

Schlußbericht: Therapieaufwand und -erfolg bei der psychodynamischen Therapie von Eßstörungen: Eine multizentrische Studie

BMFT-Förderkennzeichen: 01EN9410 (alt:07PSF53)

Vorhabensleiter:	Prof. Dr. med. Horst Kächele
Durchführende Stelle:	Forschungsstelle für Psychotherapie Stuttgart
Geplante Laufzeit:	5 Jahre
Tatsächliche Laufzeit:	September 1993 – März 1999

1.1 Aufgabenstellung

Die Projektgruppe unter Leitung von Prof. Dr. H. Kächele hatte sich vorgenommen, die Beziehung zwischen Therapieaufwand und Therapieerfolg bei der psychodynamischen Behandlung von Eßstörungen, Anorexia Nervosa (AN) und Bulimia Nervosa (BN), in der klinischen Alltagspraxis zu untersuchen. Dabei gliederten sich die Aufgaben nach den folgenden Fragen:

1. Wie werden Eßstörungen in der – psychodynamischen – klinischen Alltagspraxis behandelt? Was erklärt die unterschiedliche Intensität und Dauer?
2. Welche Ergebnisse werden erreicht?
3. Welche Patienten- bzw. Krankheitsmerkmale beeinflussen die Ergebnisse?
4. Welchen Einfluß haben Dauer und Intensität der Behandlung auf die Ergebnisse?
5. Wie verlaufen Eßstörungen über einen mittelfristigen Zeitraum von 2,5 Jahren (Remission, Rückfall, Wiedererkrankung)?

1.2 Voraussetzungen

Das Projekt wurde von der Forschungsstelle für Psychotherapie Stuttgart koordiniert (durchführende Stelle). Über vier regionale Zentren wurde die Durchführung in den schließlich 45 teilnehmenden klinischen Einrichtungen moderiert.

Die Projektgruppe hatte sich eine Geschäftsordnung gegeben. Der gegenüber dem BMBF verantwortliche Leiter, Prof. Dr. H. Kächele, wurde danach von einem Geschäftsführenden Ausschuß (Vorsitz: H. Kächele), einem Mitgliederrat (Vorsitz: H. Kächele) und dem Forum, d.h. der Gesamtheit der Mitglieder (Vorsitz: H. Kächele) unterstützt.

Der Durchführung ging eine 4-jährige Vorbereitungsphase (gefördert von der Breuninger-Stiftung) voraus, in der das Netzwerk zwischen an Forschung interessierten KlinikerInnen und an Klinik interessierten ForscherInnen geknüpft wurde. Eine zentrale Rolle spielte dabei das „Forum für Eßstörungen,, über das zunächst die Projektplanung, später die Koordination und schließlich der Wissenstransfer organisiert wurde.

1.3 Planung und Ablauf des Vorhabens

Mit Eintreffen des Zuwendungsbescheides (07.09.1993) wurde das Team aufgebaut, die vertraglichen Voraussetzungen für die Durchführung geschaffen. Die Fragebogen-Batterie wurde für den Druck vorbereitet und die ModeratorInnen für ihre Aufgaben in den Kliniken geschult. Bereits im November 1993 konnten die ersten PatientInnen in die Studie aufgenommen werden. Im September 1995 waren 1246 PatientInnen für die Studie rekrutiert, von denen nach Ausschluß unbrauchbarer Fälle (wegen fehlender oder inkonsistenter Angaben) schließlich 1171 die Projektstichprobe bilden. Die 2,5-Jahresnachuntersuchung wurde im April 1996 begonnen. Bis Ende August 1998 konnte so der mittelfristige Krankheitsverlauf von 879 PatientInnen (75,1%) erfaßt werden.

In Zusammenarbeit mit den europäischen Partner aus der COST-Action B6 (s.u.) wurde im Juni 1998 in Stuttgart ein internationaler Fachkongress zur Behandlung von Eßstörungen durchgeführt (gefördert von der EU, der DFG, sowie dem Ministerium für Wissenschaft und Kultur Baden-Württembergs).

1.4 Wissenschaftliche Ausgangssituation

Die effiziente Behandlung der Eßstörungen sind aufgrund der Häufigkeit und Schwere der Erkrankung von erheblicher versorgungsepidemiologischer und gesundheitsökonomischer Bedeutung. Allerdings war das empirisch gesicherte Wissen über die Wirksamkeit der therapeutischen Versorgung unter Alltagsbedingungen und die möglichen differentiellen Effekte unterschiedlicher Behandlungsstrategien mager. Studien über den Zusammenhang zwischen Aufwand und Wirkung psychotherapeutischer Behandlungen bei Eßstörungen sind auch heute kaum zu finden (Treasure und Schmidt 1999, Herzog, Hartmann und Sandholz 1996; Mitchell et al., 1993). Dies gilt insbesondere für stationäre Therapieprogramme, die gerade für AN in der Regel zur Anwendung kommen (z.B. Treasure et al. 1995, McKenzie and Joyce, 1992). Obwohl diese in der Regel sehr intensiven Therapien den Einsatz erheblicher therapeutischer Ressourcen verlangen, gab

und gibt es so gut wie keine empirischen Daten über den Einfluß von Umfang, Dauer, Intensität oder Komposition stationärer Behandlungsprogramme auf das Therapieergebnis.

1.5 Zusammenarbeit mit anderen Stellen

Das Projekt bildete das Fundament für eine europäische Zusammenarbeit, die auf Initiative der Forschungsstelle für Psychotherapie über das BMBF als Action B6 im COST-Programm der EU etabliert werden konnte (Oktober 1994). Die Aktualität der Thematik und die Qualität des Forschungsansatzes haben schließlich Forschergruppen aus 19 europäischen Ländern zusammengeführt.

2. Ergebnisse

2.1 Stichprobe

Von November 1993 bis September 1995 wurden von 43 Kliniken aus dem gesamten Bundesgebiet 1247 PatientInnen in die Studie eingebracht. Eingeschlossen wurden alle erwachsenen PatientInnen, die im genannten Zeitraum eine der beteiligten Kliniken aufsuchten und die Kriterien für die Diagnosen Anorexia Nervosa und/oder Bulimia Nervosa nach DSM-III R erfüllten. Ausgeschlossen wurden nur solche PatientInnen, die entweder die Teilnahme verweigerten (nach ausführlicher mündlicher und schriftlicher Aufklärung) oder die physisch, psychisch oder intellektuell zu einer Teilnahme nicht in der Lage waren.

Die teilnehmenden Kliniken repräsentieren ein weites Spektrum der psychodynamisch ausgerichteten stationären Therapie von Eßstörungen in Deutschland. Es sind darunter auf diese Krankheitsgruppe spezialisierte Kliniken mit entsprechend spezialisiertem Therapieprogramm ebenso vertreten wie psychodynamische Kliniken oder Abteilungen, bei denen die Eßstörungen nur einen beschränkten, aber wichtigen Teil der PatientInnenschaft ausmacht, so daß sie für diese PatientInnen oftmals spezielle Therapieangebote bereithalten, sowie Kliniken, in denen nur vereinzelt Eßstörungen - meist ohne weitere Modifikation - nach dem allgemeinen Behandlungsprogramm der Klinik behandelt werden (vgl. den von der Studiengruppe herausgegebenen Therapieführer, Kächele et al. 1995). Der unterschiedlich starken Ausrichtung auf die Behandlung von Eßstörungen entsprechend verhalten sich die Einbringungszahlen.

PatientInnen mit der Diagnose AN machten gut ein Drittel der Stichprobe aus (n= 355); etwas mehr als die Hälfte wurden als BN diagnostiziert (n=647). Eine wichtige Teilgruppe

bildeten die PatientInnen, für die beide Diagnosen zutrafen (n=169). Das DSM IV empfiehlt, diese Gruppe getrennt zu betrachten. Sie wird im folgenden mit dem Kürzel „BN&AN„ bezeichnet. Alle drei hier genannten Eßstörungen sind bekanntlich psychische Erkrankungen, die ganz überwiegend weibliche Jugendliche und junge Frauen betreffen. In der Stichprobe fanden sich weniger als 4 % männliche Patienten. Dieser Anteil ist kleiner als in der Literatur sonst berichtet wird.

Die PatientInnen waren zwischen 17 und 57 Jahre – die AN im Mittel 24,8 (SD5,6) und die BN 25,9 (SD=6,3) - alt (nach Vereinbarung wurden auch diejenigen PatientInnen in die Studie aufgenommen, die im Laufe der stationären Behandlung 18 Jahre alt wurden und auf der Erwachsenenstation behandelt wurden).

Die mittlere Erkrankungsdauer lag bei 7,4 Jahren. Mit durchschnittlich 5,8 Jahren war die Zeit bis zur Behandlung bei den PatientInnen mit AN die kürzeste und bei den PatientInnen mit BN war sie mit durchschnittlich 8,5 Jahren die längste. Ein beträchtlicher Teil der PatientInnen aller drei Subgruppen von Eßstörungen wurde als chronisch krank angesehen, wobei als Kriterium für Chronifizierung eine Krankheitsdauer von 5 oder mehr Jahren angenommen wurde.

Entsprechend dem Studienprotokoll wurde eine Nachuntersuchung 2,5 Jahre nach Therapiebeginn durchgeführt. Die PatientInnen zeigten eine gute Compliance. Von 879 (75,1%) konnten zumindest minimale Informationen über den aktuellen Gesundheitszustand erhoben werden. Ein ausführliches semi-strukturiertes Interview, das sogenannte LIFE (Longitudinal-Interval-Followup-Evaluation), konnte mit 781 (66,7%) geführt werden, während von den übrigen 98 (8,4%) Angaben über Angehörige oder (Haus-)Ärzte eingeholt werden konnten. Nur 64 (5,5%) verweigerten die Nachuntersuchung explizit, 207 (17,7%) konnten nicht aufgefunden werden und für 11 waren aus unterschiedlichen Gründen keine ausreichenden Daten erreicht worden.

Die PatientInnen, die an der 2,5-Jahresuntersuchung teilnahmen, unterschieden sich nur wenig von den Nicht-TeilnehmerInnen (Tab.1). Teilnehmende AN wurden etwa zwei Wochen länger behandelt oder hatten einen etwas höheren Body-Mass-Index (BMI) zur Entlassung oder waren zu Beginn seltener zur Behandlung motiviert. Bei den BN hatten TeilnehmerInnen an der 2,5-Jahresuntersuchung einen geringfügig höheren BMI bei Behandlungsbeginn oder eine tendenziell längere Krankheitsgeschichte oder etwas niedrigere Scores in der Morgan-Russel "Soziale Anpassung". Insgesamt gab es wenig Anzeichen für eine mögliche Verzerrungen durch Selektionseffekte, lediglich der Anteil

chronifizierter PatientInnen in der BN-Stichprobe zur 2,5-Jahresuntersuchung bleibt eventuell zu beachten.

2.2 Stationäre psychodynamische Psychotherapie: Wie werden Eßstörungen behandelt?

Die an der Studie beteiligten Kliniken boten unterschiedliche Therapieprogramme an. Im Rahmen dieser naturalistischen Studie wurde auf jede Einflußnahme auf die üblicherweise praktizierte Therapie verzichtet. Das hatte zur Folge, daß der therapeutische Aufwand für eine Vielfalt von Therapieprogrammen zuverlässig dokumentiert werden mußte, um einen Vergleich dieser unterschiedlichen Programme auf den zentralen Aufwandsdimensionen zu gewährleisten. In der Pilotstudie hatten sich klinikspezifische Dokumentationsbögen als praktikabel erwiesen. Für jede Klinik wurde also eine solche Liste der dort zur Anwendung kommenden therapeutischen Maßnahmen erstellt. Diese Listen bildeten die Grundlage für die patientenbezogene wöchentliche Dokumentation darüber:

- welche therapeutische Maßnahme angewandt wurde
- wie viele Gruppenmitglieder daran teilnahmen (bei Gruppenmaßnahmen)
- wie lange die Maßnahme dauerte (in Minuten pro Woche)
- von wem die Maßnahme durchgeführt wurde (Code zur Berechnung eines Kostenindex)?

Darüber hinaus wurde wöchentlich dokumentiert, wieviel Zeit, evtl. auch in Abwesenheit der PatientIn, für klinische Tätigkeiten verwendet wurde (z.B. patientenbezogene Stationskonferenzen, Diagnose- oder Indikationsbesprechungen, Supervisionen etc), die in diesem Sinne indirekte Kosten verursachten.

Anhand der auf diese Weise wöchentlich dokumentierten Daten zur Art, Menge und Dauer der angewandten therapeutischen Maßnahmen wurde der therapeutische Aufwand skaliert. Als Parameter wurden dabei genutzt:

1. Behandlungsdauer
2. Stundenplan pro Woche (Intensität)
3. Kosten (Summe der Kosten in Arztstunden/BAT IIa, und zwar unter Berücksichtigung der Gruppengröße bei gruppentherapeutischen Maßnahmen).

Diese Daten zu Dauer, Intensität und Kosten demonstrieren die enorme Variation der Behandlung von Eßstörungen in der klinischen Praxis (Abb. 1). Der Median der

Aufenthaltsdauer lag bei 11 Wochen. Das Spektrum reichte von 4,3 (5. Perzentil) bis zu 27,9 Wochen (95. Perzentil). Zwischen den drei Subgruppen bestanden keine Unterschiede hinsichtlich der Behandlungsdauer, wohl aber in bezug auf die Intensität.

Der wöchentliche Stundenplan von BN-Patientinnen umfaßte mit im Median 13,6 Stunden mehr Therapie als der für die Patientinnen mit AN (Median 13,0 Stunden) und BN&AN (Median 13,1 Stunden). Auch im Hinblick auf diesen zentralen Intensitätsparameter erwiesen sich die Behandlungsprogramme für alle drei Subgruppen gleichermaßen als heterogen; sie reichten von 4,0 Stunden für BN bzw. 4,2 Stunden für AN und 5,5 Stunden für BN&AN (5. Perzentil) bis zu einer Wochenstundenzahl von 25,0 Stunden für BN bzw. 27,0 Stunden für AN und 30 Stunden für BN&AN (95. Perzentil). Die Kosten für die tatsächlich angewendeten therapeutischen Maßnahmen (ohne Fixkosten) summierten sich zu 28,5 (Median für BN) bzw. 34,0 (Median für AN) und 34,3 (Median für BN&AN) BAT IIa-Stunden. Die Behandlung für BN-Patientinnen hatte also trotz der höheren Intensität im Hinblick auf die Wochenstundenzahl die niedrigsten Kosten. Dies folgte aus dem für BN im Vergleich zu den beiden anderen Subgruppen größeren Anteil an Maßnahmen, die in Gruppenform durchgeführt wurden.

Von großer Bedeutung für die Organisation therapeutischer Versorgung ist die enorme Variation im Umfang der angewandten Behandlungen zwischen den Kliniken. In Abb. 2a und 2b sind die 21 Kliniken, die 15 oder mehr Patientinnen zur Stichprobe der multizentrischen Studie beigetragen haben, nach dem Median der Behandlungsdauer der angewendeten therapeutischen Maßnahmen geordnet. Die Behandlung in der Klinik mit der längsten Aufenthaltsdauer (Klinik 53, Median 26,6 Wochen) war 4,2 mal länger als in der Klinik mit der kürzesten (Klinik 85, Median 6,3 Wochen). Es gab Kliniken mit einer relativ homogenen Behandlungsdauer über alle Patientinnen hinweg (angezeigt durch die kurzen Boxen der Boxplots in Abb. 2a wie z.B. Klinik 32 oder Klinik 11); in anderen variierte sie stark zwischen den Patientinnen (z.B. Klinik 53 oder Klinik 19 in Abb. 2a). Vergleicht man die Kliniken in bezug auf die Intensität der angewandten Behandlung - als Intensitätsparameter wurde der Umfang des Wochenstundenplanes gewählt -, erhält man ein anderes Bild (Abb. 2b, die Kliniken sind in der selben Weise geordnet wie in Abb. 2a): die Klinik mit der zweitkürzesten Behandlung im Median (Klinik 10) steht hinsichtlich der Intensität an erster Stelle (Median 22,7 Stunden pro Woche). Beide Bilder zusammen demonstrieren, wie unterschiedlich das Therapieangebot der verschiedenen Kliniken hinsichtlich Dauer und Intensität und damit hinsichtlich der Gesamtmenge angewendeter Therapie gestaltet wurde. Diese Heterogenität ist nicht nur von akademischer Bedeutung.

Die Gesamtkosten für die tatsächlich angewendeten therapeutischen Maßnahmen pro Patient (ohne Fixkosten) z.B. waren in Klinik 53 (Median 72,1 Arztstunden/BAT IIa) etwa dreimal so hoch wie in Klinik 85 (Median 21,6 Arztstunden/BAT IIa).

2.3 Wer erhält wieviel Behandlung?

Die im folgenden berichteten Ergebnisse beziehen sich auf eine Stichprobe von 1112 PatientInnen (AN: n=338, BN: n=605, AN&BN: n=169) aus 43 Kliniken, von denen 733 (AN: n=225, BN: n=399, AN&BN: n=109) auch an der 2,5-Jahresuntersuchung teilgenommen haben. D.h. auf die Daten von 59 PatientInnen wurde verzichtet, um eine konsistente Interpretation zu ermöglichen. 20 dieser 59 PatientInnen waren nämlich primär ambulant behandelt worden, bei den übrigen 39 dauerte die stationäre Behandlung weniger als 4 Wochen. Da keine Klinik ein reguläres Behandlungsprogramm von weniger als 4 Wochen angeboten hatte, sind diese Behandlungen als Sonderfälle betrachtet worden.

Methodik

Die Behandlungsdauern variierten sowohl *innerhalb* als auch und *zwischen* den Kliniken. Die Analyse der möglichen Einflußfaktoren geschah daher mit Hilfe Hierarchisch Linearer Modellierung – HLM - (Bryk und Raudenbush 1992; Goldstein 1995). Hierbei handelt es sich um ein Mehrebenen-Regressionsverfahren. Es erlaubt die gleichzeitige Untersuchung der Variabilität der Behandlungsdauern auf beiden Ebenen. Jede PatientIn läßt sich ja genau einer Klinik zuordnen, insofern stehen PatientInnen und Kliniken in hierarchischer Ordnung. In einem Anfangsmodell ohne Prädiktoren wurden daher die Varianzen der Behandlungsdauern innerhalb und zwischen den Kliniken geschätzt. In einem zweiten Schritt wurden dann zunächst Patientenmerkmale und in einem dritten Schritt schließlich dann auch Klinikeigenschaften als Prädiktoren hinzugenommen.

Ergebnisse

Die Modellierung erfolgte in drei Stufen. Als Ausgangsmodell diente ein Modell ohne Prädiktoren ('unconditional model' nach Bryk und Raudenbush, 1992). Mit diesem wurden die Varianzen zwischen und innerhalb der Kliniken geschätzt¹ (erste Spalte in Tab. 3a). Die Bedeutung der patientenbezogenen Faktoren (Tab. 2) für die Behandlungsdauer zeigt sich in der zweiten Spalte von Tab. 3a. Schließlich wurden auf der dritten Stufe auch die Information über die Spezifität des Behandlungsangebots hinzugezogen. Die Steigerung

¹ Da die Annahme homogener Varianzen für die Behandlungsdauern in den Kliniken offensichtlich nicht zutrifft (s. Abb. 2) wird für die Schätzung der *innerhalb* Varianz eine nach Littell et al. (1996) modifizierte Technik angewendet.

der Anpassungsgüte über die Modellstufen hinweg kann der Tab. 3a entnommen werden. Wie man dort sieht, läßt sich mit den patientenbezogenen Merkmalen nur ein kleiner Teil der Varianz *innerhalb* der Kliniken erklären, zur Varianz *zwischen* den Kliniken tragen diese patientenbezogenen Merkmale nichts bei.

Tabelle 3a einfügen

Die detaillierten Ergebnisse für die Modelle der Stufen 2 und 3 sind den Tabellen 3b und 3c dargestellt. Während in Tab. 3b die statistischen Kennwerte (Omnibustests) für *alle* Prädiktoren und Interaktionen aufgeführt sind, werden in Tab. 3c die Effekte, Standardfehler und t-Werte nur der Prädiktoren aufgelistet, die einen statistisch signifikanten Effekt gezeigt haben. Der stärkste Effekt zeigte sich für die Klasse der Kliniken mit einer eigenen Eßstörungenabteilung. Dort erhielten die PatientInnen eine im Durchschnitt um sieben Wochen längere Behandlung. Daneben erwiesen sich einige wenige der patientenbezogenen Merkmale als Prädiktoren mit moderaten Effekten (Tab.3b und c): PatientInnen mit einer niedrigen Behandlungsmotivation zu Beginn, wurden im Durchschnitt um eine halbe Woche kürzer behandelt als solche mit hoher Motivation. Für bulimische PatientInnen gilt, daß ein – relativ großer - Unterschied von einem Punkt in der psychischen Gesamtbelastung (Skala GSI der SCL-90R) mit einer im Durchschnitt um eine halbe Woche längeren Behandlungsdauer korrespondiert. Für eine PatientIn mit Anorexie wird dieser (Haupt-)Effekt durch den Interaktionseffekt weitgehend ausgeglichen. Hier ist der gleiche Unterschied von einem Punkt in der psychischen Gesamtbelastung (GSI) mit einer um 0,24 Wochen kürzeren Behandlung verbunden. Für AN war ein Unterschied im Untergewicht (zu Behandlungsbeginn) von +18%-BMI mit einer um eine Woche längeren Behandlungszeit verbunden. Dagegen hatte das Untergewicht bei BN keine Bedeutung für die Behandlungsdauer. Dieses Beispiel unterstreicht noch einmal die Bedeutung der Interaktionseffekte. In dieser Hinsicht lohnt noch eine Bemerkung zu dem auf den ersten Blick beeindruckenden Effekt von 4,06 der Diagnose. Denn heißt ja noch nicht, daß AN im Durchschnitt 4 Wochen länger als BN behandelt wurden. Es müssen eben noch die Interaktionen berücksichtigt werden. Für AN und BN mit einem Gewicht von 90%BMI z.B. ergab sich lediglich ein Unterschied von 0,56 Wochen.

Tabelle 3b und 3c einfügen

2.4 Welche Ergebnisse werden in der stationären psychodynamischen Behandlung erreicht?

Die Erfolgsraten zu Behandlungsende und zur 2,5-Jahresuntersuchung sind in Tabelle 4 wiedergegeben, und zwar getrennt für die drei Untergruppen von Eßstörungen. Weitgehende Symptombefreiung zu Behandlungsende, so war Erfolg vereinfachend gesprochen hier definiert (s. Abb. Xx), findet man bei AN ebenso wie bei den PatientInnen, die sowohl die Kriterien für AN als auch für BN erfüllen, eher selten. Als erfolgreich behandelt im Sinne der Symptombefreiung wurden zu diesem Zeitpunkt nur 11% bzw. 17% eingeschätzt. Die Unterschiede zwischen der Selbsteinschätzung der PatientInnen und der Experteneinschätzung durch die Therapeuten unterschieden sich wenig. Deutlich besser stellten sich die Ergebnisse zur 2,5-Jahresuntersuchung dar. Ein deutlich größerer Teil der PatientInnen mit AN (36% aus Sicht der Experten bzw. 33% aus der eigenen Sicht der PatientInnen) oder AN&BN (26% bzw. 21%) waren weitgehend symptomfrei.

Für die PatientInnen mit BN ergab sich ein anderes Bild. Die Erfolgsraten zu Behandlungsende lagen deutlich über denen der beiden anderen Gruppen von Eßstörungen. Allerdings unterschieden sich die Einschätzungen aus Therapeutensicht (45%) beträchtlich von denen aus der Sicht der PatientInnen (31%). Bis zur 2,5-Jahresuntersuchung waren die Erfolgsraten jedoch auf 36% (aus Sicht der klinischen Experten) bzw. 22% (aus Sicht der PatientInnen) zurückgegangen.

Abbildung 4 einfügen

2.5 Wer hat welche Erfolgschancen? Welche Rolle spielen die Behandlungsdauer und –intensität?

Methodik

Im Zentrum des Interesses stand der mögliche Einfluß von Dauer und Intensität der Behandlung auf die Erfolgsraten nach 2,5 Jahren. Wegen des naturalistischen Designs der Studie mußte damit gerechnet werden, daß Schätzungen der Effekte durch konfundierende

Variablen verzerrt sein können². Zur Kontrolle der möglichen Verzerrungen wurde ein Vorschlag von Rosenbaum und Rubin (1983) aufgegriffen. Um solche in nicht-randomisierten Studien zu korrigieren, entwickelten diese Autoren in den 80er Jahren den „Propensity-Score“². Dieses Korrekturverfahren wurde für den Vergleich von zwei Behandlungsalternativen entwickelt. Es gliedert sich in zwei Schritten: Im ersten werden mit logistischer Regression oder Diskriminanzanalyse die Propensity-Scores bestimmt. Im zweiten Schritt wird die Stichprobe (z.B.) anhand der Quintile der Propensity-Score-Verteilung aufgeteilt. Innerhalb der Quintile sind die konfundierenden Variablen zwischen beiden Alternativen in der Regel besser ausbalanciert, so daß dort der Effekt (relativ) unverzerrt geschätzt werden kann. Der Schätzwert für den Gesamt-Effekt ergibt sich schließlich als gemittelter Effekt über die Quintile hinweg (vgl. Rosenbaum und Rubin, 1984).

Konkret wurde in der vorliegenden Studie die statistische Analyse von Einflußfaktoren für den Therapieerfolg nach Vorschlägen von Hosmer und Lemeshow (1989) durchgeführt (vgl. auch D'Agostino 1998). Dabei erfolgte die Analyse mit der logistischen Regression, wobei ebenfalls die Technik der Propensity-Scores zur Korrektur eingesetzt wird. Wegen der großen Anzahl der möglichen Faktoren in dieser Studie bestand zudem die Gefahr der Überspezifizierung. Deshalb wurden zusätzlich korrigierte Schätzungen mit Hilfe eines Bootstrapverfahrens (Efron und Tibshirani 1993) bestimmt.

² So könnten z.B. chronisch erkrankte Patientinnen eher längere als kürzeren Behandlungen erhalten haben. Unter der weiteren Annahme, daß diese PatientInnen eine schlechtere Prognose haben, würde sich eine Fehlschätzung der Behandlungserfolge für kurze und lange Behandlungen ergeben.

Da die Methode der Propensity-Scores nur für den Vergleich von jeweils zwei Alternativen angewendet werden kann, wurde für die folgenden Analysen die Behandlungsdauer dichotomisiert, also ein Vergleich von kurzen und langen Behandlungen durchgeführt. Dazu wurden die Behandlungen am Median aufgeteilt (AN: 11,1 bzw. BN: 11,4 Wochen). Alle in Tab. 2 aufgeführten Variablen – mit Ausnahme der Behandlungsdauer und –intensität – wurden in die Propensity-Korrektur einbezogen. Abbildung 4 illustriert den Effekt am Beispiel des Untergewichtes. Insgesamt reduzierte die Propensity-Korrektur die Unbalanciertheit der berücksichtigten Konfounder bei AN um 72% und bei BN um 88% (gerechnet nach Liebermann et al. 1996). Anzumerken ist, daß der Klinikeinfluß auf diese Weise nicht kontrolliert werden konnte, da in einigen der Kliniken nur kurze Behandlungen zur Anwendung gekommen waren.

Tabelle 4 einfügen

Ergebnisse

Die Schätzungen der Erfolgsraten für kurze und lange Behandlungsdauern bei der 2,5-Jahresuntersuchung unterschieden sich in den fünf Quintilen nur wenig. Die gemittelten Werte sind in Tabelle 5 zusammen mit den Odds-Ratios (mit 95%-Konfidenzintervallen) angegeben³. Zum Entlassungszeitpunkt ließ sich nur bei BN ein Unterschied erkennen, wie die korrigierten Odds-Ratios anzeigen. PatientInnen mit BN haben demnach bei kurzer Behandlungen eine um ca. 33% höhere Chance auf Therapieerfolg (aus Sicht der Therapeuten). Zur 2,5-Jahresuntersuchung zeigten sich keine Unterschiede zwischen kurzen und langen Behandlungen. Weder die unkorrigierten noch die korrigierten Odds-Ratios gaben einen Hinweis auf einen Einfluß der Behandlungsdauer auf die Erfolgsraten.

Tabelle 5 einfügen

³ Ein Odds-Ratio von 1 bedeutet keinen Unterschied zwischen langen und kurzen Behandlungen. Ein Odds-Ratio von 2 bedeutet eine doppelt so hohe Erfolgsrate der langen Behandlungen und ein Odds-Ratio von 0.5 eine halb so große Erfolgsrate im Vergleich zu den kurzen Behandlungen.

Diese globale Aussage muß allerdings differenziert werden. Die folgenden Analysen fokussieren auf mögliche Einflußfaktoren zum Zeitpunkt der 2,5-Jahresuntersuchung.⁴ Eine hierarchische logistische Regressions-Analyse⁵ der Erfolgsraten ließ keinen systematischen Beitrag der Variabilität zwischen den Kliniken erkennen. Die Untersuchung der prognostischen Faktoren für Therapieerfolg erfolgte daher mit einer einfachen logistischen Regression, wobei die Variabilität der Erfolgsraten zwischen den Kliniken vernachlässigt wurde.

Die Analyse wurde getrennt für AN und BN durchgeführt. Als Prädiktoren für den Behandlungserfolg dienten wiederum alle in Tabelle 2 aufgeführten Variablen; zusätzlich gingen die Behandlungsdauer (lang vs. kurz) und –intensität sowie der Propensity-Score mit ein. Um ein differenziertes Bild über den möglichen Einfluß der Behandlungsdauer zu gewinnen, wurde für alle anderen Prädiktoren die Interaktion mit der Dauer untersucht. Die resultierenden Modelle sind in Tabelle 6 beschrieben. Sie unterscheiden sich für AN und BN sowohl im Hinblick auf die eingehenden Faktoren als auch auf die Anpassungsgüte (nach Kontrolle der Behandlungsdauer). Insbesondere gilt es zu beachten, daß sich für AN keine Interaktionseffekte nachweisen ließen, während bei der BN zwei wichtige Interaktionseffekte zu berücksichtigen sind:

Für die AN zeigten sich auch nach Kontrolle der anderen Faktoren keine Effekte auf die Erfolgchancen für die Behandlungsdauer (lang vs. kurz) und –intensität. Dagegen ließ sich für die Variablen Untergewicht, Erkrankungsdauer und Persönlichkeit ein Einfluß auf die Erfolgsaussichten erkennen. Ein niedriges Gewicht (%BMI) zu Behandlungsbeginn, eine lange Erkrankungsdauer und ausgeprägte Abweichungen in den Persönlichkeitsskalen des FPI (‘Extrem-Score,‘) zeigten sich als Risikofaktoren für niedrige Erfolgsraten. Dagegen waren für PatientInnen die Erfolgsaussichten günstiger, die über einen sexuellen Mißbrauch berichtet hatten, die ein höheres Wunschgewicht angegeben hatten oder die sich durch die Eßstörung stärker beeinträchtigt beschrieben hatten.

⁴ Das liegt im wesentlichen darin begründet, daß die Menge der eingesetzten Behandlung, ihre Dauer und Intensität, dann als echter Prädiktor und potentiell kausaler Faktor interpretiert werden kann. Dagegen ist davon auszugehen, daß Entscheidungen über die Beendigung von der Therapie in einem – allerdings derzeit unbekannten und innerhalb der Studie nicht zu klärenden – Zusammenhang mit den erreichten und/oder angestrebten Ergebnissen stehen.

⁵ unter Benutzung der GLIMMIX-Macros (vgl. Littell, Milliken, Stroup und Wolfinger, 1996, Kap. 11)

Das Bild für die BN ist komplexer. Eine gute Therapiemotivation ebenso wie eine relativ kurze Erkrankungsdauer oder geringe Zahl von Vorbehandlungen waren mit leicht erhöhten Erfolgsraten verbunden. Tendenziell niedriger waren die Erfolgsraten bei PatientInnen mit hohen Werten auf den ersten drei EDI-Subskalen, d.h. mit starker Eßstörungenproblematik. Besondere Aufmerksamkeit verdient der Einfluß der Behandlungsdauer. Wegen der Interaktionseffekte mit dem Alter und der Klinik-Kategorie, in der sich die Spezialisierungsgrad des klinischen Settings abbildet, sind Differenzierungen nötig. Der komplexe Zusammenhang zwischen Behandlungsdauer, Alter und Klinik-Kategorie ist in Abbildung 5 veranschaulicht:

Beispiel 1: Eine PatientIn im Alter von 24,6 (Median der BN), die in einer Klinik mit spezifischem Eßstörungenprogramm ,lange, behandelt wurde, hat eine um xx-fach geringere Erfolgsaussicht als eine gleichaltrige BN-PatientIn in einer vergleichbaren Klinik mit ,kurzer, Behandlung. Für zwei PatientInnen im Alter von 29,6 Jahren unterscheiden sich unter sonst gleichen Bedingungen wie oben die Erfolgsaussichten um das xx-fache.

Beispiel 2: Eine PatientIn im Alter von 24,6 (Median der BN), die in einer speziellen Eßstörungsklinik, lange, behandelt wurde, hat eine um xx-fach höhere Erfolgsaussicht als eine gleichaltrige BN-PatientIn in einer vergleichbaren Klinik mit ,kurzer, Behandlung. Für zwei PatientInnen im Alter von 29,6 Jahren unterscheiden sich unter sonst gleichen Bedingungen wie oben die Erfolgsaussichten um das xx-fache.

Tabelle 6 einfügen

Gilt das noch ??? Insgesamt war die Anpassungsgüte für die beiden Modelle bescheiden. Das Modell für die AN führte zwar zu einer Varianzaufklärung von 42%, nach Korrektur durch ein Bootstrap-Verfahren reduzierte sich diese jedoch auf 12%. Für das Modell der BN ergab sich ein Anteil erklärter Varianz von 18%. Es erwies sich jedoch unter der Bootstrap-Korrektur, nach der nur 4% übrigblieben, als wenig stabil.

2.6 Exploration des mittelfristigen Verlaufes von Eßstörungen

Methodik

Das Projekt erlaubte die Beobachtung von PatientInnen mit AN oder BN über einen Zeitraum von 2,5 Jahren. Für die Exploration dieses mittelfristigen Verlaufes von AN und BN war das Longitudinal Interval Followup Evaluation-Interview (LIFE) eingesetzt

worden. Dieses semi-strukturierte Interview erlaubt neben der klinischen Einschätzung des aktuellen Gesundheitsstatus zum Zeitpunkt des Interviews, die retrospektive Identifikation sogenannter Änderungspunkte („change points,“) und liefert somit Daten über den zeitlichen Verlauf. Die aktuelle Anwendung im Projekt erfolgte auf der Basis eines monatlichen Zeitrasters und umfaßte die Zeit zwischen Behandlungsende und 2,5-Jahresuntersuchung.

In einem Konsensbildungsprozeß wurden gemeinsam mit den KollegInnen aus dem europäischen Forschungsverbund COST Action B6, die Konzepte Gesundung-Remission-Rückfall-Wiedererkrankung, die von Frank et al. (1991) für die Depression vorgeschlagen worden waren, für AN und BN adaptiert und für die im Projekt verfügbaren Daten operationalisiert (s. auch Kordy et al., submitted). Mit Hilfe dieser Konzepte ließen sich der individuellen Verläufe nachzeichnen und im Hinblick auf klinisch relevante überindividuelle Charakteristika explorieren. Für die Exploration wurden einfache Survival-Analysen (Kaplan-Meier) durchgeführt, die für diesen ersten auf Illustration zielenden Auswertungsschritt genügen sollten.

Ergebnisse

Das Erreichen einer partiellen Remission war das häufigste Ereignis. Etwa 30% der BN und 20% der AN gingen in diesem Zustand, d.h. mit leichter Symptomatik, in die nachstationäre Zeit. Die Entwicklung über die Zeit war dabei sehr ähnlich. In den ersten 12 Monaten nach Behandlungsende wuchs der Anteil derer, die mindestens einmal diesen Zustand der partiellen Remission erreichten, in beiden Gruppen um ca. 20%. Bis zum Ende der Beobachtungszeit erreichten sowohl bei den AN als auch BN ca. 60% mindestens einmal diesen Zustand (Abb. 6a,b).

Eine volle Remission (vereinfacht gesagt: Symptombefreiheit für mindestens 3 Monate) oder gar Gesundung (Symptombefreiheit für mindestens 12 Monate) bei Behandlungsende waren bei den AN die Ausnahme und auch bei der BN sehr selten (ca. 10% Remission, 4% Gesundung). Für die AN änderte sich das auch im Beobachtungsraum nicht wesentlich, während ca. 25% der BN zumindest einmal die volle Remission erreichten.

Wie bereits angedeutet wurde, illustrieren die Abb. 6a,b nur das Erreichen der Zustände, sagen aber nichts über die Stabilität oder weitere positive Entwicklung. So war etwa das Rückfallrisiko gerade für BN wie auch für AN erheblich. Offensichtlich beschreibt die partielle Remission einen fragilen Zustand. Innerhalb von 12 Monaten haben ca. 35% der partiell remittierten AN, danach traten Rückfälle kaum noch auf. Als recht stabil erwies

sich der Zustand der vollen Remission. Von den – relativ wenigen – die diesen Zustand erreichten, erlitt während der Beobachtungszeit nur eine PatientIn einen Rückfall.

Ähnlich zeigte sich bei der BN ein erhebliches Rückfallrisiko, allerdings verteilte sich das Risiko etwas anders über die Zeit. Auch bei den partiell remittierten BN hatten ca 30% innerhalb eines Jahres einen Rückfall erlitten. Danach nahm das Risiko zwar ab, aber nach einem weiteren Jahr war der Anteil der Rückfallfreien auf unter 60% gesunken.

Interessanterweise entwickelte sich das Rückfallrisiko für die voll remittierten BN in den ersten 9 Monaten ähnlich ungünstig wie für die partiell remittierten. Danach ging das Risiko praktisch auf Null zurück.

Diskussion

Die Behandlung von Eßstörungen bleibt eine Herausforderung. Die Ergebnisse der Studie sind zufriedenstellend, aber zugleich ernüchternd. Sie sind zufriedenstellend, insofern sie sich auf die Erfolgsraten beziehen. Zu Therapieende waren 12% (aus Sicht der Therapeuten) bzw. 11% (aus Sicht der PatientInnen) der AN weitgehend symptomfrei und diese Rate stieg zur 2,5-Jahresuntersuchung auf 36% bzw. 33% an; bei den BN waren zu Therapieende 45% bzw. 31% als weitgehend symptomfrei eingeschätzt worden, wobei diese Rate allerdings zur 2,5-Jahresuntersuchung auf 36% bzw. 22% sank. Dies zeigt, daß in der Alltagspraxis psychodynamischer Behandlungen von Eßstörungen Erfolgsraten erreicht werden, die in der Nähe der unter Studienbedingungen erzielten Ergebnisse liegen. So hatten Mitchell et al. (1996) in ihrer sorgfältigen Analyse der Therapiestudien zur Wirksamkeit von Psychotherapie bei Bulimia Nervosa Erfolgsquoten von 40%-60% zu Therapieende gefunden. Fichter et al. (1994) berichten für eine im Durchschnitt 12-wöchige verhaltenstherapeutisch orientierte stationäre Behandlung eine Rate von 54% symptomfreier BN-PatientInnen zwei Jahre nach Therapie. Für AN sind nur wenige Therapiestudien versucht worden, zudem sind deren Ergebnisse wegen unterschiedlicher Probleme - insbesondere Therapie- und/oder Therapieabbruch - nur schwer zu interpretieren (z.B. van Strien et al. 1992, vgl. auch Crisp et al. 1992). Insofern fehlt eine solide empirische Vergleichsbasis (Roth und Fonagy 1996). Kürzlich veröffentlichte Ergebnisse einer Langzeitstudie lenken den Blick auf die zeitliche Perspektive. Strober et al. (1997) berichten für adolozente anorektische PatientInnen, daß 21 % erst zwei bzw. 33% drei Jahre nach Beginn einer psychotherapeutischen Behandlung im Hinblick auf die Hauptsymptome Untergewicht und Amenorrhoe weitgehend symptomfrei waren. Solche Vergleiche sind nicht ganz unproblematisch, da Erfolg sehr unterschiedlich definiert wird.

Die Erfolgskriterien bei Strober et al. (1997) und Fichter et al. (1994) sind denen der MZ-ESS insofern ähnlich, als sie ebenfalls die weitgehende Symptombefreiheit zum Maßstab wählen, auch wenn dies unterschiedlich operationalisiert wurde.

Ernüchternd sind die Befunde zur Bedeutung von Dauer und Intensität in psychodynamischen stationären Behandlungen und für deren Erfolg. Die Dauer der Behandlungen variierte stark zwischen den an der Studie beteiligten Kliniken und ebenso beträchtlich innerhalb vieler dieser Kliniken. Aber nur zu einem geringen Teil lassen sich diese Unterschiede durch den klinischen oder sozio-biografischen Ausgangszustand der PatientInnen erklären. Offensichtlich spielen die lokalen klinischen Ansätze, die sich eher auf klinische Erfahrung und therapeutische Einstellung als auf empirische Studien stützen, eine größere Rolle. Insbesondere die hohe durchschnittliche Behandlungsdauer in Kliniken mit einer separaten Eßstörungenabteilung verdient Aufmerksamkeit.

Wenn nicht der Krankheitsstatus der PatientInnen die Unterschiede in der Dauer der Behandlung erklärt, inwiefern rechtfertigt der Zweck, d.h. bessere Erfolgschancen, das Mittel der längeren Therapie? Auch in dieser Hinsicht sind die Studienergebnisse ernüchternd. Lange und kurze Behandlungen unterscheiden sich kaum im Ergebnis, auch oder gerade wenn möglicher konfundierende Merkmale der PatientInnen kontrolliert werden („Propensity-Score-Korrektur“). Eingangsscharakteristika der PatientInnen tragen moderat zur Vorhersage der Therapieerfolge bei AN bei. Läßt man die Klinikmerkmale (als nicht patientenbezogene Variablen) außer Acht, so erhält man ein Modell, das 37% der Varianz erklärt (statt der berichteten 42%). Bei der Bootstrap-Korrektur bleiben davon 19% (statt 12%), was auf eine bessere Stabilität hindeutet. Nach diesem Modell erwies sich ein starkes Untergewicht erwartungsgemäß als Risikofaktor, insofern daß stark untergewichtige AN, die in der Regel zwar länger behandelt werden, nach 2.5 Jahren eher schlechtere Erfolgsraten zeigen als weniger untergewichtige Patientinnen. Hier mag der längere Aufenthalt durch das höhere Maß an notwendiger Gewichtszunahme bestimmt sein, was in aller Regel auch mehr Zeit beansprucht. In die gleiche Richtung weisen die Ergebnisse für den Einfluß der Erkrankungsdauer bzw. der Anzahl der Vorbehandlungen, die sich beide als Indikatoren des Chronifizierungsgrades auffassen lassen.

Die Suche nach prognostisch positiven oder negativen Faktoren für die Behandlung der BN erbrachte ein nur wenig ergiebiges Modell (15% erklärter Varianz, nach Bootstrap 2%). Neben einem leichten positiven Effekt der Behandlungsdauer zeigten sich Therapiemotivation als prognostisch positiver und die oben genannten Indikatoren für die

Chronifizierung als prognostisch negative Faktoren. Ein möglicherweise interessanter Interaktionseffekt deutet an, daß Kliniken mit einer eßstörungsspezifischen Abteilung im Vergleich zu anderen Kliniken eher für BN als für AN bessere Erfolgsaussichten bieten.

Die Teilnahme einer so großen Zahl an Kliniken aus dem gesamten Bundesgebiet ist auf der einen Seite ein Vorteil, begrenzt auf der anderen Seite jedoch auch wichtige Schlußfolgerungen. So unterschieden sich beispielsweise die Kliniken beträchtlich in der Zusammensetzung ihrer Klientele. Sich daraus allein oder in Zusammenhang mit klinikspezifischen Settingvarianten ergebende mögliche Verzerrungen der Effekte der Behandlungsdauer konnten nicht mit Hilfe der Propensity-Score-Technik korrigiert werden, da einige Kliniken überwiegend lange, andere überwiegend kurze Behandlungen durchführten.

Für Schlußfolgerungen über die adäquate Dauer der Behandlung von Eßstörungen ist besonders zu beachten, daß die berichteten Studienergebnisse nur für das hier untersuchte Spektrum von Behandlungen gelten, d.h. für solche mit einer Dauer von 5 bis etwa 16 Wochen. Kürzere Behandlungen waren von den Analysen ausgeschlossen worden, da fraglich erschien, inwieweit diese Dauern beabsichtigt waren bzw. sich durch irreguläre Beendigung ergeben hatten. Behandlungen von mehr als 16 Wochen Dauer waren zu selten.

Dennoch bleibt ein großer Spielraum für weitere Einflüsse, die im bisherigen Modell nicht enthalten sind (z.B. hier nicht erhobene Klinikeigenschaften, Parameter zum Therapieverlauf oder des Behandlungsprogramms oder Ereignisse im Katamnese-Zeitraum).

Aufgrund der vorliegenden Verteilung der Behandlungsdauern läßt sich dieses Ergebnis am zuverlässigsten auf Behandlungen zwischen 6 und 15 Wochen Länge anwenden. Kürzere oder längere Dauern waren zu selten, so daß zuverlässige Aussagen für diese Behandlungslängen nicht möglich sind. Nicht ausgeschlossen werden kann auch der Einfluß von einzelnen Klinik-Settings (weil nicht in der Propensity-Korrektur enthalten) und von verschiedenen häufig oder intensiven Behandlungen im Anschluß an die hier untersuchte stationäre Behandlung.