C. (1996). Agoraphobia, simple phobia, and social phobia in the National Comorbidity Survey. *Archives of General Psychiatry*, 53, 159-168.
- Margraf, J. (2000a). *Lehrbuch der Verhaltenstherapie* (Bd. 1 und 2.) (2.Aufl.). Berlin: Springer-Verlag.
- Margraf, J. (2000b). Grundprinzipien und historische Entwicklung. In J. Margraf (Hrsg.), *Lehrbuch der Verhaltenstherapie, Bd.1* (2.Aufl.) (S.1-30). Berlin: Springer-Verlag.
- Margraf, J., & Ehlers, A. (1996). *Beck Angstinventar (BAI)*. Bern: Huber.
- Margraf, J. & Lieb, R. (1994). Was ist Verhaltenstherapie? Versuch einer zukunftsorientierten Neucharakterisierung. *Zeitschrift für Klinische Psychologie*, 24, 1-7.
- Margraf, J. & Poldrack, A. (2000). Angstsyndrome in Ost- und Westdeutschland: Eine repräsentative Bevölkerungserhebung. *Zeitschrift für Klinische Psychologie*, 29 (3), 157-169.
- Margraf, J. & Schneider, S. (1990). *Panik. Angstanfälle und ihre Behandlung*. Berlin: Springer-Verlag.
- Margraf, J. & Schneider, S. (1996). *Kosten-Nutzen-Analyse psychologischer Psychotherapie: kognitive Verhaltenstherapie bei schweren Angstkrankheiten*. Forschungsbericht, Technische Universität Dresden.
- Margraf, J., Schneider, S. & Ehlers, A. (1991). *DIPS. Diagnostisches Interview bei psychischen Störungen (Handbuch, Interviewleitfaden und Protokollbogen)*. Berlin: Springer-Verlag.
- Markowitz, J. S., Weissman, M. M., Oullette, R., Lish, J. D. & Klerman, G. L. (1989). Quality of life in panic disorder. *Archives of General Psychiatry*, 46, 984-992.
- Marks, I. (1987). *Fears, phobias, and rituals*. New York: Oxford University.

- Matt, G. E. & Cook, T. D. (1994). Threats to the validity of research syntheses. In H. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis*. New York: Russell Sage Foundation.
- Mays, D. T. & Franks, C. M. (1985). *Negative Outcome in Psychotherapy*. New York: Springer Publishing.
- McFadden, D. (1994). Contingent Valuation and Social Choice. *American Journal of Agricultural Economics*, 76, 689-708.
- McGaw, B. & Glass, G. V. (1980). Choice of metric for effect size in meta-analysis. *American Educational Research Journal*, 17, 325-337.
- Merikangas, K. R., Angst, J., Eaton, W. W., Canino, G. J., Rubio-Stipec, M., Wacker, H., Wittchen, H.-U., Andrade, L., Essau, C. A., Kraemer, H., Robins, L. & Kupfer, D. J. (1996). Comorbidity and boundaries of affective disorders with anxiety disorders and substance misuse: results of an international task force. *British Journal of Psychiatry*, 168 (Suppl. 30), 58-67.
- Mitchell, R. C. & Carson, R.T. (1989). *Using surveys to value public goods*. Washington D.C.: Resources for the Future.
- Möser, K. (1998). *Verbesserte Methoden der Kosten-Nutzen Analyse bei Therapievergleichen: eine gesundheitsökonomische Evaluation am Beispiel schwerer Angstkrankheiten*. Unveröff. Diss., Technische Universität Dresden.
- Mumford, E., Schlesinger, H. J., Glass, G. V., Patrick, C. & Cuerdon, T. (1984). A new look at evidence about reduced cost of medical utilization following mental health treatment. *American Journal of Psychiatry*, 141(10), 1145-1158.
- Myers, J. K., Weissman, M. M., Tischler, G. L., Holzer, C. E., Leaf, P. J., Orvaschel, H., Anthony, J. C., Boyd, J. H., Burke, J. D., Kramer, M. & Stoltzman, R. (1984). Six month prevalence of psychiatric disorders in three communities: 1980-1982. *Archives of General Psychiatry*, 41, 959-967.
- Neumer, S. & Margraf, J. (1996). Kosten-Effektivitäts- und Kosten-Nutzen-Analyse. In J. Margraf (Hrsg.), *Lehrbuch der Verhaltenstherapie (Bd.1)* (S. 543-551). Berlin: Springer-Verlag.
- Nicholson, R. A. & Berman, J.S. (1983). Is follow-up necessary in evaluating psychotherapy? *Psychological Bulletin*, 93, 261-278.
- Noyes, R. & Glancy, J. (1976). Anxiety neurosis: A 5-year follow-up. *Journal of Nervous and Mental Disease*, 162, 200-205.
- Nuber, U. (1992). Wer darf die Seele kurieren? In M. Blum & T. Nessler (Hrsg.), *Psychische Umwelt – körperliche Gesundheit* (14-22). Freiburg: Rombach.
- O'Brien, B., Gafni, A. (1996). When do the „dollars” make sense? *Medical Decision Making*, 16, 288-299.
- O'Brien, B., Novosel, S., Torrance, G. & Streiner, D. (1995). Assessing the economic value of a new Antidepressant. A willingness – to – pay approach. *Pharmacoeconomics*, 8, 34-45.

- O'Brien, B. & ViraMontes, J. L. (1994). Willingness to pay: a valid and reliable measure of Health state preference? *Medical Decision Making*, 14, 289-297.
- Paul, G. L. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology*, 31, 109-118.
- Pieper, K. (1999). *Kosten-Nutzen-Analyse ambulanter Verhaltenstherapie – ein Längsschnitt*. Unveröff. Dipl.Arbeit, Technische Universität Dresden.
- Rachman, S. J. & Wilson, G. T. (1980). *The Effects of Psychological Therapy*. London: Pergamon Press.
- Reimer, C., Hempfing, L. & Dahme, B. (1979). Iatrogene Chronifizierung in der Vorbehandlung psychischer Erkrankungen. *Praxis Psychotherapie und Psychosomatik*, 24, 123-133.
- Reinecker, H. (1991). *Zwänge: Diagnose, Theorien und Behandlung*. Göttingen: Huber.
- Rief, W., Hiller, W., Geissner, E. & Fichter, M. M. (1994). Hypochondrie: Erfassung und erste klinische Ergebnisse. *Zeitschrift für Klinische Psychologie*, 23, 34-42.
- Ringel, E. & Kropiunigg, U. (1983). *Der fehlgeleitete Patient. Psychosomatische Patientenkarrieren und ihre Akteure*. Wien: Facultas.
- Robins, L. N., Helzer, J. E., Croughan, J. & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule. Its history, characteristics and validity. *Archives of General Psychiatry*, 38, 381-389.
- Robins, L. N., Helzer, J. E., Weissman, M. M., Orvaschel, H., Gruenberg, E., Burke, J. D. & Regier, D. A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry*, 41, 949-958.
- Robins, L. N. & Regier, D. A. (1991). *Psychiatric Disorders in America: The Epidemiologic Catchment Area Study*. New York: The Free Press.
- Robinson, J. R., Young, T. K., Roos, L. L. & Gelskey, D. E. (1997). Comparing administrative data and self-reports. *Medical Care*, 35 (9), 932-947.
- Rosen, J. C. & Wiens, A. N. (1979). Changes in medical problems and use of medical services following psychological intervention. *American Psychologist*, 34 (5), 420-431.
- Rosenthal, R. (1982). A simple, general purpose display of magnitude of experimental effect, *Journal of Educational Psychology*, 74, 166-169.
- Rosenthal, R. (1990). How are we doing in soft psychology? *American Psychologist*, 45, 775-777.
- Rosnow, R. L. & Rosenthal, R. (1988). Focused tests of significance and effect size estimation in counseling psychology. *Journal of Counseling Psychology*, 35, 203-208.
- Ruhmland, M. & Margraf, J. (2001). Effektivität psychologischer Therapien von Panik und Agoraphobie: Meta-Analysen auf Störungsebene. *Verhaltenstherapie, Special Issue, Vol. 11(1)*, 41-53.

- Runde, G. (1999). *Kosten-Nutzen-Analyse kognitiver Verhaltenstherapie: Perspektive der Volkswirtschaft und der Patienten*. Unveröff. Diplomarbeit, Technische Universität Dresden.
- Ryan, M. (1996). Using Willingness-to-Pay to assess benefits of reproductive techniques. *Economic Evaluation*, 5, 543-558.
- Salvador-Carulla, L., Segui, J., Fernandez-Cano, P. & Canet, J. (1995). Costs and offset effect in panic disorder. *British Journal of Psychiatry*, 116 (Suppl. 27), 23-28.
- Schlesinger, H. J., Mumford, E. & Glass, G. V. (1980). Mental health services and medical utilization. In G. R. Vandenbos (Eds.), *Psychotherapy. Practice, research, policy* (pp. 70-94). Beverly Hills: Sage.
- Schneider, S. & Margraf, J. (1998). *Agoraphobie und Panikstörung*. Göttingen: Hogrefe.
- Schulte, D. (1993). Wie soll Therapieerfolg gemessen werden? Überblicksarbeit. *Zeitschrift für Klinische Psychologie*, 22, 374-393.
- Sekot, C. (1999). *Messung des Erfolgs von Psychotherapie. Eine inhaltliche und methodische Entwicklung zusammengesetzter Therapieerfolgsmaße*. Unveröff. Dipl.Arbeit, Katholische Universität Eichstätt.
- Sekot, C., & Jacobi, F. (2001). Inhalte und Methoden einer evidenzbasierten Qualitätssicherung für Psychotherapie und klinische Psychologie. *Integrative Therapie*, 27(1-2), 29-62.
- Seligman, M. (1995). The Effectiveness of Psychotherapy - The Consumer Reports Study. *American Psychologist*, 50 (12), 965-974.
- Shadish, W. R., Matt, G. E., Navarro, A. M. & Siegle, G. (1997). Evidence that therapy works in clinically representative conditions. *Journal of Consulting and Clinical Psychology*, 65 (3). 355-365.
- Shapiro, S., Skinner, E., Kessler, L., van Korff, M. & German, P. (1984). Utilization of health and mental health services. *Archives of General Psychiatry*, 41, 971-978.
- Silverman, W. (1996). Cookbooks, manuals, and paint-by-numbers: Psychotherapy in the 90's. *Psychotherapy*, 33 (2), 207-215.
- Simon, G., Ormel, J., von Korff, M. & Barlow, W. (1995). Health care costs associated with depressive and anxiety disorders in primary care. *American Journal of Psychiatry*, 152, 352-357.
- Smith, M.L., Glass, G.V., Miller, T.I. (1980). Meta-analysis of psychotherapy. *American Psychologist*, 41, 165-180.
- Sommer, S. (1998). *Untersuchungen zur Behandlungsgeschichte von Angstpatienten*. Unveröff. Diplomarbeit, Technische Universität Dresden.
- Steyer, R., Hannoever, W., Telsler, C. & Kriebel, R. (1997). Zur Evaluation intraindividuelle Veränderung. *Zeitschrift für Klinische Psychologie*, 26 (4), 291-299.

- Strupp, H. H., Horowitz, L. M. & Lambert, M. J. (Eds.) (1997). *Measuring patient changes in mood, anxiety, and personality disorders: Toward a core battery*. Washington: American Psychological Association.
- Svartberg, M. & Stiles, T. C. (1991). Comparative Effects of short-term psychodynamic psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology, 59*, 704-714.
- Thompson, M. S. & Cohen, A. B. (1990). The Feasibility of the Willingness-to-Pay Measurement in Health Service Research. In U. Laaser, E. J. Roccella, J. B. Rosenfeld, & H. Wenzel (Eds.), *Costs and benefits in Health Care and Prevention* (pp. 164-177). Berlin: Springer-Verlag.
- Tschuschke, V. & Kaechele, H. (1996). What do psychotherapies achieve? A contribution to the debate centered around differential effects of different treatment concepts. In U. Esser, H. Pabst & G. W. Speierer (Eds.), *The power of the person-centered approach. New challenges, perspectives, answers* (S. 159-181) Köln: Gesellschaft für wissenschaftliche Gesprächspsychotherapie.
- Vessey, J. T. & Howard, K. I. (1993). Who seeks psychotherapy? *Psychotherapy, 30* (4), 546-553.
- Wade, W. A., Treat, T. A. & Stuart, G. L. (1998). Transporting an empirically supported treatment for panic disorder to a service clinic setting: A benchmarking strategy. *Journal of Consulting and Clinical Psychology, 66* (2), 231-239.
- Wampold, B. E. & Jensen, W. R., (1986). Clinical significance revisited. *Behavior Therapy, 17*, 302-305.
- Warner, K. E. & Luce, B. R. (1982). *Cost - Benefit and Cost - Effectiveness Analysis in Health Care*. Michigan: Health Administration Press University of Michigan.
- Weismann, H. W. (1997). Kostenanalysen sollen vergleichbarer werden. Kommentar. *Gesundheitsökonomie und Qualitätsmanagement, 2*, A63.
- Wittchen H-U. (1991). Der Langzeitverlauf unbehandelter Angststörungen: Wie häufig sind Spontanremissionen? *Verhaltenstherapie - Praxis, Forschung, Perspektiven; 1*(4), 273-282.
- Wittchen, H.-U. (1988). Natural course and spontaneous remissions of untreated anxiety disorders - Results of the Munich Follow-up Study (MFS). In I. Hand & H.-U. Wittchen (Eds.), *Panic and Phobias 2. Treatments and variables affecting course and outcome* (S. 3-17). Berlin: Springer-Verlag.
- Wittchen, H.-U. (2000). Met and unmet need for interventions in community cases with anxiety disorders. In G. Andrews & S. Henderson (Eds.), *Unmet need in psychiatry* (pp.256-276). Cambridge: Cambridge University Press.
- Wittchen, H.-U. (in Druck). Die Zukunft der Klinischen Psychologie - Zwischen Glanz und Erosion. In R. K. Silbereisen & D. Frey (Hrsg.), *Perspektiven der Psychologie*. Weinheim: Beltz.

- Wittchen, H.-U., Lieb, R., Schuster, P. & Oldehinkel, A. J. (1999). When is onset? Investigations into early developmental stages of anxiety and depressive disorders. In J. L. Rapoport (Eds.), *Childhood onset of "adult" psychopathology, clinical and research advances* (pp. 259-302). Washington: American Psychiatric Press.
- Wittchen, H.-U., Müller, N., Pfister, H., Winter, S. & Schmidtke, B. (1999). Affektive, somatoforme und Angststörungen in Deutschland - Erste Ergebnisse des bundesweiten Zusatzsurveys "Psychische Störungen". *Das Gesundheitswesen*, 61(Sonderheft 2), 216-222.
- Wittchen, H.-U., Müller, N. & Storz, S. (1998). Psychische Störungen: Häufigkeit, psychosoziale Beeinträchtigungen und Zusammenhänge mit körperlichen Erkrankungen. *Das Gesundheitswesen*, 2 (60), 95-100.
- Wittchen, H.-U. & Perkonig, A. (1996). Epidemiologie psychischer Störungen. Grundlagen, Häufigkeit, Risikofaktoren und Konsequenzen. In A. Ehlers & K. Hahlweg (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich D Praxisgebiete, Serie 2 Klinische Psychologie* (S. 69-144). Göttingen: Hogrefe Verlag für Psychologie.
- Wittchen, H.-U., Perkonig, A., Lachner, G. & Nelson, C. B. (1998). Early developmental stages of psychopathology study (EDSP) - Objectives and design. *European Addiction Research*, 4, 18-27.
- Wittchen H.-U. & Pfister, H. (Hrsg.). (1997). *DIA-X-Interviews: Manual für Screening-Verfahren und Interview; Interviewheft Längsschnittuntersuchung (DIA-X-Lifetime); Ergänzungsheft (DIA-X-Lifetime); Interviewheft Querschnittuntersuchung (DIA-X-12 Monate); Ergänzungsheft (DIA-X-12 Monate); PC-Programm zur Durchführung des Interviews (Längs- und Querschnittuntersuchung); Auswertungsprogramm*. Frankfurt: Swets, Zeitlinger.
- Wittchen, H.U. & Vossen, A. (2000). Komorbiditätsstrukturen bei Angststörungen – Häufigkeit und mögliche Implikationen. In J. Margraf (Hrsg.), *Lehrbuch der Verhaltenstherapie* (Bd.1) (2.Aufl.) (S.329-346). Berlin: Springer-Verlag.
- Wittchen, H.-U. & Zerssen, D. v. (1987). *Verläufe behandelter und unbehandelter Depressionen und Angststörungen: Eine klinisch-psychiatrische und epidemiologische Verlaufsuntersuchung*. Berlin: Springer-Verlag.
- Wittmann, W.W. (1985). *Evaluationsforschung*. Berlin: Springer-Verlag.
- Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford: Stanford University Press.
- Yates, B. T. (1985). Cost-effectiveness analysis and cost-benefit analysis: An introduction. *Behavioral Assessment*, 7, 207-234.
- Yates, B. T. & Newman, F. L. (1980). Approaches to Cost-Effectiveness Analysis and Cost-Benefit Analysis of Psychotherapy. In G. R. Vandenbos. (Eds.), *Psychotherapy: Practice, Research, Policy* (pp.103-162). Beverly Hills: Sage Publications.

- Zielke, M. (1993). *Wirksamkeit stationärer Verhaltenstherapie*. Weinheim: Psychologie Verlags Union.
- Zielke, M. & Mark, N. (1989). Effizienz und Effektivität stationärer psychosomatischer Behandlungen. *Praxis der Klinischen Verhaltensmedizin und Rehabilitation*, 7, 132-144.

Anhang

- 1. Erklärungen**
- 2. Eingesetzte Instrumente**

1 Erklärungen

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die Arbeit wurde an der TU Dresden, Lehrstuhl für Klinische Psychologie (bis 3/1999 Prof. Dr. J. Margraf, seit 4/2000 Prof. Dr. H.-U. Wittchen) angefertigt. Frühere Promotionsversuche gab es nicht. Die Promotionsordnung der TU Dresden wird anerkannt.

Dresden, 16.7.2001

Dipl.-Psych. Frank Jacobi

2 Eingesetzte Instrumente

Kostenfragebogen

„Manual III“ zur Ein-Jahres-Nachuntersuchung der Therapien in der CDS

Fragebogen zur Zahlungsbereitschaft

Datum:

A) Information zu diesem Fragebogen

In unserer Einrichtung wird zur Zeit eine Studie über die Kosten psychischer Probleme und Störungen durchgeführt. Sie können mit Ihrer Mitarbeit und Ihren Informationen einen Beitrag zu einer besseren Gesundheitspolitik leisten, denn es liegen über die Kosten psychischer Probleme bislang nur wenig genaue Angaben vor. Und die bewilligten Kosten für die Behandlung von psychischen Störungen stehen immer noch im Mißverhältnis zu denen anderer medizinischer Leistungen. So werden Behandlungskosten für körperliche Erkrankungen in der Regel ohne jede Bedenken übernommen, während die Krankenkassen bei psychotherapeutischen Maßnahmen häufig eine Reihe von speziellen Gutachten und/oder Zuzahlungen der Patienten fordern. Außerdem handelt es sich um wichtige Information für Ihren Therapeuten, da z. B. zahlreiche Vorbehandlungen oder berufliche Probleme für Patienten eine besondere Belastung darstellen können, die im Rahmen einer Therapie durchaus zur Sprache kommen sollten.

Wir möchten Sie daher bitten, an dieser Studie teilzunehmen, indem Sie folgenden Fragebogen bearbeiten.

Die Fragen beziehen sich auf bisherige ambulante Behandlungen (Teil B), stationäre Aufenthalte (Teil C) sowie um berufliche Beeinträchtigungen (Teil D). Dabei sind für uns besonders die letzten zwei Jahre interessant. Es kann mitunter sehr schwierig sein, sich über einen so langen Zeitraum zurückzuerinnern. Falls Sie sich nicht sicher sind, schauen Sie diesbezüglich einfach noch einmal in Ihren persönlichen Unterlagen, Terminkalendern oder ähnlichem nach; oder besprechen Sie dies mit Ihrem Partner oder anderen Personen, die sich daran noch erinnern können. Sicherlich werden Sie manche Angaben nur schätzen können; Sie haben in entsprechenden Zusatzfragen die Gelegenheit, Ihre Sicherheit solcher Schätzungen anzugeben. Falls Sie mit der einen oder anderen Begrifflichkeit nicht vertraut sein sollten, fragen Sie einfach bei Ihrem Therapeuten/Ihrer Therapeutin im diagnostischen Interview nach.

Selbstverständlich werden alle Ihre hier gemachten Angaben - wie alle unsere Fragebogendaten - anonymisiert und vertraulich behandelt und ausgewertet.

Bei den *kursiv* gedruckten Passagen handelt es sich um die Fragen, bei den grau unterlegten Passagen um Hinweise für die Beantwortung, also:

1. *Sind Sie bereit, an dieser Untersuchung teilzunehmen?*

JA

NEIN

Wenn JA, bitte umblättern und weiter zu **B)**

B) Praxiskontakte: ambulante Diagnostik und Heilbehandlungen

Was ist (in Stichworten) der aktuelle Anlaß, aufgrund dessen Sie sich in dieser Einrichtung in Behandlung begeben wollen?

Bitte geben Sie den Krankheitsbeginn an, d.h. wann Ihre Beschwerden und Symptome, derentwegen Sie die Behandlung in unserer Einrichtung aufgesucht haben, erstmals aufgetreten sind.

Monat/Jahr: _____

1. Vergangene Psychotherapien

1.1 Haben Sie aufgrund der Beschwerden und Symptome, derentwegen Sie die Behandlung in unserer Einrichtung aufgesucht haben, jemals vorher eine ambulante Psychotherapie durchgeführt?

JA

NEIN

Wenn JA, bitte in der folgenden Tabelle bei der entsprechenden Psychotherapieart eintragen, bei welchem Therapeuten mit wievielen Sitzungen in welchem Zeitraum; außerdem bitte ankreuzen, mit welchem Erfolg. Wenn NEIN, bitte weiter zu Seite 5.

Psychotherapieart	Berufsgruppe Behandler ¹	Zahl der Sitzungen (à 50 Minuten)	Zeitraum (von / bis)	Erfolg: Wie fühlten Sie sich nach der Therapie im Vergleich zu vorher? Wie anhaltend war der Erfolg? (Bitte auf den Skalen ankreuzen)
Entspannungstraining oder Autogenes Training				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg
Hypnose, Meditation oder ähnliches				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg

¹ Berufsgruppe Behandler: Bitte für Diplompsychologen eine "1", für Arzt eine "2" und für sonstige Berufsgruppe den Namen der Berufsgruppe oder falls nicht bekannt eine "3" eintragen

Psychotherapieart	Berufsgruppe Behandler ²	Zahl der Sitzungen (á 50 Minuten)	Zeitraum (von / bis)	Erfolg: Wie fühlten Sie sich nach der Therapie im Vergleich zu vorher? Wie anhaltend war der Erfolg? (Bitte auf den Skalen ankreuzen)
Gesprächspsychotherapie				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg
Psychoanalytische oder tiefenpsychologische Behandlung				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg
Konfrontation mit gefürchteten Reizen oder Situationen				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg
Andere verhaltenstherapeutische Methoden (außer Konfrontation)				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg
Gestalttherapie				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg
Familientherapie				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg
Körperbezogene Therapie				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechter 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg

² Berufsgruppe Behandler: Bitte für Diplompsychologen eine "1", für Arzt eine "2" und für sonstige Berufsgruppe den Namen der Berufsgruppe oder falls nicht bekannt eine "3" eintragen

sonstige: _____ _____ _____				1.....2.....3.....4.....5.....6.....7 sehr viel ...viel... etwas... unver- etwas... viel... sehr viel... besser ändert schlechte 1.....2.....3.....4 Kein... kurzfristiger... mittelfristiger... anhaltender... Erfolg
-----------------------------------	--	--	--	---

Da es manchmal schwierig zu entscheiden ist, was für ein Name zu welcher Therapieform gehört, möchten wir Sie bitten, kurz zu beschreiben, was bei Ihrer Therapie gemacht wurde:

1.2 Wie kam der Entschluß für eine Psychotherapie damals zustande?

eigene Initiative

Empfehlung wurde ausgesprochen von: _____

wurde überwiesen von: _____

Im allgemeinen fällt es schwer, sich über einen so langen Zeitraum genau zurückzuerinnern. Wie sicher sind Sie sich bezüglich der Vollständigkeit Ihrer Angaben zu vergangenen ambulanten Psychotherapien?

Bitte Sicherheit markieren; wichtiger als der exakte Zeitpunkt oder etwa der Berufsgruppe des Behandlers ist hier die Anzahl der Psychotherapiestunden

Ich bin mir bezüglich der Anzahl der Psychotherapiestunden...

völligmäßig.....ziemlich ganz
 unsicher sicher sicher sicher

2. Weitere Praxiskontakte

2.1. Haben Sie aufgrund der Beschwerden und Symptome, derentwegen Sie die Behandlung in unserer Einrichtung aufgesucht haben, jemals vorher einen Psychologen, Psychiater, einen anderen Arzt, Heilpraktiker oder anderen Heilbehandler aufgesucht? (Psychotherapien mit mehr als fünf Sitzungen nicht mitrechnen; diese wurden im vorigen Abschnitt erfragt)
Oder ist deswegen ein Notarzt bei Ihnen gewesen bzw. haben Sie eine Notaufnahme aufgesucht?

- JA NEIN

Wenn NEIN, weiter zu C) auf Seite 8, wenn JA, bitte fortfahren.

2.2 Was für ein Spezialist war das? Wie oft sind Sie wegen dieser Beschwerden bei diesem Spezialisten gewesen?

Bitte ankreuzen und (gegebenenfalls mehrere) Namen eintragen; außerdem in rechte Spalten ungefähres Datum Erstkontakt und die Gesamtzahl der Kontakte eintragen; PSYCHOTHERAPIESITZUNGEN ODER STATIONÄRE AUFENTHALTE HIER NICHT MITRECHNEN (werden in eigenen Abschnitten erfragt)

	Name des Spezialisten	Datum Erstkontakt (Monat/Jahr)	Anzahl der Praxiskontakte
<input type="radio"/>	Psychologe (nur einzelne Beratungen, keine Psychotherapien mit mehr als 5 Sitzungen)	I	I
		I	I
		I	I
		I	I
		I	I
<input type="radio"/>	Psychiater, Neurologe	I	I
		I	I
		I	I
		I	I
<input type="radio"/>	Allgemeinarzt, Hausarzt	I	I
		I	I
		I	I
		I	I
<input type="radio"/>	Internist	I	I
		I	I
		I	I
		I	I
<input type="radio"/>	Kardiologe	I	I
		I	I
		I	I
<input type="radio"/>	Notarzt/Notaufnahme (nur Erstkontakt und Anzahl angeben)	I	I
<input type="radio"/>	Gynäkologe	I	I
		I	I

O	Heilpraktiker	_____ I _____ I
		_____ I _____ I
		_____ I _____ I
		_____ I _____ I
O	Andere (z.B. Seel- sorger, Beratungs- stelle; bitte be- schreiben)	_____ I _____ I
		_____ I _____ I
		_____ I _____ I
		_____ I _____ I

2.3 Was für diagnostische Maßnahmen wurden durch diese Spezialisten durchgeführt?

Bitte beschreiben (z.B. Befragung über Krankheitsgeschichte (Anamnese), psychiatrische Diagnostik, kardiologische Diagnostik, neurologische Diagnostik, psychologische Fragebögen etc.)

Art der diagnostischen Maßnahmen: _____

2.4 Welcher Art war die Behandlung?

2.4.1 Allgemeine Beratung JA NEIN

2.4.2 Medikamente JA NEIN

Wenn JA, bitte angeben, welche Medikamente das waren und wieviel und wie lange Sie diese genommen haben sowie wer sie verschrieben hat

Medikament	Dosis	Dauer	verschrieben von
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

2.4.3 Wie erfolgreich waren bei Ihnen die Behandlung durch Beratungen und Medikamente, dh. wie fühlten Sie sich nach der Behandlung im Vergleich zu vorher?

Bitte auf den folgenden Skalen ankreuzen. Bei mehreren Behandlungen mit unterschiedlichen Ergebnissen zusammen mit dem Therapeuten ausfüllen.

1.....2.....3.....4.....5.....6.....7
 sehr viel ... viel... etwas... unver- etwas... viel... sehr viel...
 besser ändert schlechter

2.4.4 Wie anhaltend war der Erfolg?

1.....2.....3.....4
 Kein... kurzfristiger... mittelfristiger... anhaltender...
 Erfolg

Im allgemeinen fällt es schwer, sich für einen längeren Zeitraum genau zurückzuerinnern. Wie sicher sind Sie sich bezüglich der Vollständigkeit Ihrer Angaben?

Bitte Sicherheit markieren; wichtiger als der exakte Zeitpunkt ist hier die Anzahl der Praxiskontakte

Ich bin mir bezüglich der Vollständigkeit der angegebenen Praxiskontakte...

völligmäßig.....ziemlich ganz
 unsicher sicher sicher sicher

3. Hat einer dieser Spezialisten noch andere Heilmaßnahmen verordnet bzw. durchgeführt (Z.B. Massagen, Akupunktur, Heilbäder)

JA NEIN

Wenn JA, bitte auf folgender Seite angeben, welche und wieviel

Art der sonstigen Heilmaßnahmen

Menge

_____	_____
_____	_____
_____	_____

4. **Nicht immer bezahlt die Krankenkasse sämtliche Kosten einer Heilbehandlung. Wer hat die Kosten dieser bisher besprochenen Heilmaßnahmen übernommen, wenn Ihre Krankenkasse nicht zu 100% bezahlt hat?**

Bitte für die Behandlungen angeben, die **nicht** zu 100% von der Kasse gedeckt wurden, welcher Kostenträger (private Kostenübernahme oder Zuzahlung durch Sie oder Angehörige, Zusatzversicherung, Rentenversicherung, sonstige) zu wieviel % bezahlt hat

Kostenträger	Heilbehandlung (bitte von oben übernehmen)	% bezahlt
_____	_____	_____
_____	_____	_____
_____	_____	_____

C) Stationäre Aufenthalte

1. **Sind Sie aufgrund der Beschwerden und Symptome, derentwegen Sie unsere Einrichtung aufgesucht haben, jemals schon einmal stationär behandelt oder untersucht worden, oder haben Sie eine Kur gemacht?**

 JA

 NEIN

Wenn NEIN, bitte weiter zu D) auf Seite 9 unten, wenn JA, bitte angeben, was für eine Klinik das war und wieviele Tage Sie dort waren

Art der Klinik (z.B. Kurklinik, Psychiatrie, Psychosomatik, Rehabilitationsklinik ...)	Name und Ort der Klinik	Aufenthaltsdauer
_____	_____	_____
_____	_____	_____
_____	_____	_____

2. **Wer hat Sie dorthin überwiesen?** _____

3. **Was wurde dort mit Ihnen gemacht?**

Bitte kurz beschreiben, was für diagnostische, psychotherapeutische (Einzel- und Gruppengespräche) und sonstige therapeutische Maßnahmen (z.B. Entspannung, Arbeitstherapie, Massagen etc) dort durchgeführt wurden

Zahl der psychotherapeutischen Einzelsitzungen: _____

Zahl der psychotherapeutischen Gruppensitzungen: _____

weitere stationäre diagnostische und therapeutische Maßnahmen (z. B. Medikamente, Physiotherapie, Maltherapie etc.):

_____**4. Wie fühlten Sie sich nach der stationären Behandlung im Vergleich zu vorher? (Bei mehreren Behandlungen mit unterschiedlichem Erfolg mit Therapeuten zusammen ausfüllen)**

1.....2.....3.....4.....5.....6.....7
 sehr viel ... viel... etwas... unver- etwas... viel... sehr viel...
 besser ändert schlechter

5. Wie anhaltend war der Erfolg?

1.....2.....3.....4
 Kein... kurzfristiger... mittelfristiger... anhaltender...
 Erfolg

Im allgemeinen fällt es schwer, sich für eine so lange Zeit genau zurückzuerinnern. Wie sicher sind Sie sich bezüglich der Vollständigkeit Ihrer Angaben zu stationärer Behandlung?

Bitte Sicherheit angeben; wichtiger als der exakte Zeitpunkt ist hier die Anzahl der Tage, die Sie stationär verbracht haben

Ich bin mir bezüglich der Vollständigkeit der angegebenen stationären Aufenthalte...

völligmäßig.....ziemlich ganz
 unsicher sicher sicher sicher

D) Arbeitsfähigkeit**1. Sind Sie aufgrund Ihrer Beschwerden, derentwegen Sie unsere Einrichtung aufgesucht haben, berentet worden?**

JA NEIN

Wenn JA, bitte angeben seit wann bzw. für welchen Zeitraum

Berentung seit/von _____ bis _____

PAT:

Datum:

Untersuchungsmanual III

Ein-Jahres-Follow-up

Der folgende Fragebogen enthält eine Reihe von Fragen zu Ihren Erfahrungen und Verhaltensweisen in unterschiedlichen Lebensbereichen. **Alle Fragen beziehen sich auf den Zeitraum seit Ende Ihrer Therapie bei der Christoph-Dornier-Stiftung.** Bitte füllen Sie den Fragebogen allein aus, ohne andere Personen um ihr Urteil oder ihre Meinung zu bitten. Beantworten Sie bitte jede Frage bzw. kreuzen Sie die jeweils zutreffende der vorgegebenen Antwortmöglichkeiten an. Sie können selbstverständlich weitere Bemerkungen am Rand hinzufügen, falls Sie dies wünschen. Sämtliche Angaben werden streng vertraulich behandelt. Um die Lesbarkeit Ihrer Antworten sicherzustellen, schreiben Sie bitte mit Druckbuchstaben.

Teil 1: Angaben zur Person

Bitte tragen Sie die folgenden Informationen ein:

Name: _____ Anschrift: _____

Telefon (tagsüber / abends): _____

Wann sind Sie am besten telefonisch erreichbar? _____

Planen Sie, in der nächsten Zeit umzuziehen? Nein Ja

Wenn ja:

Wann? _____ 19 ____ Wohin? _____

Familienstand (Mehrfachangaben möglich):

- | | |
|--|---------------------------------------|
| <input type="radio"/> ledig | <input type="radio"/> verheiratet |
| <input type="radio"/> zum 2. Mal oder öfters verheiratet | <input type="radio"/> getrennt lebend |
| <input type="radio"/> geschieden | <input type="radio"/> verwitwet |
| <input type="radio"/> unverheiratet mit festem Partner zusammenlebend | |
| <input type="radio"/> feste Partnerbeziehung, aber in getrennten Haushalten lebend | |
| <input type="radio"/> keine feste Partnerbeziehung, aber sexuelle Kontakte | |
| <input type="radio"/> weder feste Partnerschaft noch sexuelle Kontakte | |

Sind Sie zur Zeit berufstätig? Nein Ja (ganztags) Ja (Teilzeit)

Teil 2: Angaben zu Therapieerfolg und zu immer noch bestehenden Beeinträchtigungen durch die Problematik

Bitte geben Sie auf dieser Seite an, wie Sie sich heute **im Vergleich zum Beginn der Therapie** fühlen. Kreuzen Sie zu diesem Zweck bei den folgenden Skalen jeweils die Zahl an, die am besten auf Sie zutrifft. Die Bedeutungen der Zahlen sind jeweils unter den Zahlen angegeben.

a) Therapieerfolg

Im Vergleich zum Beginn der Therapie fühle ich mich:

1 2 3 4 5 6 7
sehr viel... viel... etwas... unverändert etwas... viel... sehr viel...
b e s s e r *s c h l e c h t e r*

b) Arbeit oder Ausbildung

(Falls Sie keine Vollzeit- oder Teilzeitarbeit haben und auch nicht in einer Ausbildung stehen, streichen Sie bitte die ganze Skala durch.)

Schätzen Sie Ihre immer noch bestehende Beeinträchtigung (durch die Probleme, derentwegen Sie die Behandlung in unserer Einrichtung aufgesucht haben) bei Arbeit oder Ausbildung auf dieser 0-4 Skala ein:

0 1 2 3 4
gar nicht... ein wenig... mäßig... schwer... sehr schwer/
massiv beeinträchtigt
(kann nicht arbeiten)

c) Freizeit bzw. Sozialleben

(Zum Beispiel Zusammenkünfte mit Freunden, Feste, Ausgehen, Unterhaltung, Theater, Kino, Konzerte, Besuche, Vereine, Sport, Urlaub etc.)

Schätzen Sie Ihre immer noch bestehende Beeinträchtigung (durch die Probleme, derentwegen Sie die Behandlung in unserer Einrichtung aufgesucht haben) in Ihrer Freizeit auf dieser 0-4 Skala ein:

0 1 2 3 4
gar nicht... ein wenig... mäßig... schwer... sehr schwer/
massiv beeinträchtigt
(kann nichts davon tun)

d) Familienleben und häusliche Pflichten

(Zum Beispiel Ihr Verhältnis zu anderen Familienmitgliedern, Hausarbeiten, Rechnungen bezahlen, Einkaufen, Besorgungen, Saubermachen etc.)

Schätzen Sie Ihre immer noch bestehende Beeinträchtigung (durch die Probleme, derentwegen Sie die Behandlung in unserer Einrichtung aufgesucht haben) in Familienleben oder häuslichen Pflichten auf dieser 0-4 Skala ein:

0 1 2 3 4
gar nicht... ein wenig... mäßig... schwer... sehr schwer/
massiv beeinträchtigt
(kann nichts davon tun)

Teil 3: Angaben zur körperlichen Gesundheit

Bitte tragen Sie die folgenden Informationen ein:

Wann sind Sie das letzte Mal von einem Arzt gründlich untersucht worden? _____

Hatten oder haben Sie ernsthafte Probleme mit Ihrer Gesundheit? Nein Ja

Wenn ja: Welcher Art waren diese Gesundheitsprobleme?

Seit wann haben Sie diese Probleme? _____

Teil 4: Behandlung alter oder neu aufgetretener psychischer Probleme

Haben Sie seit Ende der Therapie in der Christoph-Dornier-Stiftung wegen Ängsten, Depressionen oder anderen psychischen Beschwerden noch weitere Behandlungen aufgesucht?

Nein Ja

Wenn Nein, gehen Sie bitte zum nächsten Abschnitt (Teil 5) auf Seite 6 über.

Wenn ja:

Um was für Beschwerden handelte es sich? _____

Wer führte die Behandlung(en) durch? Bitte geben Sie die jeweilige Berufsgruppe an
(Psychologe, Psychiater, Allgemeinmediziner, Facharzt (bitte spezifizieren), andere)

Die Folgenden Fragen beziehen sich auf die Art weiterer Behandlungen.

a) Psychotherapie: Haben Sie seit Ende der Therapie weitere psychotherapeutische Behandlungen in Anspruch genommen? Nein Ja

Wenn Ja, bitte Psychotherapieart angeben:

- Entspannungstraining oder Autogenes Training
- Hypnose, Meditation oder ähnliches
- Gesprächspsychotherapie
- Psychoanalytische oder tiefenpsychologische Behandlung
- Konfrontation mit gefürchteten Reizen oder Situationen
- andere verhaltenstherapeutische Methoden
- Gestalttherapie
- Familientherapie
- körperbezogene Therapie
- andere psychotherapeutische Behandlungen

Wenn ja: Welcher Art? _____

Bitte geben Sie die Zahl der Sitzungen (à 50 Minuten) an: _____

b) **Allgemeine Beratung** Nein Ja wenn ja, bitte Anzahl angeben: _____

c) **Medikamente** Nein Ja

Wenn ja: Geben Sie bitte an, welche Medikamente das waren, wieviel und wie lange Sie diese eingenommen haben und durch wen sie verschrieben wurden.

Medikament (Name)	Dosierung (Menge)	Dauer	verschrieben von
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

Haben Sie darüber hinaus rezeptfreie Medikamente eingenommen?

_____	_____	_____
_____	_____	_____
_____	_____	_____

d) **andere Heilmaßnahmen** (z.B. Massagen, Akupunktur, Heilbäder) Nein Ja

Wenn ja: Welche? _____

e) **stationäre(r) Aufenthalt(e)** Nein Ja

Wenn ja: Art der Klinik(en): _____
Name und Ort der Klinik(en): _____
Aufenthaltsdauer : _____

Bitte geben Sie die Zahl der Arztbesuche aufgrund alter oder neu aufgetretener psychischer Probleme seit Therapieende an (ausgenommen Psychotherapie): _____

Wir möchten Sie nun nach dem Erfolg dieser weiteren Behandlungen befragen. Wie fühlten Sie danach im Vergleich zu vorher?

Bitte tragen Sie die Behandlung(en) in die obere Zeile ein und kreuzen das jeweils zutreffende an:

Behandlung:				
sehr viel besser				
viel besser				
etwas besser				
unverändert				
etwas schlechter				
viel schlechter				
sehr viel schlechter				

Teil 4: Arbeitsfähigkeit

Bitte tragen Sie die folgenden Informationen ein:

Sind Sie wegen alten oder neu aufgetretenen psychischen Beschwerden berentet worden? Nein Ja

Wenn ja: für welchen Zeitraum? _____

Wie viele Tage waren Sie seit Ende der Therapie in der Christoph-Dornier-Stiftung wegen dieser Beschwerden nicht in der Lage, Ihrer Arbeit nachzugehen (diese Frage betrifft auch nichterwerbstätige Arbeit)?

_____Tage

Wie viele Tage davon waren Sie krankgeschrieben?

_____Tage

Wie viele Tage haben Sie seit Ende der Therapie in der Christoph-Dornier-Stiftung wegen dieser Beschwerden zusätzlich Urlaub genommen (bezahlt oder unbezahlt)?

_____Tage bezahlt unbezahlt

Gab es für Sie bezüglich Ausbildung und Beruf Veränderungen, die Sie mit der Therapie in Verbindung bringen (z. B. neue berufliche Möglichkeiten, höherer Verdienst o.Ä.)?

Datum:

Code:

1. Zeitlicher Aufwand

Stellen Sie sich vor, daß Sie vor Beginn der Behandlung erfahren, daß die Therapie nur sehr weit von Ihrem Wohnort entfernt stattfinden kann. Die Krankenkasse wäre zwar bereit, die Fahrtkosten und Kosten für anfallende Übernachtungen zu übernehmen, doch es entstünde ein erheblicher zeitlicher Aufwand.

Angenommen, die Therapie würde ein halbes Jahr dauern:

a) *Wieviel Zeit für Therapiesitzungen, Hin- und Rückfahrt sowie Übernachtung wären Sie in diesem halben Jahr wöchentlich zu opfern bereit, angefangen von einer Stunde bis zu wieviel Tagen pro Woche?*

_____ pro Woche

b) *Wieviel von Ihrer frei verfügbaren Zeit (Zeit, frei von absolut notwendigen beruflichen und familiären Verpflichtungen) ist der Zeitbetrag in a) in etwa?*

wenig eher wenig die Hälfte viel alles

Freizeit bedeutet jedem Menschen unterschiedlich viel. Manche brauchen daher viel Freizeit, andere sind mit weniger zufrieden.

Wie beurteilen Sie das Ausmaß Ihrer Freizeit?

zu wenig eher zu wenig gerade richtig eher zu viel zu viel

2. Maximale Zahlungsbereitschaft

Stellen Sie sich vor, es gäbe keine Krankenkassen. Sie müssen sich vor Beginn der Therapie entscheiden, ob Sie auch bei Selbstzahlung durchführen möchten. Mit dem heutigen Wissen darum, was die Therapie Ihnen gebracht hat:

a) *Wieviel DM wären Sie unter Berücksichtigung Ihrer finanziellen Umstände höchstens für die gesamte Therapie zu zahlen bereit?*

Lassen Sie dabei die tatsächlichen Kosten einer Therapie außer Acht. Uns interessiert nur Ihre subjektive Einschätzung. Bedenken Sie, daß Sie diese Ausgaben an anderer Stelle einsparen müßten!

_____ DM

Ich würde nicht mehr zahlen, weil _____

Ich würde so viel zahlen, weil _____

Datum:

Code:

Da ein und dieselbe Summe für jeden unterschiedlich viel bedeutet, benötigen wir nun noch Angaben zu Ihrer finanziellen Situation.

b) Wie zufrieden sind Sie mit Ihrer gegenwärtigen finanziellen Situation?

Bedenken Sie, daß dabei neben dem Einkommen u. a. auch Kredite, Schulden, Ersparnisse, Nebeneinkünfte etc. eine Rolle spielen.

sehr gut gut zufriedenstellend ausreichend (knapp) schlecht

c) Wie hoch ist Ihr monatliches Gesamthaushalts-Nettoeinkommen?

*(also das Einkommen aller Mitverdienenden nach Abzug der Steuern zusammengerechnet)
Bei Wohngemeinschaften, in denen das Einkommen nicht geteilt wird: nur das eigene Einkommen!*

unter 1000 DM
1000 - 2000 DM
2000 - 3000 DM
3000 - 4000 DM
4000 - 5000 DM
5000 - 6000 DM
6000 - 7000 DM
7000 - 8000 DM
8000 - 9000 DM
9000 - 10000 DM
über 10000 DM Wieviel? _____DM

Wie viele Personen leben von diesem Einkommen? _____

Haben Sie im Rahmen der Therapie finanzielle Ausgaben gehabt, wie Selbstbeteiligung, wenn die Krankenkasse nicht zahlte, Fahrtkosten, weitere durch die Therapie entstandene Nebenkosten?

Nein Ja

Wenn ja, wie viel haben Sie ausgegeben? _____DM

Wie beurteilen Sie diesen Fragebogen? Gab es Verständnisprobleme? Wie schwierig war es den Fragebogen auszufüllen? Finden Sie, daß derartige Fragebögen genutzt werden können, um zu messen, wie gut eine Therapie ist?

Vielen Dank für Ihre Mitarbeit!

Cost-Benefit-Calculator

1. Kosten der untersuchten Behandlung

Definieren Sie die Behandlungseinheiten (z.B. Psychotherapiesitzungen á 50 Minuten, Tage stationären Aufenthalts, Tage in Tageskliniken, etc.), und geben Sie deren Preis und Anzahl ein. Achten Sie bitte bei allen Angaben darauf, Dezimalstellen mit einem Punkt (nicht mit Komma) einzutragen.

Behandlungseinheit	Preis der Behandlungseinheit in DM	(durchschnittliche) Anzahl der Behandlungseinheiten
VT à 50 Minuten	117.30	41.5

Geben Sie ggf. den (durchschnittlichen) Gesamtbetrag weiterer Kosten der Behandlung in DM an (z.B. Nebenkosten durch Anreise, Kosten für Konfrontationsübungen):

450

2. Störungskosten vor und nach der Behandlung

Geben Sie den untersuchten Zeitraum *vor* der Behandlung in Jahren an (Standardwert = 1 Jahr):

1

Geben Sie den untersuchten Zeitraum *nach* der Behandlung in Jahren an (Standardwert = 1 Jahr):

1

Geben Sie den Zeitraum nach der Behandlung in Jahren an, für den das Kosten-Nutzen-Verhältnis berechnet werden soll (Standardwert = 1 Jahr):

1

Geben Sie die Diskontierungsrate an (empfohlen: 0.05):

0.05

Definieren Sie Kostenfaktoren der Störung (z.B. Psychotherapiesitzungen, Tage stationären Aufenthalts, ambulante Arztkontakte, Arbeitsunfähigkeitstage), und geben Sie deren Preis und Anzahl *vor* und *nach* der Behandlung ein.

Kostenfaktor	Preis pro Einheit des Kostenfaktors in DM	(durchschnittliche) Menge <i>vor</i> der Behandlung für den obenstehenden Zeitraum	(durchschnittliche) Menge <i>nach</i> der Behandlung für den obenstehenden Zeitraum
Arzt 1	60	4.25	0.5
stationär	275	2.8	1.5
PT	120	4.5	0.5
AU	390	11.0	0.2

Die Kosten der Behandlung betragen: DM.

Der Nutzen im ersten Jahr beträgt: DM.

Der Nutzen im Berechnungszeitraum beträgt: DM.

Die Kosten-Nutzen-Relation der untersuchten Behandlung beträgt 1:

