

**Die Konzeptualisierung und Validierung von
Therapie-Response und seine Bedeutung für die
Behandlungsplanung in einem verlaufsorientierten
Qualitätsmanagement am Beispiel von Eßstörungen**

Inauguraldissertation
zur Erlangung des Grades eines Dr. phil.
an der Fakultät für Sozial- und Verhaltenswissenschaften
der Universität Heidelberg

von
Matthias Richard
2001

Gutachter:
Prof. Dr. Peter Fiedler (Berater)
Prof. Dr. Horst Kächele

Datum der Prüfung: 2. Februar 2001

Danksagung:

An dieser Stelle möchte ich allen Personen herzlich danken, die mich im Verlauf dieser Arbeit durch Anregungen, Kommentare und auf andere Art und Weise unterstützt haben:

Ich danke Prof. Dr. Peter Fiedler für die unkomplizierte Beratung und Betreuung dieser Arbeit.

Mein besonderer Dank gilt Dr. Hans Kordy, von dem ich die Anregung zu diesem Thema erhielt; darüber hinaus danke ich für vielfältige und wertvolle Diskussionen zum Thema Qualitätsmanagement und longitudinaler Forschung während meiner Tätigkeit an der Forschungsstelle für Psychotherapie, Stuttgart. Ich bedanke mich bei Prof. Dr. Kächele und dem geschäftsführenden Ausschuß der multizentrischen Studie, MZ-ESS, für die Überlassung der Daten.

Prof. Judith Singer (Harvard University, Boston) und Ph.D. Zoran Martinovich (Northwestern University, Chicago) danke ich für ihre methodischen Bemerkungen und Vorschläge zur Untersuchung von longitudinalen Fragestellungen.

Prof. Michael Lambert (Brigham Young University, Provo) danke ich dafür, daß ich während meines Gastaufenthaltes an der Brigham Young University einen Einblick in die Art und Weise bekam, wie verlaufsorientiertes Qualitätsmanagement in der Praxis der Psychotherapie funktionieren kann.

Sylvi, vielen Dank für Dein 'spell checking'

Mein Dank gilt am meisten meiner Frau, Dipl.-Psych. Marion Schowalter, die mir genügend Raum gab, um zu Hause arbeiten zu können und noch viel mehr für den Ausgleich in unserer gemeinsamen Zeit!

Inhaltsverzeichnis

1 Überblick	5
TEIL I : DIE KONZEPTUALISIERUNG VON THERAPIE-RESPONSE	7
2 Argumente für eine stärkere Verlaufsorientierung bei der Therapieplanung	7
2.1 Qualitätsmanagement in der Psychotherapie	7
2.1.1 Grundelemente eines Qualitätsmanagement-Systems am Beispiel des Stuttgart - Heidelberger Modells	8
2.1.2 Die Umsetzung der Informationen aus dem QM	9
2.2 Indikation	12
2.2.1 Differentielle und störungsspezifische Indikation	12
2.2.2 Adaptive Indikation	14
3 Das Konzept des Therapie-Response	16
3.1 Explikation des Begriffs ‘Therapie-Response’	16
3.2 Therapie-Response in schulenübergreifenden Psychotherapiemodellen .	18
3.2.1 Das Modell der Systematischen Behandlungsplanung	19
3.2.2 Das Allgemeine Modell der Psychotherapie / Phasenmodell . . .	21
3.2.3 Dreikomponentenmodell zur Wirkungsweise von Psychotherapie	23
3.2.4 Das Transtheoretische Modell der Psychotherapie	25
3.3 Empirische Hinweise für die prädiktive Validität von Therapie-Response	27
3.3.1 Studie von Tang & DeRubeis (1999)	28
3.3.2 Studie von Quitkin et al. (1996)	28
3.3.3 Studie von Breslin et al. (1997)	29
3.3.4 Studie von Olmsted et al. (1996)	30
4 Gliederung des empirischen Teils	31
4.1 Fragenblock zur prädiktiven Validität von Therapie-Response	31
4.2 Fragenblock zur Anwendung von Therapie-Response in einem QM-System	33
4.3 Auswertungsstrategien - Methodik	34

TEIL II : DIE VALIDIERUNG VON THERAPIE-RESPONSE	36
5 Beschreibung der Studie	36
5.1 Versuchsplan und Stichprobe	36
5.1.1 Versuchsplan	36
5.1.2 Erhebungsinstrumente	37
5.1.3 Stichprobenbeschreibung	37
6 Die Erfassung von Therapie-Response und Therapieerfolg	39
6.1 Die Symptombereiche	40
6.1.1 Anorexie	40
6.1.2 Bulimie	41
6.2 Die Erfassung von Therapieerfolg	42
6.3 Erfassung von Veränderungen der Symptomschwere während der Therapie	43
6.4 Ausgewählte Operationalisierungen des Therapie-Response	45
6.4.1 Verbessertes Wohlbefinden	46
6.4.2 Moderate Symptomatik (Cutoff)	47
6.4.3 Verringerung der Symptome ('In-Gang-Kommen' der Therapie)	48
7 Ergebnisse	50
7.1 Häufigkeiten von Therapie-Response	50
7.2 Therapie-Response und Therapieerfolg	52
7.2.1 Ohne Kontrolle der Eingangsinformationen	52
7.2.2 Mit Kontrolle der Eingangsinformationen	55
8 Exkurs I: Kann Therapie-Response vorhergesagt werden?	65
9 Exkurs II: Therapie-Response und längerfristiger Erfolg	69
10 Therapie-Response und Therapieverlauf	73
10.1 Der Verlauf auf Gruppenebene	74
10.2 Der Verlauf auf individueller Ebene	78

TEIL III : THERAPIE-RESPONSE INNERHALB EINES VERLAUFSORIENTIERTEN QM	79
11 Verlaufsorientiertes Qualitätsmanagement	79
11.1 Anforderungen an ein verlaufsorientiertes QM-System	79
11.1.1 Erfassung der Therapieergebnisse auf individueller Ebene	80
11.1.2 Bewertung der Ergebnisse auf dem Hintergrund angemessener Standards	80
11.1.3 Rückmeldung der Ergebnisse an die behandelnden Personen	81
11.1.4 Leichte Integrierbarkeit des Systems in den therapeutischen Alltag	82
11.1.5 Indikatoren zur Prognose des weiteren Verlaufes	82
11.2 Kritische Würdigung des COMPASS QM-Systems	83
12 Illustration der Anwendung von Therapie-Response im QM	87
12.1 Grundidee und Methodik	87
12.1.1 Grundidee	87
12.1.2 Methodik	88
12.2 Konstruktion der Referenzmodelle	91
12.2.1 Festlegung der Verlaufsformen der 4 Referenzmodelle	91
12.2.2 Vorhersage des Verlaufs aufgrund der Anfangsinformation	94
12.3 Vergleich von 'neuen' Verläufen mit den Referenzmodellen	96
12.3.1 Beispiel 1: Responder mit Therapieerfolg	96
12.3.2 Beispiel 2: Non-Responder mit Therapieerfolg	99
12.3.3 Beispiel 3: Responder ohne Therapieerfolg	100
12.3.4 Beispiel 4: Non-Responder ohne Therapieerfolg	102
12.3.5 Beispiel 5: unklare Klassifizierung	103
13 Die Bandbreite positiver Verläufe als Prognose-Indikator	104
13.1 Der Prognose-Indikator als zu überprüfende Hypothese	104
13.2 Vorgehen und Ergebnisse der Analysen mit dem Prognose-Indikator	105
13.2.1 Vorgehen	105
13.2.2 Ergebnisse	106
13.3 Anwendungsbeispiele	111

14 Diskussion und Ausblick	113
14.1 Diskussion	113
14.1.1 Der Zusammenhang zwischen Therapie-Response und Therapie- erfolg	113
14.1.2 Zeitpunkt des Therapie-Response:	114
14.1.3 Response und weiterer Verlauf der Symptomatik:	115
14.1.4 Zur Methodik der Konstruktion von Referenzverläufen	115
14.1.5 Prognose-Indikatoren	116
14.1.6 Ist ein verlauforientiertes QM durchführbar?	117
14.2 Zusammenfassung der wichtigsten Ergebnisse	118
14.2.1 Konzeptualisierung und Operationalisierung von Therapie-Response	118
14.2.2 Prädiktive Validität von Therapie-Response	118
14.2.3 Vorhersage von Therapie-Response	119
14.2.4 Therapie-Response und Therapieverlauf	120
14.2.5 Eigenschaften eines verlaufsorientierten QM-Systems	120
14.2.6 Konstruktion von Verlaufsstandards	121
14.2.7 Verbesserte Abschätzung der Erfolgswahrscheinlichkeit durch Prognose- Indikatoren	121
14.3 Ausblick	122
Literatur	124
15 Anhang	131

1 Überblick

Diese Arbeit unterbreitet den Vorschlag, bei der Planung und routinemäßigen Evaluati-
on psychotherapeutischer Behandlungen innerhalb eines Qualitätsmanagements (QM)
die Entwicklung des Therapieverlaufs systematischer mit einzubeziehen: Die Abwägung
der Erfolgchancen der Patienten kann mit der Hinzunahme des Konzeptes ‘Anspre-
chens auf Therapie’ bzw. Therapie-Response¹ weiter präzisiert werden, so daß Ent-
scheidungen für den Fortgang der Therapie auf eine sicherere Basis gestellt werden
können.

Die Arbeit gliedert sich in drei Teile: Teil I erläutert das Konzept des Therapie-
Response und legt die Zielsetzung der eigenen empirischen Arbeit dar. In Abschnitt
2 wird beschrieben, welcher potentielle Gewinn in einem verstärkt verlaufsorientierten
Qualitätsmanagement liegt und skizziert die Berührungspunkte zur adaptiven Indika-
tion in der Psychotherapie. Abschnitt 3 expliziert den Begriff des Therapie-Response,
ordnet ihn in schulenübergreifende Psychotherapiemodelle ein und beschreibt einige
empirische Untersuchungen zu Therapie-Response unter der Perspektive der prädik-
tiven Validität. In Abschnitt 4 werden die Fragestellungen und Vorgehensweisen der
eigenen empirischen Untersuchung vorgestellt.

Teil II berichtet die Ergebnisse der Validierung des in Teil I beschriebenen Be-
griffs ‘Therapie-Response’. Abschnitt 5 beschreibt Hintergrund und Design der Stu-
die, Abschnitt 6 drei verschiedene Operationalisierungen von Therapie-Response und
die Operationalisierung von Therapieerfolg. Anschließend wird der Zusammenhang von
Therapie-Response zu verschiedenen wichtigen ‘Psychotherapie-Parametern’ berichtet:
zum einem zum Therapieerfolg am Ende der Behandlung (Abschnitt 7), zum ande-
ren zu den verschiedenen Eingangscharakteristika der Patientinnen (Abschnitt 8) und
schließlich zum längerfristigen Therapieerfolg nach 2.5 Jahren (Abschnitt 9). Abschlie-
ßend wird Therapie-Response hinsichtlich seines Zusammenhangs zum Therapieverlauf
untersucht (Abschnitt 10).

Teil III demonstriert die Anwendung von Therapie-Response innerhalb eines verlauf-
orientierten QM. In Abschnitt 11 werden die Anforderungen an ein solches System
beschrieben und ein bereits bestehendes System anhand dieser Kriterien bewertet. Ab-
schnitt 12 beschreibt die Grundidee und die Methodik eines solchen Systems und zeigt
anhand von Beispielen, wie Therapie-Response zur Bewertung des bisherigen Verlaufs

¹Die Begriffe ‘Ansprechen auf Therapie’ und ‘Therapie-Response’ werden hier und im weiteren
synonym gebraucht. Der englische Begriff ‘Response’ fügt sich einfacher in den Text ein als der Begriff
des ‘Ansprechens’, der ohne den erläuternden Zusatz ‘auf Therapie’ m.E. zu vage bleibt. In beiden
Fällen ist damit immer die (meist symptomatische) Entwicklung des Patienten während der Therapie
gemeint. Der Genus des Begriffs ist maskulin: ‘der Therapie-Response’.

eingesetzt werden kann. Schließlich werden in Abschnitt 13 aus diesem System Prognoseindikatoren abgeleitet, um sicherere Rückschlüsse auf den Therapieerfolg abzuleiten. Die komprimierte Zusammenfassung der Ergebnisse dieser Arbeit, sowie Ausblicke für weitere Untersuchungen finden sich in Abschnitt 14.

TEIL I:

DIE KONZEPTUALISIERUNG VON THERAPIE-RESPONSE

2 Argumente für eine stärkere Verlauforientierung bei der Therapieplanung

Dieser Abschnitt erläutert die Hintergründe, die zur Beschäftigung mit dem Konzept des Therapie-Response geführt haben. Der Ausgangspunkt ist dabei die aktuelle Diskussion und eigene Erfahrungen zur Qualitätssicherung (QS) bzw. zum Qualitätsmanagement (QM) in der psychotherapeutischen Versorgung, sowie damit verbundene Fragen der Zuweisung von Patienten zu geeigneten Behandlungen. Es wird deutlich werden, daß eine stärkere Verlauforientierung (im Gegensatz zu einer statischen Zustandsorientierung) bei der routinemäßigen Evaluation von Psychotherapien innerhalb des Qualitätsmanagements Vorteile in sich birgt, von denen sowohl Patienten als auch Therapeuten profitieren können. Dabei wird ein bisher nur spärlich untersuchtes Feld der Psychotherapie berührt, nämlich die Anpassung des therapeutischen Vorgehens an die Veränderungen, die sich während der Behandlung ergeben (adaptive Indikation). Ein Qualitätsmanagement, das stärker als bisher den Verlauf einer Therapie berücksichtigt, hat eine wichtige Aufgabe mit der adaptiven Indikation gemeinsam: die Notwendigkeit, den weiteren Therapieverlauf auf der Grundlage der bisherigen Entwicklung zu prognostizieren. Dieser Schnittpunkt zwischen verlauforientiertem Qualitätsmanagement und adaptiver Indikation wird im folgenden Abschnitt beleuchtet.

2.1 Qualitätsmanagement in der Psychotherapie

Das Gesundheitswesen in Deutschland befindet sich gegenwärtig im Wandel. Um die explodierenden Kosten der Gesundheitsversorgung in den Griff zu bekommen, steigen die Bemühungen, neue Organisationsformen und -strukturen zu etablieren, welche den Anforderungen der Zukunft (hohe Effektivität der Versorgung bei gleichzeitiger Wirtschaftlichkeit) besser gewachsen sind als die bestehenden. Die Einführung der Pflegeversicherung, freie Wahl des Krankenversicherungsträgers, aber auch die Kürzungen der Ausgaben im Bereich Rehabilitation sind einige Beispiele hierfür. Der Druck, die knapper werdenden Mittel ohne Einbußen in der Qualität der Versorgung zu verteilen, zwingt zur Suche nach Verbesserungsmöglichkeiten innerhalb dieses komplexen Systems. Im Bereich der psychotherapeutischen Versorgung findet dies in der Einführung der Pflicht zur Qualitätssicherung (§137 Sozialgesetzbuch) im Rahmen des Gesundheitsreformgesetzes seinen Niederschlag. So hat sich 'Qualitätssicherung' inzwischen

von einem Reiz- und Modewort (Kächele & Kordy 1992) zu einem festen und ernstgenommenen Bestandteil der Versorgung entwickelt – sowohl bei den Versorgungsanbietern, als auch im Bereich der klinischen Evaluations- und Psychotherapieforschung. Eine ganze Reihe von Entwürfen, was unter QM in der Psychotherapie zu verstehen ist und wie es durchgeführt werden sollte, sind inzwischen entwickelt und untersucht worden (Laireiter & Vogel 1998, Kordy & Lutz 1995, Kordy et al. 1999, Grawe & Braun 1994, Fiegenbaum, Tuschen & Florin 1997). In den USA gibt es aufgrund der anderen Struktur des Gesundheitssystems ('managed care') schon mehr Erfahrungen im Umgang mit 'qualitätssichernden' Maßnahmen (Lyons, Howard, O'Mahoney & Lish 1997, Howard, Brill, Lueger, O'Mahoney & Grant 1993), wobei jedoch oft der kostensparende Aspekt sehr im Vordergrund steht (siehe Abschnitt 11). In neuester Zeit wird aber auch hier ein Trend in Richtung Ergebnisorientierung sichtbar (Sperry 1997 und weitere Beiträge im dortigen Heft).

2.1.1 Grundelemente eines Qualitätsmanagement-Systems am Beispiel des Stuttgart - Heidelberger Modells

Allgemein lassen sich vom Begriff der 'Qualität' 3 Perspektiven abheben: Struktur-, Prozeß- und Ergebnisqualität (Donabedian 1982). Die wichtigste Komponente der 'Qualität' von Psychotherapie stellt sicherlich die Ergebnisqualität dar. Eine ausreichende Anzahl an Personal oder gute Ausstattung mit diagnostischen Instrumenten (Aspekte der Strukturqualität) und geringe Wartezeiten oder schnelle Kommunikation zwischen den Behandlungspartnern (Aspekte der Prozeßqualität) alleine garantieren kein positives Behandlungsergebnis. Die Qualität der Behandlung insgesamt muß sich an der Ergebnisqualität, sprich dem Therapieerfolg, messen lassen. Folglich ist die Erhebung der Ergebnisse notwendig. Will man diese Information nutzen, um daraus Maßnahmen zur Verbesserung abzuleiten, müssen diese Ergebnisse anhand von Standards bewertet und in leicht verständlicher Form an die Beteiligten zurückgemeldet werden (vgl. Kordy & Lutz 1995).

Ein System mit dem Anspruch, bei der Sicherung der Behandlungsqualität von Psychotherapie einen Beitrag zu leisten, ist durch folgende Eigenschaften gekennzeichnet (vgl. Lyons et al. 1997, Lutz 1997):

- Erfassung der Therapieergebnisse auf individueller Ebene mit Hilfe von standardisierten psychometrischen Instrumenten zu mindestens zwei Meßzeitpunkten (prä und post)
- Bewertung der Ergebnisse auf dem Hintergrund von angemessenen Standards
- Rückmeldung der Ergebnisse an die behandelnden Personen

- leichte Integrierbarkeit des Systems in den therapeutischen Alltag

Die genannten Punkte werden in Abschnitt 11.1 noch vertiefender dargestellt, wenn es darum geht, was diese Kriterien für ein verlaufsorientiertes QM bedeuten. Das Stuttgart-Heidelberger Modell der aktiven internen Qualitätssicherung kommt den gerade aufgezählten Anforderungen im Wesentlichen nach und kann daher als Beispiel einer gelungenen Umsetzung von QM dienen:

Bei Beginn und am Ende der Behandlung werden Patienten und Therapeuten Fragebögen ausgehändigt, die den Zustand des jeweiligen Patienten beschreiben. Auf Seiten der Patienten werden hierzu u.a. normierte psychometrische Instrumente eingesetzt, die die Beschwerden der Patienten auf körperlicher, psychischer und interpersonaler Ebene erfassen (Gießener Beschwerdebogen, Symptomcheckliste SCL-90-R, Inventar Interpersoneller Probleme). Nach Abschluß der Behandlung werden diese Fragebögen eingesammelt und in ein Computersystem eingegeben. Dieses bewertet in einem von Klinikern und Forschern gemeinsam entwickelten Algorithmus die einzelnen Veränderungen und setzt sie dabei in Beziehung zu den Populationsparametern von Nicht-Patienten. Die automatische Bewertung des Computers der einzelnen Patienten wird in sog. 'Qualitätszirkeln' an das therapeutische Team zurückgemeldet. Dabei werden die Informationen des Computers mit dem klinischen Eindruck verknüpft und es besteht nochmals die Möglichkeit, die Bewertung des Computers zu korrigieren. Alle Informationen werden in Datenbanken anonymisiert gespeichert und stehen weiteren Auswertungen zur Verfügung (vgl. im einzelnen: Kordy & Lutz 1995, Lutz, Stammer, Leeb, Dötsch, Bölle & Kordy 1996, Lutz 1997, Kordy & Hannover 1998, Kordy et al. 1999).

2.1.2 Die Umsetzung der Informationen aus dem QM

Die Nutzung der gesammelten Information geschieht in der Regel auf 2 Ebenen: zum einen auf der Ebene des einzelnen Patienten und zum anderen auf einer aggregierten Ebene. Auf der Ebene des einzelnen Patienten werden die in der Behandlung eingetretenen Veränderungen bewertet, zurückgemeldet und im Team (beispielsweise in Form eines Qualitätszirkels) reflektiert. Die Ergebnisse dieses Prozesses können daraufhin wieder in die Behandlung einfließen. Auf der aggregierten Ebene werden die Ergebnisse der einzelnen Patienten zusammengefaßt. Das ermöglicht die Erstellung von Vergleichen verschiedener Jahrgänge oder bestimmter Patientengruppen hinsichtlich der Erfolgsraten oder anderer Parameter (z.B. Patientenzufriedenheit). Auf dieser Ebene können mit Hilfe statistischer Methoden u.U. Gruppen identifiziert werden, die vergleichsweise schlecht abschneiden. Daraufhin lassen sich entsprechende Interventionen treffen, um dieser Gruppe eine bessere Behandlung zu ermöglichen (z.B. durch gezielte

Weiterbildung oder Einrichtung eines speziellen Therapiebausteins für diese Gruppe). Der Nachteil eines solchen Vorgehens besteht jedoch in 2 Punkten:

1. Für solche Vergleiche werden relativ große Stichproben benötigt und daher liegen die Resultate erst dann vor, wenn genügend Information gesammelt worden ist.
2. Aufgrund der nicht-experimentellen Datenerhebung bleibt unklar, ob die gefundenen Unterschiede zwischen den Gruppen tatsächlich auf Behandlungseffekte oder evtl. auf andere nicht erhobene Einflüsse zurückzuführen sind (vgl. Sperry, Brill, Howard & Grissom 1996).

Nach wie vor ist unklar, ob und wie diese Systeme tatsächlich zur Verbesserung der Behandlungsqualität beitragen. Entsprechende Untersuchungen sind noch nicht durchgeführt worden bzw. werden gerade erst durchgeführt (Lambert, Vermeersch & Hansen 1999). In der Beurteilung des Stuttgart-Heidelberg Systems (Lutz 1997) schätzten die Therapeuten die Praxistauglichkeit und Verständlichkeit der individuellen Ergebnisrückmeldungen deutlich positiv ein, ebenso empfanden sie die Teilnahme am QM als motivierend. Ein Großteil der am QM beteiligten Therapeuten konnten jedoch kaum eine Wirkung der Rückmeldungen über die Therapieergebnisse auf ihr therapeutisches Vorgehen erkennen. Dies ist höchstwahrscheinlich darauf zurückzuführen, daß zum Zeitpunkt der Rückmeldung die Behandlung i.d.R. schon beendet ist. Folglich können eventuelle Änderungen im therapeutischen Vorgehen diesem speziellen Patienten nicht mehr zugute kommen und es ist schwierig, das Resultat auf andere Patienten zu übertragen. An dieser Stelle bedarf das Stuttgart-Heidelberger Modell m.E. noch der Entwicklung eines Vorgehens, wie behandlungsrelevante Ergebnisse aus den Qualitätszirkeln systematisch gesammelt und einer Umsetzung näher gebracht werden können. Dieser Punkt macht nochmals unter anderer Perspektive das Problem deutlich, daß es u.U. lange dauert, bis Patienten von den Erkenntnissen aus dem QM profitieren können. In der gleichen Untersuchung stellte sich ebenfalls heraus, daß für den Großteil der Therapeuten das QM mit einem spürbaren Mehraufwand an Arbeit verbunden war. Angesichts der gleichzeitig berichteten relativ geringen Auswirkungen auf die tägliche Arbeit, deutet sich hier aus der Sicht der Therapeuten eine Diskrepanz zwischen Aufwand und Ertrag an.

Die Rückmeldung der Ergebnisse ist jedoch dann am meisten relevant, wenn die Information möglichst aktuell und noch auf den Patienten selbst anwendbar ist. Die zentrale Frage, die sich während der Behandlung für die Therapeuten stellt, besteht dann darin: 'Woran läßt sich erkennen, ob sich die Therapie erfolgreich entwickelt oder nicht? Welche Anzeichen zeigen zu welchem Zeitpunkt, daß sie *nicht* wie gewünscht verläuft und daher das Vorgehen neu reflektiert werden sollte?' Diese Fragen können nur dann beantwortet werden, wenn *im Verlauf* der Therapie zusätzliche Erhebungen und

Rückmeldungen vorgenommen werden. Allerdings ist dann nicht sofort offensichtlich, was die Bewertung des (Zwischen-)Ergebnisses anhand der für den Abschluß einer Therapie entwickelten Erfolgskriterien bedeuten. Verlaufsbezogene Standards, welche Art von Veränderung zu welchem Zeitpunkt zu erwarten wäre, sind meines Wissens für Psychotherapie noch nicht entwickelt. Die Implikationen für die Organisation und Auswertung der Informationen eines solchen *an einzelnen Verläufen orientierten QM* sind weitreichend und werfen unter den o.g. Anforderungen an ein QM-System neue Fragen auf: Welche Bewertungsstandards sind für eine *Verlaufsdiagnostik* relevant? Wie können verlaufsorientierte Standards gewonnen werden, aus denen ersichtlich wird, zu welchem Zeitpunkt welcher Fortschritt eingetreten sein sollte? Welche Eigenschaften müssen verlaufsorientierte Meßinstrumente aufweisen? Welche Auswirkungen hat eine Verlaufsorientierung auf die Rückmeldestrukturen? Diese und weitere Fragen werden in Abschnitt 11.1 diskutiert.

Erste Erfahrungen mit der eßstörungsspezifischen Version des Stuttgart-Heidelberger Systems zeigen, daß sowohl Patienten als auch die weiterbehandelnden Therapeuten die Zusammenfassung der Ergebnisse aus der (bereits abgeschlossenen) stationären Behandlung als hilfreich empfinden (Kordy et al. 1999). Das Stuttgart-Heidelberger Modell erlaubt Mehrfacherhebungen im Verlauf der Behandlung und die Bewertung der Ergebnisse zwischen zwei beliebigen Meßzeitpunkten; allerdings erfolgt die Bewertung anhand der Standards, die für den Entlassungszeitpunkt entwickelt wurden. Das COMPASS QM-System von (Howard et al. 1993) bietet die bisher deutlichste Verlaufsorientierung: Entsprechend den Ergebnissen aus den Phasenmodell der Psychotherapie (Howard, Lueger, Maling & Martinovich 1993, vgl. Abschnitt 3.2.2) wird der momentane Zustand des Patienten zu mehreren Zeitpunkten während der Therapie erhoben. Dieser wird mit dem üblicherweise zu erwartenden Verlauf von Patienten mit ähnlichen Eingangsbedingungen verglichen und ggf. an die Therapeuten zurückgemeldet. In Abschnitt 11.2 wird dieser Ansatz dargestellt und kritisch gewürdigt.

Zusammenfassung:

In den letzten Jahren sind eine Reihe von QM-Systemen für Psychotherapie entwickelt worden, deren tatsächlicher Beitrag zur Verbesserung der Behandlungsergebnisse bisher noch nicht ausreichend untersucht worden ist. Die Prinzipien der Auswertung der in solchen Systemen gesammelten Informationen (Bewertung der Ergebnisse *nach* Abschluß der Therapie, Datenanalyse auf aggregiertem gruppenstatistischem Niveau) erfordern jedoch eine lange Zeitspanne, bis deren Ergebnisse in konkrete Maßnahmen zur Verbesserung der Behandlung umgesetzt werden können. Einen direkteren Einfluß in Bezug auf die Verbesserung der Behandlung kann dagegen von verlaufsorientierten Systemen erwartet werden, in denen der Verlauf eines Patienten kontinuierlich erhoben und bewertet wird (Einzelverlaufsorientierung). Dadurch können die erhobenen

spezifischen Informationen und deren Implikationen für die Behandlung dem gleichen Patienten noch *während* der Behandlung wieder zugute kommen. Ein *verlaufsorientiertes QM-System* wird voraussichtlich deutliche Vorteile gegenüber bisherigen QM-Systemen aufweisen; gleichzeitig erfordert es die Weiterentwicklung von geeigneten Standards und Instrumenten zur Bewertung von Therapieverläufen.

2.2 Indikation

2.2.1 Differentielle und störungsspezifische Indikation

Vor dem Hintergrund der Qualitätssicherung kommen auch den Untersuchungen zur differentiellen und störungsspezifischen Indikation (Bastine 1992, Fiedler 1997b) aus der Psychotherapieforschung eine besondere Bedeutung zu. Diese gehen der oft zitierten Frage nach, wie sie von Paul (1969) erstmalig formuliert wurde:

‘Welche Behandlung, von welchem Behandler ausgeführt, ist für diese Person und seine spezielle Problematik unter welchen Bedingungen am wirksamsten - und wieso?’

Das Ziel liegt dabei u.a. in der Verbesserung der psychotherapeutischen Versorgung durch eine *gezieltere Zuweisung* der Patienten zu den für sie am erfolgversprechendsten Therapien. Der Grundgedanke liegt darin, durch eine möglichst optimale Zuweisung der Patienten zu den geeignetsten Therapien, die psychischen Störungen langfristig wirksam zu behandeln und den Patienten lange Irrwege durch das Gesundheitssystem zu ersparen. Dies verminderte einerseits die Inanspruchnahme weniger wirksamer therapeutischer Leistungen als auch die Häufigkeit von Rückfällen und den deshalb notwendig werdenden erneuten Behandlungen. Überprüft man die beeindruckende Vielzahl von Ergebnissen auf ihre praktische Anwendbarkeit, so stellt sich jedoch – auch nach einer Zusammenschau der Ergebnisse – folgendes heraus:

1. Lambert & Bergin (1994) weisen in ihrem Überblick über die Ergebnisforschung darauf hin, daß die im Durchschnitt zwar positiven Therapieeffekte nicht über die große Variabilität der Therapieergebnisse hinwegtäuschen dürfen. Nicht alle Patienten profitieren von Psychotherapie, ein gewisser Teil zeigt auch negative Therapieeffekte (vgl. auch Mohr 1995).
2. Die gefundenen Zusammenhänge zwischen den Eingangsbedingungen der Patienten und dem Therapieergebnis sind nicht deutlich genug ausgeprägt, um für den *Einzelfall* eine hinreichend gute Entscheidung für die eine oder andere Therapie zu treffen. Sie erklären meist nur einen geringen Teil der Unterschiede im Ergebnis: auch wenn innerhalb einer Untersuchung deutliche Unterschiede zwischen zwei Gruppen gefunden werden, gibt es doch Patienten der erfolgreicheren

Gruppe, die sich nicht verbessern und umgekehrt gibt es in der weniger erfolgreichen Gruppe auch Patienten, die sich genauso viel verbessern wie Patienten der erfolgreicheren Gruppe. Mit anderen Worten: die Ergebnis-Verteilungen beider Gruppen überlappen sich (Kächele & Kordy 1996a).

3. Außerdem sind die Forschungsergebnisse bisher wenig eindeutig oder noch nicht ausreichend gesichert: oft fehlen Replikationen, welche die Zusammenhänge bestätigen oder in einer Anzahl von Replikationen finden sich gleichviel bestätigende und widersprechende Befunde.

Die verbesserte Diagnostik und das vermehrte Wissen über die Ätiologie verschiedener Störungsbilder haben dazu geführt, daß inzwischen eine Vielzahl von empirisch untersuchten Therapieansätzen für spezielle psychische Störungen vorliegen (Reinecker & Fiedler 1997, Fiedler 1997b, Fiedler 1997a, Margraf 1996). Mit der Zuweisung von Patienten zu den genau auf ihre Störung zugeschnittenen Behandlungen (entsprechend guter Diagnostik im Alltag vorausgesetzt) ist ein ähnlicher Effekt der insgesamt besser eingesetzten und gezielter in Anspruch genommenen Therapien zu erwarten.

Allerdings verhindern pragmatische Gründe solch eine gezielte Zuweisung: Kächele & Kordy (1996b) beschreiben den Prozeß der Indikation als einen sozialen Aushandlungsprozeß, der versucht, unter den gegebenen Rahmenbedingungen im Versorgungssystem (Wartezeiten, Entfernung zum Therapeuten, Kostenerstattung der Krankenkassen usw.) und dem medizinisch-therapeutisch Notwendigen einen optimalen Kompromiß zu finden. D.h. selbst wenn die Forschungsergebnisse eine gezielte Zuweisung zu entsprechenden Therapeuten zuließen, gäbe es noch eine Menge weiterer 'versorgungstechnischer' Gründe, die letztlich den Ausschlag geben, in welcher Therapie und bei welchem Therapeuten der Patient schließlich eine Behandlung beginnen wird.

Zusätzlich entsteht durch die fast ausschließlich schulengebundene Ausbildung der Psychotherapeuten der Effekt, daß mit der Entscheidung für einen bestimmten Therapeuten in den meisten Fällen die prinzipielle Behandlungsform entschieden ist. Nur bei explizit eklektisch arbeitenden Therapeuten wird der Patient die Chance haben, eine den Erfordernissen seiner individuellen Eingangssituation angepaßte Behandlung zu erhalten. Häufig arbeiten die Therapeuten mit den schulenspezifischen Methoden ihrer Ausbildung und passen diese den Erfordernissen der Patienten an – dabei bleibt jedoch von vornherein eine Vielzahl anderer und u.U. für diesen Patienten wirksameren Herangehensweisen unberücksichtigt (vgl. Grawe 1998)

So kommt es, daß in der Praxis (bzw. den Praxen) bei der Therapieplanung die Ergebnisse zur differentiellen Indikation wohl kaum eine Rolle spielen. Ebenso werden störungsspezifische Behandlungsangebote bisher nicht flächendeckend angeboten, und es ist nicht immer leicht in Erfahrung zu bringen, an welcher Stelle eine spezialisierte Behandlung angeboten wird.

Die Vorhersagbarkeit des Erfolgs oder Mißerfolgs einer Therapie aufgrund der *vor Beginn der Behandlung* verfügbaren Information ist also einerseits begrenzt und andererseits aus praktischen Gründen für die Zuweisung der Patienten zu den Behandlungen im Alltag (momentan) kaum relevant. Eine Optimierung der Versorgung über die gezieltere Zuweisung der Patienten zu den bestmöglichen Behandlungen dürfte unter den gegenwärtigen Umständen nur schwer zu erreichen sein. Auch Grawe (1998) resümiert:

Ich glaube, daß in einem Versorgungssystem, in dem Psychotherapien in der Regel ein solcher Abklärungsprozeß [zur differentiellen Indikation] vorgeschaltet würde, wesentlich bessere Therapieergebnisse resultieren würden, als es gegenwärtig der Fall ist. Außerdem wäre ein solches System viel kostengünstiger, weil die enormen Kosten, die bisher durch Fehlindikation und nicht optimale Behandlung entstehen (Meyer et al. 1991), beträchtlich reduziert würden. Die Kosten, die durch einen solchen Abklärungsprozeß entstünden, wären jedenfalls gering im Vergleich zu diesen Einsparungen.

Mir ist klar, daß ich hier gerade ein psychotherapeutisches Utopia ausgemalt habe. (Grawe 1998, S.168f)

Es stellt sich die Frage, wo und wie im Prozeß der Behandlungsplanung angesetzt werden kann, um ungünstige Therapieergebnisse möglichst frühzeitig zu vermeiden. Wenn mit dem Beginn einer Therapie noch ein hinreichend großes Risiko des Scheiterns oder einer nur geringen Besserung besteht, so liegt es nahe, so früh wie möglich *während* der Behandlung abzuklären, ob sie tatsächlich hilfreich ist oder nicht. Dieser Gedanke führt direkt zu Überlegungen, wie in der laufenden Behandlung Strategie und Methode systematisch an die Erfordernisse der Patienten angepaßt werden können.

2.2.2 Adaptive Indikation

Die Anpassung von therapeutischen Interventionen im Verlauf einer Behandlung (adaptive Indikation) dient dem Ziel, die Behandlung gemäß den eingetretenen (oder auch nicht eingetretenen!) Fortschritten des Patienten zu verändern, so daß ein erfolgreicher Abschluß der Behandlung erreicht werden kann. Diese Anpassungen nehmen wohl alle therapeutisch Arbeitenden mehr oder weniger systematisch vor (z.B. durch Supervision); im Prozeß der adaptiven Indikation kann eine vorläufige Bestandsaufnahme vorgenommen werden, um abzuklären, ob die Therapie in dem gewünschten Rahmen verläuft oder ob Veränderungen in der Behandlung anstehen. Empirische Untersuchungen zu diesem Thema sind selten, da sie einerseits aufwendige verlaufsorientierte Erhebungen erfordern und andererseits auch die Absichten und Handlungspläne der Therapeuten betreffen, die sich rasch ändern können (Schulte 1996, Vogel & Schulte 1997).

Die Anpassung der Therapie an das individuelle Geschehen erfordert in jedem Fall eine Bewertung der bisherigen Ereignisse auf dem Hintergrund ihrer Bedeutung für den

weiteren Therapieverlauf. Es geht also darum, aus dem bisherigen *individuellen* Therapieschehen zu lernen, um für den weiteren Verlauf die geeigneten Weichenstellungen zu treffen.

Schulte (1996) spricht in diesem Zusammenhang von *Vorausplanung* und *Verlaufsanpassung* innerhalb der Behandlung. Verlaufsanpassungen werden nötig, sobald ein Therapeut feststellt, daß das Resultat des bisherigen Vorgehens unbefriedigend ist und er darüberhinaus bei Beibehaltung des aktuellen Vorgehens keine Besserung prognostiziert. Befunde aus einer Studie zu diesem Thema (Schulte-Bahrenberg & Schulte 1993) zeigen jedoch, daß Therapeuten das voraussichtliche Therapieergebnis mit nur unzureichender Genauigkeit (selbst gegen Ende der Therapie) abschätzen können. Schulte kommt zu dem Schluß:

Von daher scheint es als äußerst wichtig, daß der Therapeut sich bei der Einschätzung des Verlaufs der Therapie nicht nur auf seinen subjektiven Eindruck verläßt, sondern möglichst objektive Wege wählt und objektive Testverfahren einsetzt, um bereits während des Therapieverlaufs die jeweils erreichten Veränderungen der Symptomatik oder Probleme zu erfassen (Schulte 1996, S.47).

Mit der Erfassung der Veränderungen alleine ist jedoch zunächst keine Prognose möglich; um die prognostische Bedeutung von Ereignissen oder Veränderungen während der Therapie zu erfahren, müssen diese mit dem Therapieergebnis in Beziehung gesetzt werden. An dieser Stelle gibt es bislang wenig empirisch gesichertes Wissen darüber, welche Ereignisse oder Entwicklungen innerhalb der Therapie eine positive oder negative prognostische Bedeutung bzgl. des weiteren Therapieverlaufes haben (vgl. Abschnitt 3.3). Noch unklarer ist zwangsläufig, zu welchem Zeitpunkt solche Prognosen überhaupt möglich sind. Gelänge es, solche prognostisch relevanten Ereignisse zu identifizieren, könnten sie die Planung innerhalb einer laufenden Behandlung unterstützen. Die prognostische Bedeutung oder 'prädiktive Kraft' (siehe Abschnitt 3.1) solcher Ereignisse wird als zentrales Kriterium bei der Definition von Therapie-Response eine Rolle spielen, da sie die Brücke zur Behandlungsplanung schlägt.

Zusammenfassung:

An dieser Stelle bleibt festzuhalten, daß ein verlaufsorientiertes Vorgehen im QM die unmittelbarste Relevanz für Therapeut und Patient besitzt, da die erhobene Patienteninformationen dem selben Patienten wieder zugute kommen kann. Es ist zu erwarten, daß solch ein System – relativ wenig organisatorischen Aufwand für die Therapeuten vorausgesetzt – praxisnähere Informationen zur Verfügung stellt und daher auch für die Beteiligten attraktiver und motivierender wirkt, ohne daß die Möglichkeiten der aggregierten Verwendung der Daten beschnitten werden. Allerdings eröffnet dies Vorgehen auch neue Fragen: insbesondere die Erarbeitung von bedeutungsvollen Verlaufsstandards ist noch offen.

Das Wissen aus der Psychotherapieforschung über eine optimale Zuteilung von Patienten zu den für sie am geeignetsten Therapien (differentiellen Indikation) ist noch mit vielen Unsicherheiten behaftet, so daß es in der Alltagspraxis kaum zum Zuge kommt; zusätzlich verhindern versorgungstechnische Gründe eine gezielte Zuweisung der Patienten. Ein ergänzender Zugang zur Verbesserung der therapeutischen Versorgung liegt daher in der Anpassung der laufenden Behandlung an die Gegebenheiten des Patienten. Dies erfordert eine Prognose über den weiteren Verlauf, um bei negativer Prognose die Behandlung ändern zu können; dem Konzept des Therapie-Response kann hierbei die Rolle zukommen, diese Prognose mit empirischen Hinweisen zu unterstützen.

3 Das Konzept des Therapie-Response

Eine systematische Untersuchung von Therapie-Response erfordert eine genaue Explikation des Begriffs und anschließend eine Operationalisierung anhand der vorliegenden Daten. Dieser Abschnitt befaßt sich mit der Explikation des Begriffs vor dem Hintergrund verschiedener schulenübergreifender Psychotherapietheorien (Abschnitt 3.2) und stellt einige Ergebnisse verlaufsorientierter Untersuchungen zusammen, welche Hinweise über den Zusammenhang zwischen Therapie-Response und endgültigem Therapieergebnis geben (Abschnitt 3.3).

3.1 Explikation des Begriffs ‘Therapie-Response’

Eine erste Annäherung an den Begriff des Therapie-Response kann über das Alltagsverständnis des ‘Ansprechens auf eine Behandlung’ (oder ein Medikament) geschehen. Üblicherweise spricht man dann von einem Ansprechen auf die Behandlung, wenn ein Medikament erste Anzeichen einer Wirkung zeigt. So etwa wenn nach der Einnahme eines fiebersenkenden Präperates die Körpertemperatur zu sinken beginnt; oder wenn bei einem Bluthochdruck-Patienten nach der Umstellung der Ernährung, vermehrter körperlicher Aktivität und Einnahme entsprechender Medikamente die Blutdruckwerte zu sinken beginnen. Dafür brauchen die Werte noch nicht im Normalbereich zu sein, es genügt die Feststellung einer ausreichend großen Reduktion. Das Ansprechen läßt einen weiteren Fortschritt bei gleicher Behandlung erwarten; der Patient scheint auf dem Wege der Besserung zu sein.

Frank et al. (1991) geben einige Hinweise zur Konzeptualisierung von Therapie-Response: Sie präsentieren ein Schema zum einheitlichen Gebrauch der Begriffe ‘Remission’ (voll und teilweise; partial remission, full remission), ‘Rückfall’ (relapse), ‘Heilung’ (recovery) und ‘Wiederauftreten einer Störung’ (recurrence). Bei der Definition der Begriffe müssen dabei sowohl Symptomschwere als auch Dauer des symptomatischen Zustandes beachtet werden. Das Ziel in der Vereinheitlichung dieser Begriffe liegt

u.a. darin, einen besser auf den Verlauf abgestimmten Einsatz von Medikamenten zu ermöglichen (vgl. Lavori, Dawson & Mueller 1994; Dawson et al. 1998).

Response besteht bei ihnen *im Beginn der partiellen Remission unter stattfindender Behandlung*. Response in diesem Sinne ist charakterisiert durch folgende Kriterien:

- Die Patienten erfüllen nicht mehr alle Kriterien der Diagnose, aber es liegt noch mehr als nur minimale Symptomatik vor. Es ist damit eine Art ‘Zwischenzustand’ zwischen vollem Krankheitsbild und vollständiger Besserung gemeint.
- Die Verbesserung hat während der Behandlung stattgefunden und ihr Eintreten wird daher als direkte Wirkung der Behandlung angesehen. Ein Response ohne Behandlung ist folglich nicht möglich, bzw. eine Verbesserung ohne Behandlung sollte eine andere Bezeichnung erhalten.
- Wichtigstes Validitätskriterium aller von Frank et al. (1991) aufgeführten Operationalisierungen ist die ‘Vorhersagekraft’ (‘predictive power’) in Bezug auf das Behandlungsergebnis. Die Begriffe erlauben nur dann eine besser abgestimmte Behandlungsplanung, wenn das Erreichen eines Zustandes auch Auskunft über den weiteren Verlauf gibt. So sollte z.B. das Auftreten von Rückfällen nach einer Recovery (‘Heilung’) deutlich geringer sein als nach einer vollen Remission. Diese Unterscheidung hat direkte Implikationen für die weitere Behandlung: Ein Patient in voller Remission steht demnach nach wie vor unter dem Risiko eines Rückfalls und braucht ggf. sporadische ‘Auffrischungen’ der Behandlung, während dies für einen Patienten nach einer ‘Heilung’ nicht mehr notwendig ist.

Frank et al. (1991) beschreiben mögliche Konsequenzen eines Therapie-Response für die weitere Behandlungsplanung:

‘Wenn der Patient sich in der partiellen Remission befindet (egal ob dies dem natürlichen Verlauf entspricht oder durch die Behandlung geschehen ist), wird der Therapeut eher abwarten als die Behandlungsstrategie ändern. Erst wenn sich die partielle Remission nach einer angemessenen Zeit nicht zu einer vollen Remission entwickelt, wird der Therapeut üblicherweise die Behandlung ändern: entweder wird er die Intensität erhöhen (durch eine Höherdosierung der Medikamente oder Erhöhung der Anzahl an Therapiesitzungen) oder er wird zu der Behandlung, die mit der partiellen Remission verbunden war, zusätzliche Behandlungen anordnen – oder er wird zu einer neuen Behandlungsart wechseln’. (Frank et al. 1991, S. 853, eigene Übersetzung)

Damit kann der Begriff des Therapie-Response wie folgt beschrieben werden:

Therapie-Response besteht in einer Reduktion der Symptomatik, die anzeigt, daß der Patient sich auf dem Wege der Besserung befindet. Er besteht in einer Art Zwischenzustand zwischen vollem Syndrom und minimaler Symptomatik, der unter dem Einfluß von Therapie eingetreten ist. Seine Definition sollte so festgelegt werden, daß das Auftreten von Therapie-Response Aussagen über den weiteren Gesundungsverlauf zulassen, damit er zur Unterstützung weiterer Behandlungsentscheidungen herangezogen werden kann. Ob eine spezielle Definition für Response dem Kriterium der ‘prädiktiven Validität’ genügt, muß in eigenen Untersuchungen empirisch überprüft werden, so wie es in dieser Arbeit vorgenommen wird (s. Abschnitt 7).

3.2 Therapie-Response in schulenübergreifenden Psychotherapiemodellen

Aus dem vorigen Abschnitt wird deutlich, daß sich Therapie-Response durch den Beginn der Besserung – also durch eine Abnahme der Symptome – bemerkbar macht. Die Frage, *wie* Therapie-Response durch die Behandlung zustande kommt und welche Bedingungen dafür erfüllt sein müssen, geschieht durch die Untersuchung der Wirkfaktoren in der Psychotherapie und ihre Bedeutung für das Therapieergebnis. Versuche, die Resultate der Psychotherapie-Ergebnis und -Prozeß-Forschung zu übergreifenden Modellen zusammenzufassen, sind an dieser Stelle besonders geeignet, die einzelnen Bereiche und Konstrukte vor Augen zu führen, die mit dem Verlauf und dem Ergebnis einer Therapie zusammenhängen. Sie bieten dadurch einen Rahmen, um die zum Ablauf und Planung einer Behandlung relevanten Bereiche in einem Ganzen übersichtlich zusammenzufassen. Es besteht somit die Möglichkeit, Therapie-Response in das Geflecht der wichtigsten Konstrukte über Psychotherapie einzuordnen. In den folgenden Abschnitten sollen 4 Modelle kurz skizziert und im Hinblick auf ihre Aussagen und Ansatzpunkte bzgl. Therapie-Response untersucht werden.

3.2.1 Das Modell der Systematischen Behandlungsplanung

Beutler & Clarkin (1990) legen ein Modell zur systematischen Behandlungsauswahl vor, in dem sie die verfügbaren Forschungsergebnisse zur differentiellen Wirksamkeit von Psychotherapie integrieren, um ‘das anzuwenden, was wir darüber wissen, was wann für wen wirken kann’ (S. XI). Ihr Anliegen ist weniger, eine umfassende Theorie über das Zusammenwirken der in der Psychotherapie beteiligten Faktoren zu beschreiben, sondern aus den Ergebnissen der Psychotherapieforschung diejenigen Dimensionen zusammenzustellen, welche hinsichtlich der Auswahl der erfolgversprechendsten Behandlungs- und Interventionsstrategien relevant sind. Sie markieren 4 Bereiche, die

bei der Planung jeder Therapie – unabhängig von der ‘Therapieschule’ – zu berücksichtigen sind.

In diesem Modell spielen zunächst die Ausgangsbedingungen der Patienten (z.B. Diagnose, Schweregrad, Persönlichkeit, soziales Umfeld) eine Rolle, die eine erste Einschätzung über das zu erwartende Ergebnis ermöglichen. Damit verbunden ist weiterhin die Auswahl des geeignetsten Behandlungssettings (z.B. Gruppen- oder Einzeltherapie, Dauer und Frequenz der Behandlung) zu treffen, wobei hier auch Kosten und Nutzen gegeneinander abgewogen werden. Hinweise über diese Zusammenhänge stammen aus den Untersuchungen zu prognostischen Faktoren bzgl. des Therapieerfolgs so wie zur differentiellen Wirksamkeit unterschiedlicher Therapieformen. Als drittes ist die ‘Passung’ zwischen Patient und Therapeut zu beachten, worunter eine Reihe von Aspekten der Beziehungsgestaltung fallen (z.B. Auswahl des Therapeuten/Patienten, Antipathie und Sympathie, Aufbau einer förderlichen Beziehung, Abstimmung der Therapie auf die Bedürfnisse des Patienten).

Als letzten und für diese Arbeit relevantesten Punkt führen die Autoren die angemessene Durchführung der therapeutischen Strategien und Techniken an (z.B. Auswahl des Therapiefokus, der Interventionen und deren Reihenfolge). Diese müssen mit dem aktuellen Zustand des Patienten abgestimmt werden und können sich daher im Verlauf der Behandlung oder auch innerhalb einer Sitzung verändern. Das entspricht damit den Aufgaben der adaptiven Indikation, wie oben beschrieben. Die Autoren beschreiben 3 wichtige Dimensionen, die während der Behandlung kontinuierlich zu beachten sind:

- Der aktuelle Leidensdruck (‘Arousal’):
Damit ist die momentane Belastung der Patienten gemeint, die das Problem verursacht, wegen dessen er in Behandlung kam. Ist die Belastung gerade zu stark, so hat der Patient keine Ressourcen zur Verfügung, um sich auf die Lösung des Problems zu konzentrieren – ist der Leidensdruck gerade zu gering, besteht die Gefahr, daß der Patient die Motivation verliert und Veränderungen blockiert.
- Die Anzeichen für Reaktanz seitens des Patienten:
Fühlen sich Patienten zu sehr in ihrem Handlungsspielraum eingeengt, so reagieren sie ablehnend gegenüber den angebotenen Interventionen. Daher sollte der Therapeut ein Vorgehen wählen, daß bei den Patienten möglichst wenig Reaktanz aufkommen läßt.
- Fortschritt bei der Problemlösung:
In diesem Punkt heben die Autoren hervor, das Vorgehen dem Stand der Problemlösung anzupassen. Ist der Patient schon darangegangen, konkrete Schritte zur Problemlösung zu tun, so ist es förderlicher, den Fokus auf das Geschehen außerhalb der Therapie zu setzen – ist das Problem noch nicht so deutlich, sollte

der allgemeine Fokus eher auf dem Geschehen innerhalb der Therapie und dem Erleben des Patienten liegen.

Beutler und Clarkin geben gleichzeitig Empfehlungen darüber ab, welche Art von therapeutischen Strategien zu wählen sind, um den Patienten je nach aktueller Lage auf diesen Dimensionen ein Fortkommen in der Therapie zu erleichtern. Beispielsweise empfehlen sie bei gerade reaktanten Patienten ein eher non-direktives Vorgehen. Bei Patienten, die in einer Therapiephase sind, wo es darum geht, die erreichten Fortschritte zu verfestigen, schlagen sie einen eher nach außen (d.h. auf konkrete Situationen und Gegebenheiten) gerichteten Therapiefokus vor, im Gegensatz zu einem nach innen auf das gerade aktuelle Erleben und Empfinden gerichteten Fokus. Die besten Chancen auf eine positive weitere Entwicklung bestehen danach, wenn es dem Therapeuten gelingt, das 'Arousal' des Patienten auf einem mittleren Niveau zu halten und den Therapiefokus entsprechend dem Stand der Problembewältigung zu setzen, ohne bei der Wahl des Vorgehens zu viel Reaktanz hervorzurufen. Ihre Empfehlungen fußen allerdings mehr auf Untersuchungen, die einen generellen Zusammenhang der genannten Dimensionen mit dem Therapieergebnis belegen, als auf solchen, in denen das Vorgehen nach den von ihnen vorgeschlagenen Kriterien kontinuierlich verfolgt und überprüft wurde.

Im Blick auf Therapie-Response liegen keine direkten Anhaltspunkte darüber vor, welche 'Mechanismen' hinter dem Beginn der Symptomreduktion liegen könnten. Immerhin bezeichnen die drei Dimensionen 'Leidensdruck', 'Reaktanz' und 'Fortschritt bei der Problemlösung' Variablen, die Therapie-Response ermöglichen oder auch behindern können, aber sie beschreiben nicht, welche Prozesse dabei am Fortschritt einer Therapie beteiligt sind.

Beutler und Clarkin streichen die Notwendigkeit heraus, die Interventionen gemäß des Fortschritts der Patienten anzupassen (adaptive Indikation). Dadurch deuten sie Möglichkeiten bzgl. des Verlaufs einer Therapie an: Gelingt es von Anfang an nicht, auf den genannten Dimensionen einem Patienten durch die geeigneten Intervention entgegenzukommen, so stellt sich u.U. gar kein Therapie-Response ein; hat der Patient schon auf die Behandlung angesprochen, so ist nicht automatisch Therapieerfolg zu erwarten, da sich das Befinden der Patienten auf den o.g. Dimensionen von Sitzung zu Sitzung oder auch innerhalb einer Therapiesitzung ändern kann. Das Risiko einer erneuten Verschlechterung bleibt bestehen, was einen ggf. nicht-linearen Therapiefortschritt impliziert.

3.2.2 Das Allgemeine Modell der Psychotherapie / Phasenmodell

Das Allgemeine Modell der Psychotherapie Orlinsky & Howard (1987) versuchen in ihrem 'Allgemeinen Modell der Psychotherapie' (generic model of psychotherapy)

die wichtigsten Elemente einer Psychotherapie in ihrer Wirkung untereinander generell zu beschreiben. Sie unterscheiden zwischen den Ausgangsbedingungen, dem Prozeß und dem Ergebnis einer Therapie. Wichtiges Merkmal dieses Modells sind Rückkopplungsschleifen zwischen einzelnen Komponenten, die den sukzessiven Fortschritt einer Therapie erklären können.

Zu den Ausgangsbedingungen gehören hier neben den Patienten- und Therapeuteigenschaften auch der gesellschaftliche Kontext (kulturelle Normen, soziale und institutionelle Strukturen) und insbesondere auch die Beteiligung/Interessen Dritter an der Therapie (Kostenträger, Arbeitgeber, Familie) im Sinne des Drei-Parteien-Modells (Strupp & Hadley 1977).

Zu den Therapieprozeßvariablen gehört der therapeutische Vertrag, die hilfreiche Arbeitsbeziehung und die Realisierung der therapeutischen Interventionen. An dieser Stelle ist insbesondere der Rückkopplungsprozeß zwischen Therapiebeziehung, Offenheit der Patienten und Wirksamkeit der Interventionen zu nennen: Eine positive Beziehung zwischen Patient und Therapeut fördert die Offenheit seitens des Patienten für das weitere therapeutische Vorgehen, welches die Wirksamkeit der Interventionen erhöht ('Remoralization', s.u.). Diese erreichten sog. 'Mikroergebnisse' in der Behandlung stärken wiederum die therapeutische Beziehung, so daß hier ein Aufschaukelungsprozeß in Gang kommen kann.

Auf der Ergebnisseite stehen lang-, mittel-, und kurzfristige Ergebnisse (Mikro-, Meso- und Makroergebnisse). Es wird angenommen, daß sich die Mikroergebnisse (z.B. neue Einsichten, Katharsiserlebnisse) mit der Zeit zu Mesoergebnissen (z.B. mehr soziale Kompetenz, selbständigere Lebensführung) kumulieren. Diese können wiederum Auswirkungen auf die Umwelt und andere Beteiligte haben, die in einem zweiten Rückkopplungsprozeß günstig (oder auch ungünstig) auf den Zustand des Patienten zurückwirken. Ein Patient mag z.B. durch die Therapie befähigt werden, seine Umwelt so zu verändern, daß weitere Fortschritte möglich sind (z.B. eine Mitgliedschaft in einem Verein beginnen, um die soziale Isolation zu mindern). Die beiden Rückkopplungsprozesse (siehe I., II. in Abbildung 1) liefern Anhaltspunkte dafür, wie Therapie-Response zustandekommen kann: In einer gelingenden Behandlung sollte sich nach gewisser Zeit eine positive therapeutische Arbeitsbeziehung etabliert haben und der Patient oder die Patientin Offenheit für die Interventionen zeigen. Sie sind wichtige Voraussetzungen für weitere Verbesserungen und deren Kumulation – der Patient 'spricht auf die Behandlung an'. Noch deutlicher läßt sich Therapie-Response über die Kumulation bisheriger Behandlungsergebnisse verstehen: sie führen zu einer Verbesserung des Zustandes (patient's on-going) und damit zu einer Verminderung der Symptome.

Das Phasenmodell Das Phasenmodell (Howard, Lueger, Maling & Martinovich 1993) beruht auf den Ideen von Frank (1973) und Untersuchungen zur Dosis-Wirkung-

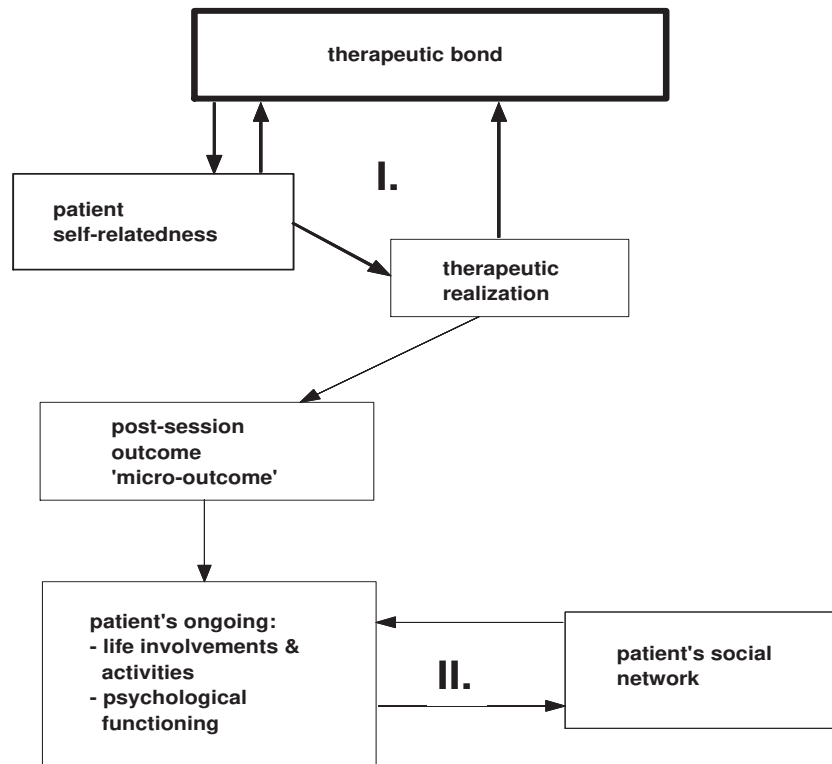


Abbildung 1: Ausschnitt aus dem 'Allgemeinen Modell der Psychotherapie'

Beziehung von ambulanter Psychotherapie (Howard, Kopta, Krause & Orlinsky 1986). Es beschreibt, wie sich unterschiedliche Problembereiche der Patienten verschieden schnell entwickeln. Der Ablauf einer Therapie kann danach in drei Phasen eingeteilt werden: *Remoralisierung*, *Remediation* und *Rehabilitation*.

In der *Remoralisierungsphase* verbessert sich das momentane Wohlbefinden der Patienten. Diese Verbesserung findet innerhalb der ersten Sitzungen statt und wird hauptsächlich auf positive Erwartungen der Patienten an die Therapie zurückgeführt. Die anschließende *Remediationsphase* ist durch Verbesserungen in der Symptomatik der Patienten gekennzeichnet. Sie wird für die ersten vier Monate der Therapie (ca. zwölf Sitzungen) angenommen. Die Veränderungen sind die Auswirkungen der therapeutischen Arbeit an den unmittelbaren Problemen der Patienten. Daraufhin folgt die *Rehabilitationsphase*, in der sich das allgemeine Funktionsniveau bessert. In dieser Phase findet sich der Patient wieder zunehmend besser mit seinem sozialen Umfeld und mit seinen Rollen in den verschiedenen Lebensbereichen (Arbeit, Familie, Freizeit usw.) zurecht. Die Verbesserung des allgemeinen Funktionsniveaus wird dabei nach ca. einem halben Jahr (26 Sitzungen) erwartet. Je nach Diagnose oder Schwere der Störung finden die Veränderungen schneller oder langsamer statt (Lueger 1995). Das Modell postuliert gleichzeitig auch eine kausal-probabilistische Abhängigkeit dieser Phasen (Howard et al. 1993). Demnach ist die Verbesserung des Wohlbefindens eine *notwendige* Voraussetzung dafür, Verbesserungen auf symptomatischer Ebene zu erzielen; diese wiederum

erst ermöglichen eine Verbesserung des allgemeinen Funktionsniveaus.

Die Verbindung des ‘Allgemeinen Modells der Psychotherapie’ und des Phasenmodells liefert einen guten Hintergrund zur Konzeptionalisierung von Therapie-Response; ersteres beschreibt die inhaltlichen Komponenten, während letzteres den zeitlichen Rahmen absteckt. Ein ‘Ansprechen auf die Therapie’ würde danach mit dem ersten Rückkopplungsprozeß des Allgemeinen Modells beginnen: Eine gelingende therapeutische Beziehung erleichtert die Selbstöffnung der Patienten und bereitet damit den Boden für die Wirksamkeit der Interventionen. Gemäß des Phasenmodells sollte das in einer Besserung des Wohlbefindens nach ca. vier Sitzungen sichtbar werden. Laut Modell wird dadurch die Wahrscheinlichkeit auf eine Besserung der Symptome erhöht; die Besserung des Wohlbefindens wäre demnach – wie oben gefordert – ein Ereignis mit ‘prädictiver Validität’ für den weiteren Verlauf (vgl. auch Abschnitt 11.2). Die Mikroergebnisse kumulieren sich auf zu Meso- und Makroergebnissen, so daß sich Verbesserungen in der Symptomatik ergeben. Diese Verbesserung sollte sich nach ca. zwölf Sitzungen bemerkbar machen.

Therapie-Response kann also einmal als die ‘Verbesserung des Wohlbefindens nach ca. vier Wochen’ oder als die ‘Verbesserung der Symptomatik nach drei Monaten’ (in diesem Modell ambulanter) Therapie verstanden werden. Es ist das einzige Modell, das Annahmen über den zeitlichen Rahmen des Auftretens von Therapie-Response macht.

3.2.3 Dreikomponentenmodell zur Wirkungsweise von Psychotherapie

Grawe (1998) entwickelt in seiner ‘Psychologischen Therapie’ ebenfalls ein elaboriertes Modell zur allgemeinen Wirkungsweise von Psychotherapien. Sein Buch beginnt mit dem ‘Phänomen schneller Besserungen zu Beginn einer Therapie’, welches er als Aufhänger für das sich entwickelnde fiktive Gespräch über die Wirkfaktoren der Psychotherapie benutzt – er gebraucht in diesem Zusammenhang allerdings nicht den Begriff des Therapie-Response. Grawe nimmt explizit Bezug auf das o.g. Phasenmodell und führt den Effekt des verbesserten Wohlbefindens zunächst auf eine ‘Induktion positiver Erwartungen’ zurück. Ganz ähnlich wie im ‘Allgemeinen Modell der Psychotherapie’ von Orlinsky und Howard enthält Grawes Modell zwei Rückkoppelungsschleifen, in denen teilweise die gleichen Konstrukte zusammenwirken wie im ‘Allgemeinen Modell’: Der erste Rückkopplungsprozeß (siehe Schleife I. im Schema) ist fast identisch und ist auf der Ebene der ‘Ressourcenaktivierung’ verortet. Zu Beginn der Therapie entstehen beim Patienten durch ressourcenaktivierende Interventionen bedürfnisbefriedigende Erfahrungen (Bedürfnis nach Orientierung und Kontrolle, nach Selbstwerterhöhung und das Bindungsbedürfnis), die alle zusammen das Bedürfnis nach Konsistenz zwischen den Grundbedürfnissen befriedigen. Dies führt zu einem verbesserten Wohlbefinden, zu erhöhter Aufnahmebereitschaft und aktiver Mitarbeit und leistet einen positiven

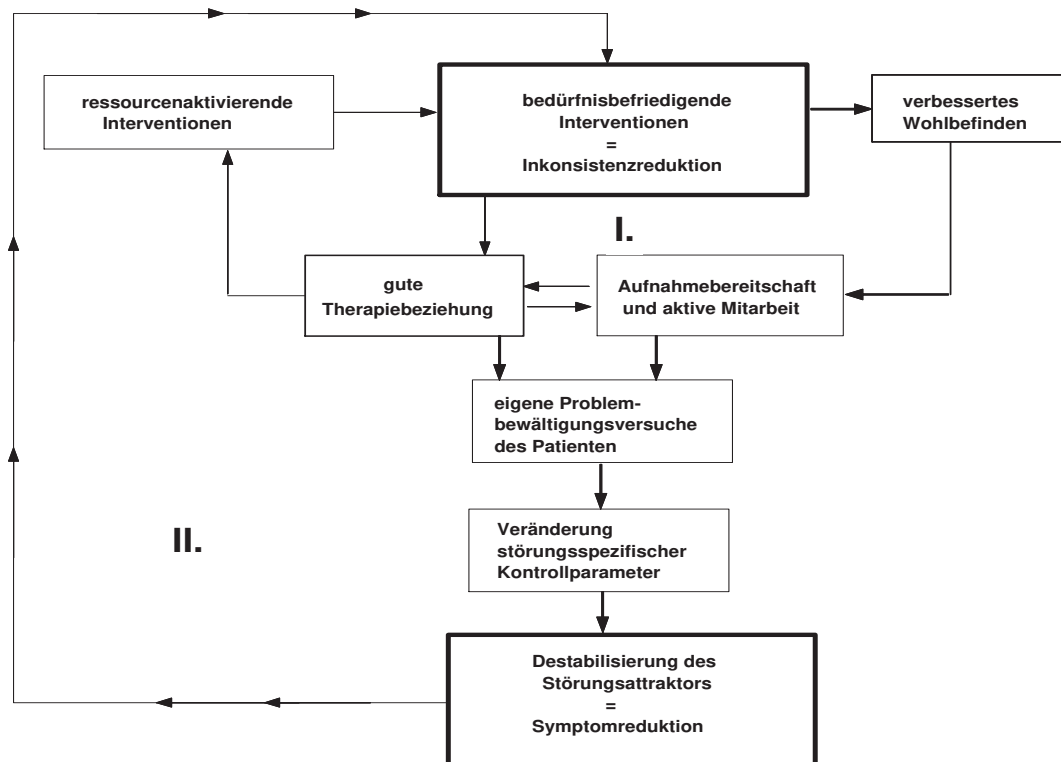


Abbildung 2: Ausschnitt aus dem Dreikomponentenmodell

Beitrag zur Arbeitsbeziehung. Dies verstärkt die bedürfnisbefriedigenden Erfahrungen und damit wieder das Wohlbefinden.

Auf der Ebene der 'Destabilisierung von Störungsattraktoren' befindet sich der zweite Rückkopplungsprozess (siehe Schleife I. im Schema): hier reduzieren sich die Symptome durch Problembewältigungsversuche, korrektive Erfahrungen und vertieftes Bewusstsein für problemrelevante Zusammenhänge. Dies fördert erneute bedürfnisbefriedigende Erfahrungen und regt damit die erste Rückkopplungsschleife an mit den oben beschriebenen positiven Wirkungen.

Die dritte Ebene oder Komponente ist die 'Veränderung motivationaler Schemata' (nicht im obigen Schema enthalten). Durch das Geschehen in der Therapie werden diese Schemata aktiviert, die sowohl förderlich oder durch ihre potentielle Konflikthaf-tigkeit auch hemmend wirken können. Wird ein konflikthafteS motivationaler Schema aktiviert, führt dies zu einer Erhöhung der Konfliktspannung und damit auch zu einer Inkonsistenz zwischen den Bedürfnissen: die Folge ist eine Verschlechterung des Wohlbefindens.

Dieses Modell führt durch die große Ähnlichkeit mit dem 'Allgemeine Modell der Psychotherapie' zu den gleichen Konzeptionalisierungen für Therapie-Response, auch wenn sich die Prozesse, die zu Therapie-Response führen, zwischen beiden Modellen teilweise unterscheiden. Therapie-Response kann auch hier als Verbesserung des Wohl-

befindens (erster Rückkopplungsprozeß) oder als Symptomreduktion (zweiter Rückkopplungsprozeß) verstanden werden.

3.2.4 Das Transtheoretische Modell der Psychotherapie

Prochaska & Norcross (1994) beschreiben in ihrem 'Transtheoretischen Modell der Psychotherapie' verschiedene Stufen der Veränderung ('Stages of Change'), die den Grad an Veränderungsbereitschaft angeben und den Verlauf in der Psychotherapie kennzeichnen. Dieser Verlauf umfaßt 5 Stufen: präkontemplative Phase, kontemplative Phase, Vorbereitungsphase, Handlungsphase und Erhaltungsphase. Die meisten empirischen Untersuchungen zu diesem Modell stammen aus dem Bereich klar umrissener Problembereiche (Raucherentwöhnung, Gewichtskontrolle oder Delinquenz bei Jugendlichen), sind jedoch auch allgemein auf die Änderungsprozesse in der Psychotherapie übertragen worden:

In der *präkontemplativen Phase* denken die Patienten nicht daran, in absehbarer Zeit ihr problematisches Verhalten (z.B. Rauchen) zu ändern. Das Problem wird gar nicht wahrgenommen, negiert oder verharmlost. Die *kontemplative Phase* ist dadurch gekennzeichnet, daß ernsthaft in Erwägung gezogen wird, das problematische Verhalten innerhalb eines halben Jahres zu ändern. Beim Abwägen der Vor- und Nachteile einer Veränderung besteht jedoch noch eine große Ambivalenz, so daß noch kein fester Entschluß gefaßt wird. In der *Vorbereitungsphase* befinden sich solche Patienten, die sich eine Veränderung innerhalb des nächsten Monats ernsthaft vorgenommen oder sogar schon erste Schritte zur Veränderung unternommen haben – allerdings sind noch keine entscheidenden Fortschritte zu verzeichnen. Die *Handlungsphase* zeichnet sich durch ein aktives Umsetzen von Verhaltensänderungen aus: Probleme werden klar benannt, systematisch angegangen und gelöst. An dieser Stelle finden die ersten Verbesserungen des Problemverhaltens statt. In der *Erhaltungsphase* ist das Problemverhalten weitgehend verschwunden und nun müssen die erreichten Veränderungen stabilisiert werden, um Rückfälle zu vermeiden.

Je nach dem, in welcher Phase sich ein Patient bei Beginn der Behandlung befindet, bedarf er unterschiedlicher Behandlungsstrategien (Heidenreich & Hoyer 1998) und wird unterschiedlich schnell auf die Behandlung ansprechen. Dabei kann sowohl der gesamte Therapieverlauf in diese Phasen eingeteilt werden, ebenso ist es aber auch möglich, bei höherem Auflösungsgrad diese Phasen als Veränderung auf verschiedenen Symptomebenen nacheinander zyklisch zu durchlaufen (vgl. Prochaska et al. 1991). Eine Besonderheit des Modells besteht darin, von jeder Stufe wieder in eine der vorhergehenden zurückschreiten zu können, so daß es Rückfälle und Stagnationen im Verlauf zuläßt. Wichtige moderierende Variablen in diesem Prozeß sind eine parallel zu den Stadien wachsende Selbstwirksamkeitserwartung und eine weniger ambivalente Bewertung

der Vor- und Nachteile einer Veränderung (Velicer, Rossi, Prochaska & DiClemente 1996).

Therapie-Response unter dieser Perspektive bestünde im Voranschreiten von einer Stufe der Veränderungsbereitschaft zur nächsten. Jeder Schritt in die nächste Stufe, bringt den Patienten einer Veränderung des Problemverhaltens näher. Der markanteste Sprung dürfte dabei der Übergang von der Vorbereitungsphase zur Handlungsphase sein, da mit dem Eintritt in die Handlungsphase erstmals eine deutliche Reduktion der Symptome zu erwarten ist.

Das Modell liefert einen zusätzlichen *motivationalen* Aspekt im Veränderungsgeschehen. Anknüpfend an die erste Annäherung an den Begriff des Therapie-Response in Abschnitt 3.1, ist die Veränderungsbereitschaft eher als eine Vorbedingung für Response aufzufassen, auch wenn diese direkte Folge einer gelungenen therapeutischen Intervention ist. Unter der Arbeitshypothese, daß Therapie-Response ein Zwischenzustand zwischen vollem Syndrom und Minimalsymptomatik darstellt, kann die Veränderungsmotivation trotz voll ausgeprägter Symptomatik vorhanden sein.

Zusammenfassung:

Der Blick auf schulenübergreifende Psychotherapietheorien legt es nahe, Therapie-Response als eine Verbesserung des momentanen Wohlbefindens bzw. (noch nicht vollständige) Besserung des symptomatischen Zustandes aufzufassen. Response ist dabei Resultat von sich gegenseitig verstärkenden Elementen (positive Erfahrungen, Öffnung für die Interventionen, Stärkung der Arbeitsbeziehung und symptomatische Verbesserungen). Das 'Allgemeine Modell der Psychotherapie' (Orlinsky & Howard 1987) in Verbindung mit dem Phasenmodell (Howard et al. 1993) bietet den elaboriertesten Zugang: es enthält die Mechanismen für zwei Arten von Therapie-Response (Verbesserung des Wohlbefindens, Verbesserung der Symptomatik) und erlaubt Hypothesen darüber, zu welchem Zeitpunkt er üblicherweise geschieht .

Darüberhinaus beschreiben die Theorien, welche Konstrukte als *Vorbedingungen* für Therapie-Response anzusehen sind (z.B. eine gute therapeutische Beziehung, Selbstöffnung des Patienten, vorhandene Änderungsmotivation). Sie können im Prozeß der adaptiven Indikation Anhaltspunkte dafür bieten, an welcher Stelle evtl. eine Änderung im therapeutischen Vorgehen vorgenommen werden kann, um die Erfolgchancen zu erhöhen.

3.3 Empirische Hinweise für die prädiktive Validität von Therapie-Response

In den letzten Jahren ist das Interesse an Studien, die versuchen, Verläufe von Behandlungen näher zu beschreiben und zu untersuchen – mitbedingt durch die Entwicklung

neuer statistischer Verfahren (Gibbons et al. 1993; Bryk & Raudenbush 1987) – gewachsen. Allerdings bedeutet eine längsschnittliche Forschungsstrategie einen höheren Zeit-, Kosten- und Organisationsaufwand, so daß gerade im klinischen Bereich Studien, die die Untersuchung des Therapieverlaufs ermöglichen, relativ selten sind (z.B. de Jong-Meyer et al. 1996 und Elkin 1994, für groß angelegte klinische Untersuchungen mit Verlaufserfassung).

Längsschnittliche Studien, die Informationen über den gesamten Therapieverlauf bereitstellen, konnten häufiger einen Zusammenhang zwischen früherem Therapieverlauf und dem künftigen Therapieergebnis feststellen. Dieser Zusammenhang scheint sowohl für langfristige Krankheitsverläufe, die von wiederkehrenden Episoden der Störung gekennzeichnet sind, zu bestehen, als auch für kurzfristige Behandlungsverläufe. In der Forschung zu chronischen psychischen Erkrankungen wie Schizophrenie und Depression gibt es schon seit längerem Untersuchungen, die den bisherigen Krankheitsverlauf mit dem Therapieergebnis in Beziehung setzen: Angst (1988) nennt den bisherigen Krankheitsverlauf (bei Depression) als besten Prädiktor für den weiteren Verlauf. Patienten mit einer hohen Anzahl von depressiven Episoden in der Zeit vor der Behandlung stehen unter einem erhöhten Risiko, nach einer Therapie erneut einen Rückfall zu erleiden, als Patienten mit weniger früheren Episoden.

Die Studien von Dew et al. (1997) und de Jong-Meyer et al. (1996) zeigen, daß sich die Symptomverläufe erfolgreich und nicht erfolgreich behandelter depressiver Patienten in der Behandlungsepisode hinsichtlich ihrer Verlaufsscharakteristika unterscheiden: Im Gegensatz zu den nicht erfolgreich behandelten Patienten ließen sich bei den erfolgreichen schon nach 3-4 Wochen Therapie erste Verbesserungen erkennen. Doch auch die erfolgreich behandelten Patienten folgen keinem homogenen Verlauf: Dew et al. (1997) konnten auch hier Gruppen identifizieren, die unterschiedlich schnell Verbesserungen zeigen.

Bei der Behandlung von Depressionen mit kognitiver Verhaltenstherapie gibt es eine Reihe von Studien, die den Symptomverlauf wöchentlich verfolgt haben. Dabei stellte sich heraus, daß viele Patienten den Großteil ihrer Symptomatik (gemessen am Beck Depressions Inventar, BDI) innerhalb der ersten 4 Wochen verlieren (Ilardi & Craighead 1994). In diesem Kontext entwickelte sich eine Diskussion darüber, ob diese frühen Veränderungen durch unspezifische Wirkfaktoren zustande kommen oder ob in dieser Phase schon spezielle Techniken der kognitiven Verhaltenstherapie zum tragen kommen (Tang & DeRubeis 1999). Die sich entfaltende Diskussion in der Zeitschrift 'Clinical Psychology: Science and Practice' (Ausgabe 6, 1999) berührt daneben auch den für diese Arbeit relevanten Themenkreis der Bedeutung früher Therapiefortschritte für die Therapieplanung.

3.3.1 Studie von Tang & DeRubeis (1999)

Tang & DeRubeis (1999) präsentieren die individuellen Verläufe von depressiven Patienten, die mit kognitiver Verhaltenstherapie behandelt wurden. Sie weisen innerhalb der ersten 10 Sitzungen sog. 'plötzliche Verbesserungen' (sudden gains) auf. Diese bestehen darin, daß sich der Wert des BDI von einer zur anderen Sitzung um ein Vielfaches im Vergleich zur Veränderung zwischen anderen Sitzungen verbessert, ohne daß sich die Symptome anschließend wieder auf das Niveau vor der plötzlichen Verbesserung verschlechtern. Die Folge ist eine deutliche 'Stufe' im Verlauf. Patienten mit diesen plötzlichen Verbesserungen hatten am Ende der Therapie und 18 Monate nach der Therapie ein besseres Ergebnis als Patienten ohne plötzliche Verbesserungen.

Wilson (1999) bringt in seiner Diskussion der Studie die Implikationen von Therapie-Response für die Behandlungsplanung auf den Punkt: er zitiert weitere (unten beschriebene) Untersuchungen, welche die Bedeutung des Behandlungsverlaufs für das Therapieergebnis demonstrieren und sagt in Bezug auf die Behandlungsplanung:

Die praktische Bedeutung von schnellem Ansprechen [rapid response] auf manuelle kognitive Verhaltenstherapie bei bestimmten Störungen (z.B. Depression und Bulimie) besteht darin, daß die Behandlung einer Änderung bedarf, wenn sich nach einer identifizierbaren Anzahl von Sitzungen der Zustand des Patienten nicht gebessert hat (Wilson 1999, S.291f, eigene Übersetzung).

3.3.2 Studie von Quitkin et al. (1996)

In der medikamentösen Behandlung von Depression konnten Quitkin et al. (1996) zeigen, daß die Erfolgchancen von Patienten, die innerhalb von vier Wochen Behandlung noch keine Anzeichen einer Besserung aufweisen, nicht mehr höher sind als die einer Placebo-Kontrollgruppe. Sie erhoben wöchentlich über einen Zeitraum von sechs Wochen den Status depressiver Patienten, von denen ein Teil mit Antidepressiva und ein anderer Teil mit Placebos behandelt wurde. Zu jedem Zeitpunkt wurden die Erfolgchancen der 'Non-Responder' (d.h. derjenigen Patienten, die zu diesem Zeitpunkt noch keine Besserung zeigen) in beiden Gruppen miteinander verglichen: Patienten die in den ersten drei Wochen noch keine Besserung aufwiesen, hatten unter medikamentöser Behandlung eine größere Chance auf Erfolg als Patienten der Placebo-Gruppe. Patienten, die nach vier oder fünf Wochen immer noch nicht auf die Behandlung ansprachen, hatten allerdings keinen Vorteil mehr gegenüber der Kontrollgruppe. Der Zeitpunkt, zu welchem der Vorteil der medikamentösen Behandlung verschwindet, variiert in Abhängigkeit des vorigen Verlaufes. Bei Non-Respondern, die auch in einem Placebo-Vorlauf keine Besserung zeigten, verschwand der Unterschied nach vier Wochen; wiesen die Patienten im Vorlauf eine minimale Besserung auf, verschwand der

Unterschied erst nach fünf Wochen. In ersterem Fall ziehen die Autoren die Konsequenz, daß bei Patienten, die nach vier Wochen noch nicht auf die Behandlung angesprochen haben, die Behandlungsstrategie zu ändern sei (Wechsel zu einem anderen Medikament oder Gabe von zusätzlichen Medikamenten).

Diese Studie fokussiert zwar den Vergleich zwischen Behandlungs- und Kontrollgruppe und nur indirekt den zwischen Respondern und Non-Respondern, aber sie macht deutlich, daß es unter bestimmten Bedingungen möglich ist, in Abhängigkeit des Ansprechens auf die Medikamente weitere Behandlungsempfehlungen abzuleiten. Eine Übertragung dieser Herangehensweise an psychotherapeutische Behandlungen steht aufgrund der erheblich komplexeren Zusammensetzung der Wirkfaktoren vor zusätzlichen Schwierigkeiten, aber die für Psychotherapie zu beantwortende Frage bleibt:

Wie können Therapeuten feststellen, daß Patienten, die keine Besserung zeigen, auch bei Fortsetzung der gleichen Behandlung nur wenig Chancen auf Besserung haben und daher eine Änderung der Behandlungsstrategie benötigen? (Wilson 1998, S.370)

3.3.3 Studie von Breslin et al. (1997)

Eine weitere von Wilson zitierte Studie ist die von Breslin et al. (1997): Ihr Anliegen besteht darin, die Vorhersage des Therapieerfolgs unter Zuhilfenahme von Prädiktoren aus dem Behandlungsverlauf zu verbessern. Sie untersuchten das Ausmaß an Alkoholkonsum bei Problemtrinkern ein halbes Jahr nach einem vierwöchigen angeleiteten Selbsthilfeprogramm. Zur Vorhersage des Konsums benutzten sie 1) die Anfangsinformationen der Patienten, 2) die Prognose des Therapeuten direkt nach Abschluß des Programms und 3) das Trinkverhalten und andere Maße aus der 4-wöchigen Behandlungsphase. Die 3 Gruppen von Prädiktoren wurden nacheinander in eine hierarchische Regressionsanalyse eingesetzt. Die Anfangsinformationen allein erklärten 36% der Variabilität des Alkoholkonsums nach einem halben Jahr, die Therapeutenprognose erklärte zusätzlich weitere 11% und die Prädiktoren aus der Behandlungsphase nochmals 11%. Wurden die Prädiktoren aus der Behandlungsphase vor der Therapeutenprognose ins Modell genommen, erklärten sie zusätzlich 24%, und die Therapeutenprognose trug im Anschluß daran nicht mehr zu einer Verbesserung des Modells bei. Der wichtigste Zusammenhang zwischen Alkoholkonsum und Informationen aus der Behandlungsphase bestand dabei im Anteil an abstinenten Tagen während des Selbsthilfeprogramms: je mehr abstinente Tage die Personen während der Behandlung hatten, desto niedriger war der Alkoholkonsum nach einem halben Jahr.

Breslin et al. (1997) folgern aus den Ergebnissen, daß Informationen aus der Behandlungsphase aufgrund ihrer prognostischen Bedeutung für die Behandlungsplanung relevant sind: Eine Person, die auch während des Selbsthilfeprogramms noch viel trinkt,

benötigt möglicherweise eine andere, intensivere Behandlung. Der Vorteil der Erfassung des Zustandes während der Behandlung besteht dann darin, das erhöhte Risiko noch *während der Behandlung* zu erkennen und sie entsprechend verändern zu können. Die Tatsache, daß sich die Prognose der Therapeuten als weniger stabiler Prädiktor erwies, interpretieren die Autoren damit, daß Therapeuten bei ihrer Prognose evtl. auf andere Aspekte achten, aber nicht unbedingt die bedeutsamsten Faktoren berücksichtigen.

3.3.4 Studie von Olmsted et al. (1996)

Bei Eßstörungen wurde ebenfalls ein Zusammenhang zwischen der Entwicklung der Symptome während der Behandlung und dem Therapieerfolg nach zwei Jahren gefunden: In einer Studie mit bulimischen Patientinnen (Olmsted, Kaplan, Rockert & Jacobsen 1996) zeigte sich eine sehr rasche Reduktion der Anzahl von Eß- und Brechanfällen innerhalb der ersten drei Wochen der Behandlung. Olmsted et al. (1996) teilten die Patienten in Gruppen ein, die ihre Symptome unterschiedlich schnell reduzierten. Von denjenigen Patienten, die ihre Symptome am schnellsten verloren hatten, erlebten innerhalb von zwei Jahren weniger einen Rückfall (15%), als diejenigen, die ihre Symptome nur langsam oder teilweise verloren hatten (über 55%). Es gelang den Autoren nicht, mit Hilfe der zu Beginn der Therapie verfügbaren Information die 'Ansprechgeschwindigkeit' der Patienten vorherzusagen.

Zusammenfassung:

Die hier angeführten Untersuchungen lassen die Vermutung zu, daß die Art und Weise des Ansprechens auf die angebotene Therapie für das Ergebnis von entscheidender Bedeutung ist. Das Ansprechen auf die Therapie war in den verschiedenen Studien recht unterschiedlich definiert: einmal waren es plötzliche Verbesserungen (Tang & DeRubeis 1999), ein anderes mal die Geschwindigkeit der Symptomreduktion (Olmsted et al. 1996) und wieder ein anderes mal die Differenz zum Eingangszustand nach einer gewissen Zeit (Quitkin et al. 1996). Wenn Therapie-Response tatsächlich mit dem Therapieergebnis zusammenhängt, kann er – so die Vermutung – im Prozeß der Behandlungsplanung zur Anpassung der Behandlung (adaptive Indikation) an den aktuellen Zustand des Patienten eingesetzt werden. Unter der Annahme, daß es für die Behandelnden schwierig ist, zutreffende Prognosen über die weitere Entwicklung der Patienten zu treffen (Schulte-Bahrenberg & Schulte 1993; Breslin et al. 1997), kann das Ansprechen auf die Behandlung bei der Erstellung einer Prognose und darauf aufbauend bei der weiteren Planung der Behandlung eine empirische Stütze liefern.

4 Gliederung des empirischen Teils

Aus dem bisher erarbeiteten wird deutlich, was unter Therapie-Response zu verstehen ist (Abschnitt 3.1 und 3.2): Es ist ein Zwischenzustand zwischen voller Symptomatik und der vollständigen Abwesenheit von Symptomen, der auf die Behandlung zurückzuführen ist. Das Ansprechen auf die Behandlung drückt sich dabei in einer Reduktion der Symptomatik aus. Zentrales Validitätskriterium von Therapie-Response ist die prognostische Bedeutsamkeit für den weiteren Verlauf. Der Blick in einige empirische Untersuchungen (Abschnitt 3.3) zeigt, daß es auch bei diesem Verständnis verschiedene Möglichkeiten gibt, Therapie-Response zu operationalisieren. Um Response für die Behandlungsplanung einzusetzen, muß die prädiktive Validität der Operationalisierungen untersucht werden; nur wenn anhand des Therapie-Response eine Aussage über den weiteren Therapieverlauf möglich wird, bieten sich Anhaltspunkte zur weiteren Planung. Zur Validierung muß daher Therapie-Response auf den Zusammenhang mit dem Therapieergebnis (am Ende der Behandlung und zur Katamnese) überprüft werden.

Im empirischen Teil dieser Arbeit wird genau dieser Zusammenhang zwischen Therapie-Response und Therapieerfolg untersucht. Allgemein formuliert lauten die Leitfragen diese Arbeit:

Inwiefern läßt sich aus der Entwicklung der Patienten in den ersten Therapiewochen (Therapie-Response) das endgültige Therapieergebnis vorher-sagen? Wie lassen sich die Ergebnisse für die weitere Behandlungsplanung fruchtbar machen?

Diese zwei Fragenkomplexe müssen in mehrere Teilfragen zerlegt werden, um sie einer Lösung näher zu bringen. Der erste Block behandelt die prädiktive Validität der Therapie-Response Operationalisierungen, der zweite Block befaßt sich mit der Umsetzung der Ergebnisse in ein verlaufsorientiertes QM-System.

4.1 Fragenblock zur prädiktiven Validität von Therapie-Response

1. Wie soll Therapie-Response operationalisiert werden?

Diese Frage wird im Abschnitt 6 angegangen. Es werden drei verschiedene Operationalisierungen erstellt, die sich sowohl auf das momentane Wohlbefinden, als auch auf die symptomatische Entwicklung beziehen. Dies ermöglicht die Untersuchung beider im Phasenmodell (vgl. Abschnitt 3.2.2) enthaltenen Konstrukte ‘Verbesserung des Wohlbefindens’ und ‘Verringerung der Symptome’, an denen sich Therapie-Response ablesen läßt.

2. Welche prognostische Bedeutung hat Therapie-Response?

Diese Frage zielt auf den Zusammenhang zwischen Therapie-Response und Behandlungsergebnis. Hier interessiert zum einen, ob Patienten, die Therapie-Response zeigen ('Responder'), sich hinsichtlich der Wahrscheinlichkeit auf Therapieerfolg von den Patienten ohne Therapie-Response ('Non-Responder') unterscheiden. Dieser Zusammenhang wird sowohl für den Erfolg zum Ende der Behandlung als auch für den Erfolg nach 2.5-Jahren untersucht. Wenn sich ein Zusammenhang zwischen Therapie-Response mit dem Zustand nach 2.5 Jahren herstellen läßt, ist dies ein starker Hinweis auf die prädiktive Validität des Therapie-Response.

3. Zu welchem Zeitpunkt sollte Therapie-Response erfaßt werden?

Es gibt kaum Untersuchungen, die Aufschluß darüber geben, bis zu welchem Zeitpunkt eine Veränderung eingetreten sein sollte, bzw. nach welchem Zeitraum ohne sichtbare Fortschritte keine Besserung mehr erwartet werden kann (vgl. Abschnitt 3.3.2). Das Phasenmodell macht die in Abschnitt 3.2.2 erwähnten zeitlichen Aussagen, die eine Verbesserung des Wohlbefindens nach 4 Sitzungen (also etwa vier Wochen) nahelegen. Es wäre durchaus möglich, daß eine Verbesserung nach zwei oder sechs Wochen eine größere prädiktive Validität aufweist, als eine Verbesserung nach vier Wochen. Es ist zu erwarten, daß sich die Vorhersage des Therapieergebnisses verbessert, je mehr man sich dem Zeitpunkt des Behandlungsendes nähert. Die Vorhersage des Therapieergebnisses nach 16 Wochen Therapie auf der Grundlage der Veränderungen bis einschließlich Woche 15, kann schwerlich noch als Vorhersage gewertet werden. Aus der Sicht der Behandlungsplanung ist ein möglichst früher Zeitpunkt wünschenswert, zu dem aber dennoch ein hoher Zusammenhang mit dem Behandlungsergebnis besteht.

Hier wäre ein möglichst kontinuierliches 'monitoring' des Verlaufes wünschenswert, um die Veränderungen lückenlos abzubilden und zu den verschiedenen Zeitpunkten die prädiktive Validität des Therapie-Response zu untersuchen. Leider sind im vorliegenden Datensatz nur vierwöchentliche Erhebungen enthalten, so daß der Zeitpunkt des Therapie-Response nur zu Woche 4 und Woche 8 untersucht werden kann; Woche 12 liegt bei den meisten Patientinnen zu nahe am Entlassungstermin, um das Therapieergebnis vorherzusagen zu können (vgl. Abschnitt 5.1).

4. Wird durch die Hinzunahme von Response die Vorhersage des Erfolgs verbessert?

Üblicherweise wird in vielen Studien versucht, aus der zu Beginn der Behandlung vorhandenen Information das Behandlungsergebnis vorherzusagen. Wie die Studie von Breslin et al. (1997) zeigt, kann die Vorhersage durch die Hinzunahme von

Information über den Therapieverlauf verbessert werden (vgl. Abschnitt 3.3.3). Ein mit dieser Studie vergleichbares Vorgehen bietet sich unter dieser Fragestellung an und ermöglicht eine Aussage über zwei Aspekte: zum einen eine Aussage darüber, ob sich mit der Hinzunahme von Therapie-Response die Vorhersage des Therapieerfolgs verbessert (gemessen am Anteil erklärter Varianz) – zum anderen wird damit gleichzeitig überprüft, ob Therapie-Response auch noch unter Berücksichtigung der Eingangsinformationen seine prädiktive Validität behält.

5. Kann man vorhersagen, ob ein Patient auf die Therapie ansprechen wird oder nicht?

In diesem Abschnitt wird der Zusammenhang zwischen den Eingangsinformationen der Patienten und dem Therapie-Response untersucht. Diese Frage wirft noch einmal von einer anderen Seite Licht auf den Beitrag des Therapie-Response zur Vorhersage des Behandlungserfolgs: Besteht zwischen Therapie-Response und Therapieerfolg ein enger Zusammenhang und läßt sich gleichzeitig Response sehr gut aus der anfänglichen Ausgangslage der Patienten ableiten, so ist fraglich, welchen *spezifischen* Beitrag Therapie-Response zur Vorhersage des Ergebnisses leistet. Daher soll hier Therapie-Response als *abhängige* Variable betrachtet werden, um herauszufinden, welche der Eingangsbedingungen mit einer erhöhten Wahrscheinlichkeit auf Therapie-Response verbunden ist.

6. Unterscheiden sich Responder und Non-Responder bzgl. des zeitlichen Verlaufes?

Unter der Perspektive des Therapieverlaufs, stellt sich die Frage, ob es bestimmte Eigenarten in der zeitlichen Entwicklung der Symptome gibt, die Responder von den Non-Respondern typischerweise unterscheiden. Wenn sich typische Verlaufskurven identifizieren lassen, so brächte die Identifikation eines Patienten als Responder oder Non-Responder schon zu diesem Zeitpunkt weitere Anhaltspunkte über den weiter zu erwartenden zeitlichen Verlauf. Diese Fragestellung erfordert intensiven Einsatz explorativer Analysemethoden zur Darstellung der einzelnen Verläufe.

4.2 Fragenblock zur Anwendung von Therapie-Response in einem QM-System

1. Welche allgemeinen Anforderungen bestehen für ein verlaufsorientiertes QM?

In Abschnitt 2.1.1 wurden vier grundlegende Anforderungen an ein QM-System aufgelistet. In einem System mit dem Anspruch, den bisherigen Verlauf einer noch nicht abgeschlossenen Therapie zu bewerten, haben diese Anforderungen Implikationen hinsichtlich der Organisation, des Ablaufes, der Instrumentenauswahl und der Vergleichsstandards. Diese werden in Abschnitt 11.1 diskutiert werden.

2. Wie können sinnvolle Verlaufsstandards erstellt werden, die Therapie-Response berücksichtigen?

Der Vergleich des frühen Symptomverlaufes einer Therapie mit Normen, die aus einer Population von Nicht-Patienten stammen, hat nur beschränkte Aussagekraft, da er lediglich die Entfernung des momentanen Zustandes zum Ziel angibt. Wünschenswert wären Vergleichsstandards, die Auskunft darüber geben, ob die Therapie ‘auf dem richtigen Weg’ ist, also eine Entscheidung darüber erlauben, ob der Verlauf für eine weitere Verbesserung spricht oder ob er sich eher im Rahmen dessen bewegt, wie es negativ verlaufende Therapien tun. Es geht also um die Frage, wie Standards oder Normen für *den Verlauf* einer Therapie entwickelt werden können, im Gegensatz zu Standards, die den Erfolg zum Behandlungsende bewerten. Insbesondere wird dabei die Rolle des Therapie-Response für die Konstruktion dieser Standards zu klären sein. In Abschnitt 12.1 wird die entsprechende Grundidee und methodische Herangehensweise dazu erläutert.

3. Wie könnte der Vergleich von aktuellen Therapieverläufen mit den Verlaufsstandards im Alltag aussehen?

Aus der Beantwortung der vorhergehenden Frage soll an einigen Beispielverläufen demonstriert werden, wie ein Vergleich des aktuellen Therapieverlaufs mit den verlaufsorientierten Standards in der Praxis aussehen kann. Anhand der vorliegenden Stichprobe werden die Verlaufsstandards erstellt, wobei einige Patienten für den nachfolgenden Vergleich zurückbehalten bleiben – sie dienen anschließend als Patienten, an die die Standards angelegt werden. An ihnen kann demonstriert werden, welche Informationen aus der Rückmeldung über den Vergleich ‘bisheriger Verlauf’ und ‘Verlaufsstandard’ abzuleiten sind.

4.3 Auswertungsstrategien - Methodik

Zur Beantwortung des ersten Frageblocks werden vor allem Verfahren der kategorialen Analyse benutzt, deren abhängige Variable aus zwei Kategorien besteht, insbesondere sind dies die logistische Regression und Odds-Ratios (Hosmer & Lemeshow 1989). Die Auswertung erfolgt in mehreren Schritten: der erste Schritt besteht darin, für die vorgeschlagenen Operationalisierungen anhand von Kreuztabellen und einfachen Odds-Ratios den Zusammenhang zwischen Therapie-Response und Therapieerfolg zu bestimmen (Frage 2 im ersten Frageblock). Im zweiten Schritt wird der gleiche Zusammenhang mit Hilfe von schrittweiser logistischer Regression untersucht. Dies geschieht über den Vergleich zweier hierarchischer Modelle: im ersten Modell werden schrittweise die zur Vorhersage des Therapieerfolgs relevanten Variablen der Eingangsinformation ausgewählt. Die Resultate dieses Modells werden im Anschluß daran mit einem zweiten Modell verglichen, das alle Variablen des ersten Modells zuzüglich des Therapie-

Response enthält. Der Vergleich dieser beiden Modelle gibt Aufschluß darüber, ob sich 1) die Vorhersage des Therapieerfolgs verbessert und 2) ob sich Therapie-Response auch *nach* der Kontrolle relevanter Eingangsvariablen als bedeutsamer Prädiktor für den Therapieerfolg erweist (vgl. Frage 4 im ersten Frageblock). Dieses Vorgehen wird sowohl mit dem Therapieerfolg zu Behandlungsende als auch mit dem Therapieerfolg nach 2.5 Jahren angewandt (vgl. Frage 2 im ersten Frageblock). Die fünfte Frage im ersten Frageblock wird ebenfalls über eine logistische Regression angegangen; in diesem Fall ist Therapie-Response die abhängige Variable ('Response', 'kein Response').

Die Untersuchung der Therapieverläufe in Abhängigkeit vom Therapie-Response (Frage 6 im ersten Frageblock) geschieht unter Anwendung explorativer Analyseverfahren (Cleveland 1993), vor allem mit Boxplots und Panel-Plots. Dazu werden möglichst viele Einzelverläufe einer Gruppe in einem Bild auf der gleichen Skala dargestellt, um so Unterschiede und Gemeinsamkeiten in der Verlaufsgestalt besser entdecken zu können.

Der zweite Fragenblock erfordert eine Analyse der Therapieverläufe anhand von Mehrebenen-Regressionsverfahren ('Hierarchische Lineare Modelle', Bryk & Raudenbush 1992, Goldstein 1995), in denen die einzelnen Verläufe durch mathematische Modelle und Parameter dargestellt werden (Ebene 1). Die individuellen Verlaufsparameter können in diesen Verfahren auf der nächsten Ebene hinsichtlich ihres Zusammenhangs mit anderen Kovariaten untersucht werden. Da dieser Methode eine wichtige Rolle bei der Konstruktion der Verlaufsstandards zukommt, wird sie soweit nötig in Abschnitt 12.1.2 näher erläutert werden.

Anmerkung zur Bedeutung von Signifikanztests: Der Einsatz von statistischen Methoden geschieht hier nicht im strengen Sinne zum Testen klar spezifizierter Hypothesen. Die vorliegenden Fragestellungen sind zu unspezifisch und die Theorien erlauben keine genauere Festlegung von Effektstärken, als daß sie in exakte Hypothesen zu überführen wären. Hinzu kommt, daß die verfügbaren Daten nicht explizit zur Beantwortung dieser Fragen erhoben wurden und daher deren Anforderungen nicht optimal erfüllen. So ließen sich zumindest aus dem Phasenmodell der Psychotherapie durchaus sehr konkrete Hypothesen ableiten, da aber die Stichprobe nur aus Patienten mit Eßstörungen besteht, die Daten aus stationärem Setting stammen und keine standardisierten Erhebungsinstrumente zur Erfassung des Therapieverlaufs benutzt wurden, eignen sich diese Daten nicht zur Überprüfung dieser Hypothesen. In solchen Situationen ist der Signifikanztest wegen der Gefahr der Fehlinterpretation kritisiert worden (Cohen 1990; Cohen 1994). Wie von Cohen empfohlen, sollen daher in dieser Arbeit verstärkt Methoden der grafischen Datenanalyse (Cleveland 1993) zum Einsatz kommen, bzw. die Ergebnisse mit Konfidenzintervallen präsentiert werden.

TEIL II: DIE VALIDIERUNG VON THERAPIE-RESPONSE

5 Beschreibung der Studie

Die Daten zu dieser Arbeit stammen aus der *Multizentrischen Studie zur psychodynamischen Therapie von Essstörungen (MZ-ESS)*, durchgeführt von der Forschungsstelle für Psychotherapie, Stuttgart (Kächele 1999). Obwohl diese Studie nicht explizit zur Untersuchung der vorliegenden Fragestellung durchgeführt wurde, lassen die erhobenen Informationen über den Therapieverlauf in der Behandlungsepisode und des Therapieergebnisses die Untersuchung von Therapie-Response zu.

5.1 Versuchsplan und Stichprobe

5.1.1 Versuchsplan

Die Studie wurde als eine longitudinale Anwendungsbeobachtung konzipiert. Im Mittelpunkt steht die Modellierung der Aufwand-Ergebnis-Beziehung der klinischen Praxis, daher folgen die Behandlungen weder einem standardisierten Manual, noch wurden die Patienten zufällig auf die verschiedenen Einrichtungen verteilt. Gegenstand ist meist eine stationäre, in einigen Zentren auch eine ambulante Behandlungsepisode. Die Benennung 'psychodynamische Therapie' bezeichnet zuallererst das Selbstverständnis der Therapeuten. Das Studiendesign gibt Abbildung 3 in schematischer Form wieder. Die

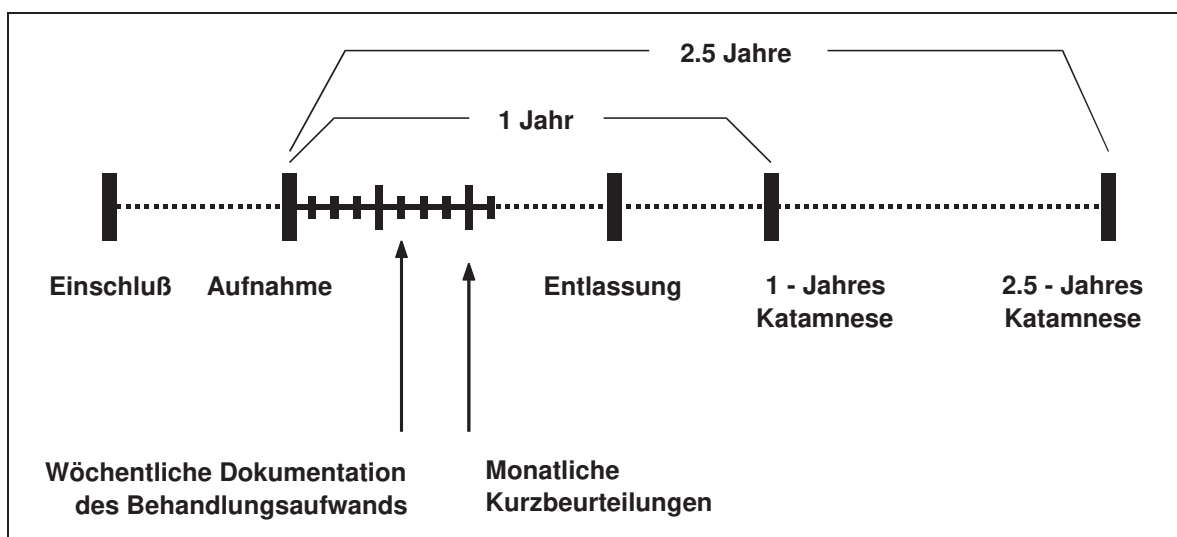


Abbildung 3: Schematische Darstellung des Studiendesigns

Datenerhebung erfolgte nach einem festen Zeitschema über einen Beobachtungszeitraum von 2,5 Jahren. Zu Behandlungsbeginn und -ende wurden jeweils detaillierte

Erhebungen zum Krankheitsstatus durchgeführt. Während der Behandlung erfassen standardisierte Kurzbeurteilungen monatlich den Krankheitsstatus. Wegen des Beobachtungscharakters der Studie variiert die Behandlungsdauer von Patient zu Patient. Ein Jahr nach Beginn der Behandlung wurde eine postalische Nachbefragung durchgeführt und 2.5 Jahre nach Behandlungsbeginn zu einem ausführlichen Katamnese-Interview eingeladen.

In die Studie wurden alle erwachsenen Patientinnen eingeschlossen, die zum Zeitpunkt des Screenings die Kriterien für die Diagnosen Anorexia Nervosa und/oder Bulimia Nervosa nach DSM-III-R erfüllten. Insgesamt liegen Behandlungsepisoden von 1171 Patientinnen (AN=355, BN=647, AN/BN=169) aus 43 Kliniken des gesamten Bundesgebietes vor (außerdem auch noch 215 Behandlungsverläufe aus der Pilot-Phase der Studie).

5.1.2 Erhebungsinstrumente

Der Krankheitsstatus wurde multidimensional und multiperspektivisch zu den 4 Haupterhebungszeitpunkten gemessen, d.h. auf den Dimensionen Symptomatik, Persönlichkeit und Familie/soziales Umfeld. Zur Erfassung der Symptomatik aus Patientenperspektive dienten die Symptom-Checklist-90 (revidierte Fassung, SCL-90-R, Franke 1995), das Eating Disorder Inventory (EDI, Thiel & Paul 1988) und aus Therapeutenperspektive die Erhebungen zum Krankheitsstatus und die Morgan-Russell-Skalen (Morgan & Hayward 1988). Jeweils zu Beginn und am Ende der Behandlung füllten die Patienten einen umfangreichen Fragebogen zu ihrem symptomatischen Status aus und nahmen an einem Interview mit dem Therapeuten teil, welches ebenfalls den gesundheitlichen Status erhob. Dieses Interview orientiert sich an den Fragen des 'Eating Disorder Examination' (EDE, Fairburn & Cooper 1993), das als Standard-Instrument zur Diagnostik bei Untersuchungen von Essstörungen zur Anwendung kommt. Zentrales Instrument in der 2.5-Jahres-Katamnese ist das für die Studienzwecke adaptierte Longitudinal Interval Follow-up Evaluation Interview (Keller et al. 1987), das die Identifizierung von Änderungszeitpunkten ('change points') und damit die Beschreibung des Verlaufs zwischen den fixen Erhebungszeitpunkten zu Therapieende und 2.5-Jahres-Katamnese erlaubt. Diese Instrumente dienen zur Feststellung des Therapieerfolges; als Instrument zur Erfassung des Therapieverlaufs – und damit von Therapie-Response – wird die Kurzbeurteilung eingesetzt, welches die Schwere und Häufigkeit der symptomatischen Diagnosekriterien (nach DSM-III-R, Wittchen 1989) erhebt.

5.1.3 Stichprobenbeschreibung

An dieser Stelle folgen die Angaben über die dieser Arbeit zugrundeliegenden Stichprobe. Zur Modellierung des Zusammenhangs von Therapie-Response und Therapieerfolg

können nur solche Behandlungsverläufe herangezogen werden, die eine genügend lange Behandlungsdauer hatten. Es muß sichergestellt sein, daß a) mindestens *eine* Kurzbeurteilung während der Behandlungszeit vorhanden ist, um Therapie-Response überhaupt erfassen zu können und b) diese weit genug vom Ende der Behandlung entfernt ist (mindestens 2 Wochen), um noch von einer 'Vorhersage' des Therapieerfolges sprechen zu können. Patientinnen, welche die Kriterien beider Diagnosen gleichzeitig erfüllen, finden wegen der relativ kleinen Fallzahl ($N=169$) und des getrennten Symptomschwere-Scores für Anorexie und Bulimie (siehe Abschnitt 6.3, Tabelle 4) in dieser Arbeit keine Berücksichtigung.

Von 67 Patienten war keine Kurzbeurteilung während der Behandlung vorhanden und bei weiteren 65 Patienten war der Abstand zwischen der ersten Kurzbeurteilung und Behandlungsende geringer als 2 Wochen, sodaß zur Analyse des Therapie-Response nach diesen Ausschlußkriterien noch 870 Patienten in Frage ($AN=309$, $BN=561$) kommen. Die Ausschlußrate unterscheidet sich nicht zwischen den beiden Diagnosen: jeweils 13% mußten aufgrund der o.g. Kriterien ausgeschlossen werden. Die ausgeschlossenen Patientinnen unterschieden sich nicht hinsichtlich der Eingangsvariablen, aber zwangsläufig hinsichtlich der Behandlungsdauer: sie hatten notwendigerweise kürzere Behandlungen.

Tabelle 1 faßt die wichtigsten Angaben für den Behandlungsbeginn dieser reduzierten Stichprobe zusammen. Aufgrund des für die Studie geforderten Mindestalters von 18 Jahren liegt das Durchschnittsalter recht hoch. Die Tabelle verdeutlicht auch den relativ schweren Beeinträchtigungs- und Chronifizierungsgrad der Patientinnen: sie sind im Durchschnitt schon mehr als 3 (AN) bzw. 6.5 (BN) Jahre erkrankt und mehr als zwei Drittel haben schon mindestens eine, und mehr als die Hälfte schon mindestens zwei Behandlungen wegen einer Eßstörung hinter sich. Die durchschnittlichen Skalenwerte des SCL-90-R belegen eindrücklich die Beeinträchtigungsschwere: sie liegen für einen Großteil der Patienten weit über dem für Nicht-Patienten zu erwartenden Bereich (Mittelwert und Standardabweichung bei Nicht-Patienten 0.3 und 0.4). Der durchschnittliche BMI bei Anorexie liegt mit 15.1 etwas höher als in anderen Untersuchungen (z.B. Strober, Freeman & Morrell 1997). Die Angaben zum Maximum zeigen, daß es schon zu Behandlungsbeginn einige wenige Patienten mit normalem BMI gibt – dies läßt sich durch die Wartezeit zwischen Screening und Behandlungsbeginn erklären, in dem schon Besserungen möglich sind.

Die Behandlungsdauer liegt im Durchschnitt bei gut 12 Wochen. Die Verteilung ist hier deutlich linkssteil, d.h. die meisten Behandlungen dauerten nicht länger als 14-15 Wochen (75%-Perzentil der Verteilung), der die übrigen 25% verteilen sich auf die längeren Behandlungsdauern.

Tabelle 1: Kennwerte der Stichprobe zu Behandlungsbeginn

	Anorexie (N=309)			Bulimie (N=561)		
	\bar{x}	(SD)	min.-max.	\bar{x}	(SD)	min.-max.
Alter	24.94	(5.65)	17.5 – 52.3	25.99	(6.27)	17.7 – 56.8
Erkrankungsdauer (Jahre)	3.72*	(–)	0.2 – 31.8	6.60*	(–)	0.1 – 34.3
Behandlungsdauer (Wochen)	11.86*	(–)	4 – 58.8	11.86*	(–)	4.7 – 63.1
AU-Tage (letztes Jahr)	28*	(–)	0 – 365	20*	(–)	0 – 365
BMI	15.10	(1.91)	9.8 – 21.3	21.45*	(–)	13.5 – 51.7
% Erwartungs-BMI	71.48	(9.04)	47.2 – 101.4	101.1*	(–)	64.5 – 224.7
EDI 'Drive for thinness'	9.12	(6.25)	0 – 21	11.48	(5.22)	0 – 21
EDI 'Bulimia'	0.0*	(–)	0 – 18	10.36	(4.51)	0 – 21
SCL-90: GSI	1.41	(0.68)	0.2 – 3.6	1.45	(0.66)	0.1 – 3.6
	N	%		N	%	
Vorbehandlungen						
mindestens 1	222	71.8%		430	76.6%	
mindestens 2	175	55.6%		337	60.1%	
Eßanfälle (> 2/Woche)	52	16.8%		502	89.5%	
Erbrechen (> 2/Woche)	63	20.4%		400	71.3%	
Angst vor Dickwerden (‘ständig’ oder mehr)	169	54.7%		446	79.5%	
Mißbrauchserfahrung	98	33.8%		256	45.7%	

alles Patientenangaben, außer BMI; * Median, wg. linkssteiler Verteilung

6 Die Erfassung von Therapie-Response und Therapieerfolg

In diesem Abschnitt wird beschrieben, mit welchen Instrumenten und anhand welcher Kriterien Therapie-Response und Therapieerfolg definiert wurden. Beide Konzepte – Therapie-Response wie auch Therapieerfolg – beziehen sich dabei inhaltlich auf die Schwere der Symptomatik². Der Unterschied liegt im *Zeitpunkt* der Erfassung: Therapie-Response wird innerhalb der ersten 8 Behandlungswochen und Therapieerfolg am Ende der Behandlung gemessen.

Was die Einschätzung der Symptomschwere angeht, so wird für Anorexie und Bulimie getrennt vorgegangen: Bei Patientinnen der Diagnose ‘Anorexie’ werden die Sympto-

²ausgenommen die Operationalisierung des ‘Wellbeing’ als Therapie-Response (s.u.)

me Untergewicht, Angst vor dem Dickwerden und Körperschemastörung bewertet, bei Patientinnen der Diagnose ‘Bulimie’ die Anzahl der Eßanfälle, Ausmaß der gewichtsreduzierenden Maßnahmen und die Beschäftigung mit der körperlichen Erscheinung. Innerhalb einer Diagnose wird zusätzlich eine Gewichtung durch sog. ‘Schlüsselsymptome’³ vorgenommen: dasjenige Symptom, welches die Störung am ehesten charakterisiert und den beeinträchtigendsten Effekt hat, bekommt ein stärkeres Gewicht als die übrigen. Bei Anorexie ist dies das Untergewicht und bei Bulimie die Anzahl der Eßanfälle; Sinn dieser Gewichtung ist, daß nur dann von Therapieerfolg gesprochen werden kann, wenn sich neben den anderen Symptomen auch das Schlüsselsymptom normalisiert hat – bleibt die Normalisierung des Schlüsselsymptoms aus, so wird unabhängig von den anderen Symptomen auch kein Therapieerfolg gewertet.

6.1 Die Symptombereiche

Bei der Beschreibung der einzelnen Symptombereiche und ihrer Erfassung benutze ich für jeden Bereich eine Abkürzung, um die Darstellung in den folgenden Tabellen übersichtlicher zu gestalten (Tabelle 2).

Tabelle 2: *Abkürzungen der Symptombereiche*

Anorexie	BMI*	Gewichtskriterium (Untergewicht)
	ANG	Furcht vor Gewichtszunahme
	KOE	Störung der Körperwahrnehmung
Bulimie	ESS*	Eßanfälle
	GEW	Gewichtsreduzierende Maßnahmen
	FIG	Übermäßige Beschäftigung mit Figur und Gewicht

* Schlüsselsymptom

6.1.1 Anorexie

Untergewicht (BMI):

Das Untergewicht der Patientin wird über den Body-Mass-Index (BMI) registriert. Er berechnet sich aus dem Gewicht und der Körpergröße (kg/m^2). Laut ICD-10 (Dilling, Mombour & Schmidt 1993) liegt das Gewichtskriterium für Anorexie bei einem BMI unter 17.5. In der MZ-ESS Studie wurde der aktuelle BMI zu dem aufgrund einer Normstichprobe zu erwartenden BMI in Beziehung gesetzt und in einen Prozentwert

³Diese Bezeichnung stammt aus der MZ-ESS Arbeitsgruppe, die die Operationalisierung des Therapieerfolges festlegte.

umgewandelt. Eine Patientin, deren BMI exakt dem der Norm entspricht, hat demnach einen BMI-Prozentwert von 100; ernstzunehmendes Untergewicht beginnt in dieser Einheit bei ca. 85% und weniger. Die Normen stammen aus einer nach Geschlecht und Alter ausdifferenzierten Normstichprobe nicht eßgestörter Personen (Hebebrand et al. 1994) .

Furcht vor Gewichtszunahme (ANG):

Dieser Bereich wird über die Subskala 1 des Eating Disorder Inventory (EDI) ‘Drive for Thinness’ erfaßt. Die Fragen dieser Skala kreisen um das Thema Gewichtszunahme (z.B. ‘Wenn ich ein Pfund zunehme, habe ich Angst’, ‘Ich fühle mich schon nach einer kleinen Mahlzeit aufgequollen’ oder ‘Ich habe Angst davor, zuzunehmen’). Ein hoher Score auf dieser Skala (> 4.5) zeigt Furcht vor Gewichtszunahme an.

Zu den Zwischenmessungen während der Behandlung waren nur die Kurzbeurteilungen auszufüllen, deshalb ist die Erfassung hier etwas anders: die Furcht vor dem Zunehmen wird durch eine Frage mit 5 Antwortalternativen erhoben (Skalierung in 5 Stufen von ‘Ich habe keine Angst, zuzunehmen.’ bis ‘Ich möchte eher sterben, als dick sein bzw. zuzunehmen.’, siehe Anhang Seite 131).

Störung der Körperwahrnehmung (KOE):

Für dieses Symptom wurde die Antwort auf die Frage ‘Wie finden Sie Ihren Körper?’ in Beziehung gesetzt zu dem tatsächlichen Körpergewicht. Die Antwortmöglichkeiten waren 5-stufig von ‘viel zu dünn’ zu ‘viel zu dick’. Empfindet eine Patientin sich als ‘viel zu dick’, hat aber einen niedrigen BMI, so klaffen tatsächliches Gewicht und Einschätzung weit auseinander. Entsprechend wurde den Antwortkategorien ein ‘realistischer’ BMI⁴ zugeordnet und vom tatsächlichen BMI subtrahiert. Je größer diese Differenz ausfällt, desto stärker ist die Störung der Körperwahrnehmung.

6.1.2 Bulimie**Eßanfälle (ESS):**

Die Anzahl der Eßanfälle innerhalb der letzten 4 Wochen wurde direkt erfragt; die Patientinnen hatten die Möglichkeit, die in Anzahl pro Tag, pro Woche oder pro Monat anzugeben. Als Einheit dient hier Anzahl der Eßanfälle pro Woche.

Gewichtsreduzierende Maßnahmen (GEW):

Dieser Bereich bezieht sich auf die Häufigkeit von Maßnahmen, die (trotz Eßanfällen) eine Gewichtszunahme verhindern. Im Eingangs- und Entlassungsinterview sind jeweils mehr Maßnahmen erfragt, als in den Kurzbeurteilungen, die nur Erbrechen, Einnahme von Abführmitteln, Diät oder ‘sonstiges’ erfragen. Die Patientinnen konnten zwischen

⁴d.h. es wurde für jede Antwortkategorie ein BMI festgesetzt, der in etwa bei einer realistischen Körperwahrnehmung zu erwarten ist; z.B. ein BMI von 20 für die mittlere Antwortkategorie

mehreren Antwortalternativen, die die Häufigkeit der jeweiligen Maßnahmen vorgeben, wählen: von ‘gar nicht’ bis ‘mehrmals täglich’. Da Erbrechen und Einnahme von Abführmitteln die schädigsten Verhaltensweisen darstellen, wird nur die Häufigkeit dieser Maßnahmen gewertet.

Übermäßige Beschäftigung mit Figur und Gewicht (FIG):

Dieser Bereich muß wegen der Unterschiedlichkeit der Fragebögen am Beginn/Ende der Behandlung und während der Behandlung mit 2 verschiedenen Fragen erhoben werden. Im Aufnahme- und Entlassungsbogen geschieht dies mit der Frage ‘Schenken Sie Ihrer Figur größere Beachtung?’, die mit ‘immer’, ‘oft’, ‘gelegentlich’, ‘selten’ und ‘nie’ beantwortet werden konnte. Die Kurzbeurteilung bietet für dieses Symptom nur eine indirekte Operationalisierung: Unter der Frage ‘Wie finden Sie Ihren Körper?’ konnten die Patienten sich auf 3 Polen in einer 5-stufigen Skala einschätzen: ‘anziehend’ vs. ‘abstoßend’; ‘wohlproportioniert’ vs. ‘unproportioniert’ und ‘weiblich’ vs. ‘unweiblich’. Unter der Annahme, daß Patientinnen, die ihren Körper eher negativ einschätzen, sich auch oft mit ihrer körperlichen Erscheinung beschäftigen, bedeutet hier eine negative Bewertung des Körpers das Vorliegen dieses Symptoms.

6.2 Die Erfassung von Therapieerfolg

In dieser Arbeit kommen die Erfolgskriterien aus unterschiedlichen Perspektiven zur Anwendung, beruhend auf den Interviews bzw. Fragebogen bei Entlassung und Katamnese: zum einen aus der Therapeuten- wie auch aus der Patientenperspektive. Das Erfolgskriterium orientiert sich an den Symptomen und wurde so gewählt, daß weitgehende Symptommfreiheit als Erfolg gewertet wird. Nur Patienten mit geringen Residualsymptomen werden dabei noch als Erfolg beurteilt. Dieses starke Erfolgskriterium markiert einen Status, der keine stationäre Behandlungsnotwendigkeit mehr erfordert.

Die Klassifizierung des Behandlungsergebnisses als ‘Erfolg’ oder ‘kein Erfolg’ in Abhängigkeit vom Ausmaß der (noch) vorhandenen Symptome ist in Tabelle 3 wiedergegeben. Für jeden einzelnen Symptombereich wird eine Grenze festgelegt, die die Normalisierung dieses Symptoms markiert. Die Klassifikation als ‘Erfolg’ geschieht, wenn zu Behandlungsende bzw. Katamnese nurmehr höchstens ein Symptom vorhanden ist, dies darf jedoch nicht das Schlüsselsymptom sein. Therapieerfolg heißt also, daß entweder keine Symptome mehr, oder nur noch eines der beiden ‘Nebensymptome’ (AN: Furcht vor Gewichtszunahme oder Störung der Körperwahrnehmung; BN: Gewichtsreduzierende Maßnahmen oder übermäßige Beschäftigung mit Figur und Gewicht) vorhanden ist und dies auch nicht in extremer Ausprägung.

Die Angaben zu den einzelnen Bereichen stammen aus dem Therapeuten-Interview, welches am Ende der Behandlung mit den Patienten geführt wurde. Die gleichen Infor-

Tabelle 3: Kriterien für Therapieerfolg bei AN und BN

Symptom	Normalisierung bei:	
Anorexie	BMI	mindestens 85% des laut Norm zu erwarteten BMI
	ANG	EDI 'Drive for Thinness' < 4.5
	KOE	'realistischer' vs. tatsächlicher % BMI <= 10%
Bulimie	ESS	keine Eßanfälle
	GEW	kein Erbrechen und Laxantien-Gebrauch
	FIG	'gelegentlich' oder weniger

Anzahl Symptome		Erfolgsbeurteilung
vorhanden		
keines		Erfolg
eines *		Erfolg
zwei		kein Erfolg
drei		kein Erfolg

* jedoch nicht das Schlüsselsymptom!

mationen zur Symptomatik sind auch im Patientenfragebogen bei Entlassung enthalten, folglich unterscheiden sich die Operationalisierungen für Therapieerfolg aus beiden Perspektiven nur unwesentlich.

6.3 Erfassung von Veränderungen der Symptomschwere während der Therapie

Die Operationalisierung von frühem Ansprechen auf die Therapie hat die wiederholte Messung des Patienten *während der Behandlung* zur Voraussetzung. Dazu werden in dieser Untersuchung die Kurzbeurteilungen benutzt, die die Symptomschwere und die subjektive Veränderungsbewertung der Patienten zum jeweiligen Zeitpunkt erfassen. Der Bogen erfragt die wichtigsten Informationen zur Symptomatik, aus denen dann ein Symptomschwere-Index berechnet wird⁵. Dieser Index erlaubt das Aufzeichnen von Veränderungen in der Symptomatik während der Behandlung und so die Operationalisierung von Therapie-Response.

Tabelle 4 gibt die Konstruktion dieses Scores wieder: Aus den in Abschnitt 6.1 be-

⁵ Eine Kopie eines Bogens befindet sich im Anhang. Fragen, die nicht in die Berechnung des Scores aufgenommen wurden, aber dennoch auf dem Bogen vorhanden sind, waren nicht geeignet, einen Beitrag zum Score zu leisten. Die Pole 'muskulös' vs. 'schlaff' bei der Beurteilung des Körpers wiesen keine nennenswerte Variabilität auf. Die Frage zur Regelblutung konnte wegen des hohen Anteils an Patientinnen, die die Pille nahmen, nicht benutzt werden.

Tabelle 4: Einteilung der Symptomschwere in der Kurzbeurteilung

Anorexie	
<i>BMI</i>	<i>% d. erwarteten BMI</i>
(x_1)	0 = n.v. : % BMI > 90 1 = schwach: 80 <= % BMI <= 90 2 = deutlich: 66.5 <= % BMI < 80 3 = extrem: % BMI < 66.5
<i>ANG</i>	<i>Furcht vor Gewichtszunahme (Frage 2)</i>
(x_2)	0 = n.v. : ‘...keine Angst...’ 1 = schwach: ‘...fürchte mich gelegentlich...’ 2 = deutlich: ‘... in ständiger Sorge...’ 3 = extrem: ‘...löst Panik aus...’ / ‘...lieber sterben als ...’
<i>KOE</i>	<i>Differenz tatsächlicher % BMI – ‘realistischer’ % BMI</i>
(x_3)	0 = n.v. : < 10 1 = schwach: 10 bis < 20 2 = deutlich: 20 bis < 30 3 = extrem: > 30
Bulimie	
<i>ESS</i>	<i>Anzahl der Eßanfälle pro Woche</i>
(x_1)	0 = n.v. : keine 1 = schwach: weniger als 2 mal 2 = deutlich: 2 bis 5 mal 3 = extrem: mehr als 5 mal
<i>GEW</i>	<i>Häufigkeit von Erbrechen oder Einnahme von Abführmitteln</i>
(x_2)	0 = n.v. : gar nicht 1 = schwach: bis einmal pro Woche 2 = deutlich: 2 bis 3 mal pro Woche 3 = extrem: täglich und mehr
<i>FIG</i>	<i>negative Einschätzung des Körpers bei: (Frage 3b,3c,3e)</i>
(x_3)	0 = n.v. : keiner der 3 Fragen 1 = schwach: einer der 3 Fragen 2 = deutlich: zwei der 3 Fragen 3 = extrem: allen 3 Fragen
Die beiden Scores für AN und BN	
errechnen sich nach: $(2x_1 + x_2 + x_3)/4$	

n.v. = ‘nicht vorhanden’; vgl. die Beschreibung der Symptombereiche in Abschnitt 6.1

schriebenen Symptomen errechnet sich pro Diagnose ein Symptomschwere-Index, der von 0 (keine Symptome) bis 3 (extreme Symptome) reicht. Jedes Symptom wird in vier Kategorien der Werte 0 bis 3 eingeteilt und die Werte anschließend summiert. Um gelegentliche fehlende Werte auszugleichen, wird die Summe durch die Anzahl beant-

worteter Fragen geteilt; fehlt mehr als ein Symptombereich einer Diagnose, so wird kein Index berechnet. Das Schlüsselsymptom (AN: Untergewicht, BN: Eßanfälle) erhält eine doppelte Gewichtung.

Dieser Score bildet nun zu jedem Erhebungszeitpunkt die Schwere der Symptomatik ab. Trägt man die Scores in ein Koordinatensystem ein mit der Symptomschwere als Y-Achse und der Zeit als X-Achse, so läßt sich der Verlauf eines Patienten grafisch veranschaulichen. Die Veränderung oder Ausprägung dieses Scores zu einem bestimmten Zeitpunkt dient als Grundlage zur Operationalisierung von Therapie-Response.

6.4 Ausgewählte Operationalisierungen des Therapie-Response

Unter Therapie-Response ist nach Abschnitt 3.1 ein ‘Zwischenzustand’ zwischen vollem Syndrom und minimaler Symptomatik zu verstehen, der während der Behandlung eingetreten ist. Wie in Abschnitt 3.3 angedeutet, sind die Möglichkeiten, Therapie-Response zu definieren, sehr umfangreich und hängen davon ab, welches Ereignis man für den weiteren Therapieverlauf für bedeutsam hält. Response kann auf unterschiedlichen Ebenen der Symptomatik liegen, wie z.B. dem Verlauf des momentanen Wohlbefindens oder der Symptomschwere der einzelnen Kriterien. In diesem Abschnitt werden die einzelnen Operationalisierungen beschrieben, wie sie in dieser Arbeit weiter untersucht werden.

Die drei Operationalisierungen sind:

1. verbessertes Wohlbefinden (subjektive Veränderungseinschätzung) (WOL)

Damit ist die subjektive Einschätzung der Patienten über Verbesserung oder Verschlechterung ihres Zustandes nach vier bis acht Wochen gemeint. Patientinnen, die sich zu diesem Zeitpunkt als ‘viel’ oder ‘sehr viel’ gebessert einschätzen, zählen zu den Respondern. Die Bezeichnung ‘verbessertes Wohlbefinden’ verdeutlicht, das hinter der positiven Beantwortung dieser Frage eine Verbesserung der Befindlichkeit der Patientin angenommen wird. Diese Interpretation lehnt sich an das Phasenmodell von Howard et al. (1993) an, in dem eine Verbesserung des momentanen Wohlbefindens für eine weitere Verbesserung auf symptomatischer Ebene als notwendig beschrieben wird (vgl. Abschnitt 3.2.2).

Dabei bleibt zu berücksichtigen, daß das ‘momentane Wohlbefinden’ hier über eine andere Formulierung der Fragen erhoben wird, als es die Theorie vorsieht – nämlich über eine direkte Verbesserungseinschätzung des Gesamtzustandes, im Gegensatz zu vier Fragen bzgl. des aktuellen Befindens im amerikanischen Original.

2. Moderate Symptome: (CUT)

Damit ist das Vorliegen eines moderaten Ausmaßes an Symptomen nach vier bis acht Wochen gemeint, welches einen nicht mehr akuten Zustand bzw. kein volles Syndrom mehr anzeigt. Er bezeichnet aber nicht unbedingt das gänzliche Verschwinden aller Symptome. Diese Operationalisierung kommt dem bei Frank et al. (1991) beschriebenen Therapie-Response am nächsten. Diese Operationalisierung fokussiert auf den Zustand zum Zeitpunkt der Erhebung – unabhängig vom Eingangsstatus. Die Symptomschwere muß dabei eine bestimmte Schwelle (cutoff) erreichen, daher die Abkürzung ‘CUT’.

3. Verringerung der Symptome: (DIF)

Diese Operationalisierung bezeichnet eine deutliche Verringerung der Symptomatik im Vergleich zum Beginn der Behandlung. Sie ist der vorigen sehr ähnlich, betont allerdings die *Veränderung* der Symptomatik im Vergleich zum Eingangswert um einen bestimmten Mindestbetrag. Es kann als ein ‘In-Gang-Kommen’ der Therapie verstanden werden.

Die in Klammern gesetzten Abkürzungen dienen der übersichtlicheren Bezeichnung des Therapie-Response in Grafiken und Tabellen. Dabei ist zusätzlich der *Zeitpunkt* der Erfassung relevant (vgl. Abschnitt 4, Frage 3 in Fragenblock 1). Diese Unterscheidung kann allerdings nur bei solchen Patientinnen vorgenommen werden, die länger als zehn Wochen in Behandlung waren, da nur bei ihnen Response nach vier und Response nach acht Wochen noch mindestens zwei Wochen vom Behandlungsende entfernt ist. Diese Patientinnen bilden eine Sub-Stichprobe zur Untersuchung des Zeitpunktes von Therapie-Response. Bei ihnen bedeutet ‘CUT4’ folglich das ‘Vorliegen von moderater Symptomatik nach vier Wochen’, ‘CUT8’ selbiges nach acht Wochen; die Nummern der Abkürzungen fallen immer dann weg, wenn es um die gesamte Stichprobe handelt (d.h. alle Patientinnen ab sechs Wochen Behandlungsdauer) und folglich nur Therapie-Response nach vier Wochen gemeint ist.

6.4.1 **Verbessertes Wohlbefinden: Subjektive Veränderungseinschätzung**

Die einzige Frage der Kurzbeurteilung, die nicht direkt auf die Symptome abzielt, besteht in folgender Formulierung: ‘Bitte schätzen Sie ein, wie sehr sich Ihr Befinden im Vergleich zu Beginn der Therapie verbessert/verschlechtert hat’. Die Antwortkategorien waren: ‘gar nicht’, ‘kaum’, ‘mäßig’, ‘viel’, ‘sehr viel’ gebessert bzw. verschlechtert. Es ist also eine direkte Veränderungseinschätzung seitens der Patienten, hinsichtlich ihres Befindens. Sie hat damit eine gewisse Ähnlichkeit mit dem Konzept des ‘momentanen Wohlbefindens’ (current wellbeing) aus dem Phasenmodell von Howard et al.

(1993). Wie schon erwähnt, bietet es sich damit zur Konzeptionalisierung von Therapie-Response an. Die Frage nach der Verbesserung der Befindlichkeit wird im folgenden ganz im Sinne dieses Konstruktes verstanden und interpretiert: Patientinnen, die nach 4 Wochen eine Verbesserung bezüglich ihres Befindens angeben, zählen als ‘Responder’; sie sollten laut Theorie eine höhere Chance auf Therapieerfolg am Ende der Therapie haben als diejenigen, die keine Verbesserung berichten.

Therapie-Response mit Hilfe der Frage nach der Verbesserung des Befindens wird dann registriert, wenn die Patientinnen ihr Befinden nach vier bzw. acht Wochen als ‘viel’ oder ‘sehr viel’ gebessert bezeichnen.

6.4.2 Moderate Symptomatik (Cutoff)

Zeigen Patientinnen nicht mehr das volle Bild der Eßstörung, so kann dies einen ersten Teilerfolg in der Therapie bedeuten. Die Idee ist: auch wenn noch ein einzelnes Symptom recht ausgeprägt vorhanden ist, sollten die anderen Symptome nicht mehr so deutlich vorhanden sein. Mit Hilfe der Kurzbeurteilungen kann ein entsprechender Schwellenwert definiert werden: Dieser Cutoff soll drei Kriterien genügen,

1. Das Schlüsselsymptom sollte nicht ‘extrem’ ausgeprägt sein .
2. Die beiden anderen Symptome dürfen ebenfalls nicht gleichzeitig ‘extrem’ ausgeprägt sein.
3. Es sollte nur noch ein Symptom ‘deutlich’ (oder stärker) vorhanden sein.

Ein Kurzbeurteilungs-Score ≤ 1.25 kommt diesen Bedingungen nahe: Aufgrund der Konstruktion des Scores ist jedoch alleine durch die Festlegung des Cutoffs noch nicht sichergestellt, daß alle drei Kriterien erfüllt sind. Dieser Cutoff gewährleistet zwar per definitionem das Zutreffen der ersten beiden Kriterien: eine extreme Ausprägung des Schlüsselsymptoms oder zwei extrem ausgeprägte Nebensymptome führen zu einem Score von mindestens 1.5. Allerdings wird aus Tabelle 5 (S. 48) deutlich, daß auch ein niedrigerer Score von 1.25 oder weniger trotzdem eines der genannten Kriterien verletzen kann: so kann beispielsweise ein Score von 1.00 dadurch zustande kommen, daß beide Nebensymptome ‘deutlich’ ausgeprägt sind (3 Bedingung nicht erfüllt). Um zu überprüfen, ob das dritte Kriterium nicht zu häufig verletzt wird, sind in Tabelle 5 die Häufigkeiten aller Symptomkombinationen der Scores 0 bis 1.25 aufgelistet. Es handelt sich um alle vorliegenden Messungen über alle Zeitpunkte hinweg, bei denen keines der Einzelsymptome fehlte (95% aller Messungen). Wie Tabelle 5 zu entnehmen, ist nur bei 4.6% der Anorexien das dritte Kriterium nicht erfüllt, bei den Bulimien sind es nur 1%. Auch der Anteil der Messungen mit einem extrem ausgeprägten Symptom

Tabelle 5: Symptomkombinationen der Kurzbeurteilungs-Scores

Anorexie (N=612)						Bulimie (N=1468)				
BMI	ANG	KOE	%	n	Score	ESS	GEW	ZUF	%	n
0	0	0	2.0	12	0.00	0	0	0	25.7	377
-	-	-	-	-	0.25	0	0	1	6.4	94
0	1	0	9.5	58	0.25	0	1	0	3.4	50
-	-	-	-	-	0.50	0	0	2	4.9	72
0	1	1	0.8	5	0.50	0	1	1	1.2	18
0	2	0	3.1	19	0.50	0	2	0	0.5	8
1	0	0	6.0	37	0.50	1	0	0	6.6	97
-	-	-	-	-	0.75	0	0	3	3.1	45*
0	1	2	1.5	9	0.75	0	1	2	1.1	16
0	2	1	1.5	9	0.75	0	2	1	0.7	11
0	3	0	0.7	4*	0.75	0	3	0	0.4	6*
1	0	1	4.6	28	0.75	1	0	1	3.5	52
1	1	0	6.7	41	0.75	1	1	0	14.6	214
0	1	3	0.2	1*	1.00	0	1	3	0.6	9*
0	2	2	2.9	8**	1.00	0	2	2	0.3	5**
0	3	1	0.2	1*	1.00	0	3	1	0.1	2*
-	-	-	-	-	1.00	1	0	2	2.8	41
1	0	3	0.2	1*	1.00	-	-	-	-	-
1	1	1	20.4	125	1.00	1	1	1	4.7	69
1	2	0	1.8	11	1.00	1	2	0	2.2	32
2	0	0	9.5	58	1.00	2	0	0	2.9	43
0	2	3	0.5	3**	1.25	0	2	3	0.1	2**
0	3	2	1.1	7**	1.25	0	3	2	0.5	7**
-	-	-	-	-	1.25	1	0	3	2.2	33*
1	1	2	2.9	18	1.25	1	1	2	3.2	47
1	2	1	6.9	42	1.25	1	2	1	0.8	12
1	3	0	1.5	9*	1.25	1	3	0	1.0	14*
2	0	1	6.7	41	1.25	2	0	1	2.3	34
2	1	0	9.0	55	1.25	2	1	0	4.0	58

BMI – Untergewicht

ANG – Angst vor dem Dickwerden

KOE – Körperschemastörung

ESS – Eßanfälle

GEW – gewichtsreduz. Maßnahmen

ZUF – Zufriedenheit mit d. Körper

Symptome: 0 = nicht vorhanden, 1 = mild, 2 = deutlich, 3 = extrem

** mehr als 1 Symptom deutlich vorhanden: AN 4.6%, BN 1%

* extremes Symptom: AN 4.3%, BN 8%

hält sich in Grenzen: 4.3% bei AN und 8% bei BN. Bei beiden Diagnosen sind insgesamt bei weniger als 10% die Bedingungen verletzt, so daß der Cutoff-Wert von 1.25 plausibel erscheint.

6.4.3 Verringerung der Symptome ('In-Gang-Kommen' der Therapie)

Die zweite Operationalisierung von Therapie-Response zielt auf die *Veränderung* der Symptome ab. Die Abnahme der Symptome wird hier als ein 'In-Gang-Kommen' der Therapie interpretiert, das auf eine weitere Abnahme oder Stabilisierung der Symptome auf niedrigem Niveau hinweisen könnte. Dazu muß eine *kritische Differenz* (Δ_{krit}) festgelegt werden, die eine genügend große Veränderung zwischen zwei Zeitpunkten (x_0, x_1) bedeutet.

Bei der Veränderungsmessung mit normierten psychometrischen Fragebögen wird dazu in der Regel das Konzept der reliablen Veränderung ('reliable change', nach Jacobson & Truax 1991; Kordy & Senf 1985) oder der kritischen Differenz herangezogen. Dazu werden insbesondere die Angaben zur Retest-Reliabilität benötigt. Da für die Kurzbeurteilungen der MZ-ESS Studie keine entsprechenden Normierungen vorliegen, muß ein anderer Weg eingeschlagen werden: nämlich die Nutzung der Behandlungsverläufe aus der Pilotstudie ⁶.

Um eine 'kritische Differenz' – also das Ausmaß der Veränderung zu bestimmen, das auf einen erfolgreichen weiteren Verlauf schließen lassen könnte, wurde die durchschnittliche anfängliche Veränderung von solchen Patientinnen ermittelt, die nach vier Wochen den Schwellenwert für moderate Symptomatik von 1.25 oder weniger erreicht hatten und diesen während ihrer Behandlungszeit stabilisieren oder noch weiter verbessern konnten, d.h. bis zum Behandlungsende keinen nachfolgenden Wert höher als 1.25 hatten.

Von insgesamt 238 Patientinnen der Pilotstudie (AN=95, BN=143) war dies bei 32.4% der Patientinnen der Fall (AN=14, BN=63). Die mittlere Veränderung (Median) innerhalb der ersten 4 Wochen lag für Anorexie bei 0.5 Punkten und für Bulimie bei 1.5 Punkten. Demnach muß sich der Anorexie-Kurzbeurteilungsscore einer Patientin der Diagnose Anorexie um mindestens 0.5 Punkte und entsprechend der Bulimie-Kurzbeurteilungsscore einer bulimischen Patientin um mindestens 1.5 Punkte verbessern, um als 'früher Responder' klassifiziert zu werden.

Die Einteilung in Response bzw. Non-Response geschieht also nach folgendem Schema:

$$Response = \begin{cases} 0 & (\text{kein Response}) & : & (x_0 - x_1) < \Delta_{krit} \\ 1 & (\text{Response}) & : & (x_0 - x_1) \geq \Delta_{krit} \end{cases}$$

wobei:

Anorexie: $\Delta_{krit} = 0.5$, Bulimie: $\Delta_{krit} = 1.5$

Tabelle 6 gibt die Verteilung der Differenzwerte pro Diagnosegruppe wieder. Wie man

⁶Eine andere Alternative wäre die Berechnung von Cronbach's Alpha, was allerdings bei nur drei Fragen pro Diagnose-Score wenig sinnvoll ist.

Tabelle 6: Verteilung der Veränderungen der 'frühen Responder'

Differenz	Anorexie		Bulimie	
	Häufigkeit	Prozent	Häufigkeit	Prozent
-0.25	1	7.1	1	1.6
0.00	4	28.6	2	3.2
0.25	1	7.1	3	4.8
0.50	3	21.4	4	6.3
0.75	4	28.6	4	6.3
1.00	–	–	9	14.3
1.25	–	–	5	7.9
1.50	1	7.1	13	20.6
1.75	–	–	4	6.3
2.00	–	–	6	9.5
2.25	–	–	9	14.3
2.50	–	–	1	1.6
2.75	–	–	1	1.6
3.00	–	–	1	1.6

aus der Pilotstudie (N=238)

sieht, verändern sich die frühen Responder der Bulimien wesentlich stärker als die Anorexien; ebenso ist der Anteil an frühen Respondern bei Bulimie höher als bei Anorexie. Zu Therapieende war bei insgesamt 58.4% (AN=40, BN=99) der Pilotstudien-Patientinnen ein Score von 1.25 oder weniger zu verzeichnen, allerdings erreichten die meisten Patientinnen diesen Wert erst nach mehr als 4 Wochen.

7 Ergebnisse

7.1 Häufigkeiten von Therapie-Response

Die in Abschnitt 6 beschriebenen Operationalisierungen für Therapie-Response ergeben die in Tabelle 7 zusammengefaßten Häufigkeiten. Sie zeigt die Responderaten der drei Operationalisierungen nach vier Wochen. Bulimische Patientinnen haben durchweg höhere Responderaten als anorektische Patientinnen, was den symptomatischen Response angeht. Bei AN kommen alle drei Operationalisierungen etwa gleich häufig vor, bei BN ist das Vorliegen einer moderaten Symptomatik mit 57% am häufigsten, die Symptomverringerung läßt sich bei einem Drittel feststellen und das Wellbeing bei 27%. Auffällig ist, daß es bei BN nach vier Wochen mehr Patientinnen mit nur noch moderater Symptomatik gibt als Patientinnen, die eine deutliche Verbesserung gezeigt

Tabelle 7: Häufigkeiten von Therapie-Response

	Anorexie		Bulimie	
	(N=300)		(N=550)	
	N	%	N	%
<i>Response nach 4 Wochen*</i>				
Verbesserung d. Wohlbefindens	74	24.50	150	27.27
moderate Symptome	70	23.30	318	57.82
Verringerung der Symptome	73	24.33	186	33.82

* Behandlungsdauer mindestens 6 Wochen

haben. Das legt die Vermutung nahe, daß ein bestimmter Anteil an Patientinnen nach vier Wochen zwar die Schwelle von 1.25 Punkten erreicht oder unterschritten hat, jedoch war deren Veränderung geringer als 1.5 Punkte und starten folglich auf relativ niedrigem Niveau (z.B. 2.0 Punkte bei Behandlungsbeginn und 1.0 Punkte nach vier Wochen).

Tabelle 8 vergleicht die Response-Raten nach vier bzw. acht Wochen für diejenigen Patientinnen, die mindestens zehn Wochen in Behandlung waren. In dieser Teilgruppe

Tabelle 8:

Häufigkeiten von Therapie-Response nach 4 bzw. 8 Wochen

	Anorexie		Bulimie	
	(N=192)		(N=373)	
	N	%	N	%
<i>Response nach 4 Wochen</i>				
Verbesserung d. Wohlbefindens	44	22.34	90	24.06
Moderate Symptome	47	24.10	203	54.28
Verringerung der Symptome	47	24.10	118	31.55
<i>Response nach 8 Wochen</i>				
Verbesserung d. Wohlbefindens	72	37.70	142	38.27
Moderate Symptome	58	30.21	217	58.18
Verringerung der Symptome	71	36.98	125	33.51

Behandlungsdauer mindestens 10 Wochen

von Patientinnen mit längerer Behandlung sind die Responderaten nach vier Wochen vergleichbar mit den Raten in Tabelle 7. Nach acht Wochen erhöht sich die Anzahl der Responder bei beiden Diagnosen. Der Anstieg ist bei AN etwas größer als bei BN. Der Anteil der Responder erhöht sich bei AN um 6 bis 12% auf 30 bis 36%, die größte Zunahme ist in der Verbesserung des Wohlbefindens zu verzeichnen, hier stieg der Anteil der Responder um 15% auf fast 38%. Bei BN haben nach acht Wochen ist der Anteil der Responder um nur 2-4% gestiegen, außer bei der Verbesserung des Wohlbefindens,

wo der Anteil – wie bei der Anorexie – um 14% höher liegt.

Der deutlichste Unterschied zwischen beiden Diagnosen liegt beim Response als ‘moderater Symptomatik’ (CUT): hier zeigen über die Hälfte der Bulimien nach vier bzw. acht Wochen Response, während bei AN auch nach acht Wochen noch weniger als ein Drittel Therapie-Response haben. Der Anteil der Patientinnen mit einer Verringerung der Symptome steigt bei AN stärker als bei BN, so daß der Anteil an Respondern nach acht Wochen bei AN sogar etwas höher als bei BN ist; dabei jedoch ist die geringere Differenz bei der Definition von Response zwischen AN im Vergleich zu BN zu beachten. Bezüglich der Verbesserung des Wohlbefindens bestehen zwischen den Diagnosen keine deutlichen Unterschiede.

Zusammenfassung:

Therapie-Response kommt bei AN etwas seltener vor (25%) als bei BN (30-55%); nur bei ihnen schwanken die Raten zwischen den verschiedenen Operationalisierungen: eine Verbesserung des Wohlbefindens war nach vier Wochen bei knapp 30%, das Vorliegen moderater Symptomatik dagegen bei gut 55% zu registrieren. Nach 8 Wochen erhöht sich der Anteil der Responder bei AN stärker als bei BN: nach 8 Wochen zeigen unabhängig von der Operationalisierung rund 33% der anorektischen Patientinnen Therapie-Response. Bei der Bulimie erhöhen sich die Response-Raten nur geringfügig, mit Ausnahme der Verbesserung des Wohlbefindens, wo die Rate im gleichen Ausmaß wie bei AN steigt.

7.2 Therapie-Response und Therapieerfolg

7.2.1 Ohne Kontrolle der Eingangsinformationen

Können wir aus dem (frühen) Ansprechen auf die Behandlung einer Patientin ableiten, daß sie am Ende der Therapie erfolgreich abschneidet? Sind solche Patienten, die nach vier oder acht Wochen auf die Behandlung ansprechen, auch diejenigen, die am Ende der Behandlung häufiger Therapieerfolg haben? Eine vorläufige Antwort auf diese Frage erhält man, wenn man sich die Kreuztabellen zwischen Therapie-Response und Therapieerfolg vor Augen führt. Tabelle 9 zeigt die Therapie-Responseraten und die Erfolgsraten. Therapieerfolg ist hier aus den Angaben des Patientenfragebogens entnommen und nach dem in Abschnitt 6.2 beschriebenen Schema definiert. Leichte Abweichungen in den Fallzahlen im Vergleich zu Tabelle 7 kommen durch fehlende Daten bei der Ermittlung des Therapieerfolgs zustande.

Wie schon beim Therapie-Response sind auch beim Therapieerfolg die Erfolgsraten zwischen beiden Diagnosen unterschiedlich: AN knapp 12% und BN 30% bei Behandlungsende. Wie man sieht, haben zwar nur 20% der Anorexien mit Therapie-Response auch am Ende Erfolg, doch sind dies immerhin fast die Hälfte aller beobachteten Therapieerfolge. Von den Anorexien ohne Therapie-Response haben nur 7% Therapieerfolg.

Tabelle 9:
Therapie-Response (Verringerung der Symptome nach 4 Wochen) und Therapieerfolg:

Anorexie				
		Erfolg		
		nein	ja	Σ
Response	nein	211 (92.9)*	16 (7.1)*	227 (75.7)
	ja	59 (80.8)*	14 (19.2)*	73 (24.3)
	Σ	270 (88.3)	30 (11.7)	300
		<i>OR** : 3.13 (1.44, 6.78)</i>		
Bulimie				
		Erfolg		
		nein	ja	Σ
Response	nein	282 (79.2)*	74 (20.8)*	356 (65.9)
	ja	96 (52.2)*	88 (47.8)*	184 (34.1)
	Σ	378 (70.0)	162 (30.0)	540
		<i>OR** : 3.49 (2.37, 5.14)</i>		

* Zeilen-Prozent, ** Odds-Ratio mit 95% Konfidenzintervall

Patientinnen mit Therapie-Response haben mit 19.2% eine höhere Erfolgsrate als Basisrate (in dieser Stichprobe 11.7%) und als Patientinnen ohne Therapie-Response. Der Odds-Ratio von 3.1 besagt, daß Responder eine mehr als drei mal so hohe Wahrscheinlichkeit auf Therapieerfolg haben als Non-Responder⁷.

Bei den bulimischen Patienten ergibt sich ein ähnliches Bild: knapp 48% der Patientinnen mit Therapie-Response haben am Ende Therapieerfolg (mehr als die Basisrate von 30%), dagegen nur gut 21% der Patientinnen ohne Therapie-Response. Responder haben eine 3.7 mal so hohe Wahrscheinlichkeit auf Therapieerfolg am Ende der Behandlung als Non-Responder. Auch dieser Unterschied ist bedeutsam: das Konfidenzintervall liegt bei 2.3 und 6.1.

Im folgenden sind für alle Operationalisierungen solche Kreuztabellen erstellt worden, um die jeweiligen Odds-Ratios zu vergleichen. Die Abbildungen 4 und 5 fassen die Odds-Ratios aller Operationalisierungen von Therapieerfolg für beide Therapieerfolgskriterien (Patienten- und Therapeutenperspektive) zusammen. Die beiden Ab-

⁷Ein Odds-Ratio von 1 bedeutet keinen Unterschied zwischen beiden Gruppen, ein Odds-Ratio von 0.5 bedeutete eine halb so große Erfolgswahrscheinlichkeit der Responder im Vergleich zu den Non-Respondern, ein Odds-Ratio von 2 eine doppelt so hohe Erfolgswahrscheinlichkeit der Responder. Die statistische Bewertung erfolgt hier über das 95% Konfidenzintervall; schließt es die eins *nicht* ein, kann dieser Unterschied als 'statistisch gesichert' gelten (Hosmer & Lemeshow 1989).

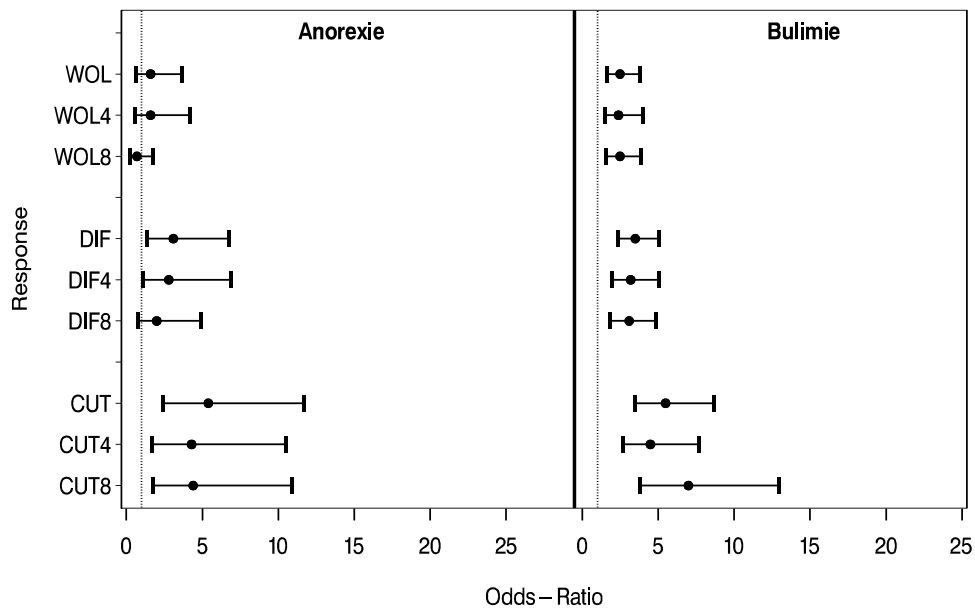


Abbildung 4: Vergleich Responder/Non-Responder (Patientenperspektive)

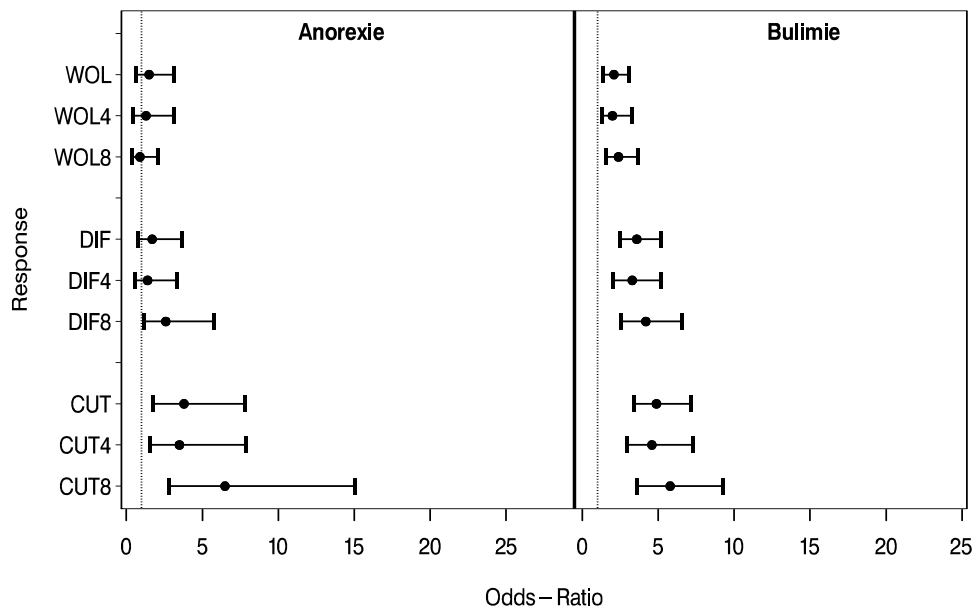


Abbildung 5: Vergleich Responder/Non-Responder (Therapeutenperspektive)

bildungen zeigen die Odds-Ratios (mit 95% Konfidenzintervallen) der verschiedenen Response-Operationalisierungen für AN und BN. Jede Operationalisierung hat 3 Odds-Ratios: die jeweils obere Zeile, zeigt Response nach vier Wochen bei mindestens sechs Wochen Behandlung (z.B. CUT), die mittlere Response nach vier Wochen bei minde-

stens zehn Wochen Behandlung (CUT4) und die jeweils untere Zeile Response nach acht Wochen bei mindestens zehn Wochen Behandlung (CUT8). Die unterbrochene vertikale Linie markiert den Odds-Ratio von eins als Referenz.

Beide Bilder zeigen ein sehr ähnliches Muster. Fast alle Odds-Ratios liegen über 1 und belegen damit durchgängig einen positiven Zusammenhang zwischen Therapie-Response und Therapieerfolg. Bei Anorexie ist er weniger deutlich: bei Response als Verbesserung des Wohlbefindens (WOL) ist der Zusammenhang am niedrigsten und durch den Einschluß der 1 in den Konfidenzintervallen am unsichersten. Die Verringerung der Symptome (DIF) weist einen deutlichen positiven Zusammenhang auf (Odds-Ratios zwischen 2.0 und 3.1), jedoch liegen die Konfidenzintervalle nahe der 1. Das Vorliegen einer moderaten Symptomatik (CUT) hat mit Odds-Ratios zwischen 4 und 5 den stärksten Zusammenhang mit dem Therapieerfolg. Bei Bulimie besteht bei allen drei Operationalisierungen ein deutlich positiver Zusammenhang; er ist beim Wellbeing (WOL) am vergleichsweise niedrigsten (2.4 - 2.5) und bei der moderaten Symptomatik (CUT) am höchsten (5.5 - 7.0). Die Unterschiede *innerhalb* der Operationalisierungen zu den verschiedenen Zeitpunkten fallen bei beiden Diagnosen nur gering aus und lassen keine Vorteil einer Erhebung von Response nach acht Wochen gegenüber Response nach vier Wochen erkennen.

Zusammenfassung:

Der einfache Vergleich von Respondern vs. Non-Respondern hinsichtlich ihres Therapieerfolges zu Behandlungsende macht einen Unterschied zwischen beiden Diagnosen deutlich: Bei BN ist Therapie-Response in *allen drei* Operationalisierungen mit einer höheren Wahrscheinlichkeit auf Therapieerfolg verbunden. In schwächerem Maße gilt dies auch für AN, jedoch ist hier ein verbessertes Wohlbefinden *nicht* mit einer erhöhten Wahrscheinlichkeit auf Erfolg verbunden. Unabhängig von der Diagnose ist der Effekt am stärksten ausgeprägt für die Cutoff-Definition von Response - am schwächsten für das verbesserte Wohlbefinden.

7.2.2 Mit Kontrolle der Eingangsinformationen

Schritt1: Vorhersage des Therapieerfolgs aufgrund der Eingangsinformation Die bisher zusammengetragenen Ergebnisse reichen noch nicht aus, um den Effekt des Therapie-Response zu belegen. So könnte es durchaus sein, daß andere Variablen zu diesem Effekt beitragen. So könnten z.B. gerade Patientinnen, die weniger schwer erkrankt sind oder weniger komplizierende Begeleitumstände mitbringen, auch diejenigen sein, die 1) von vorneherein ein günstigere Prognose für Therapieerfolg haben und 2) daher auch eher Therapie-Response zeigen. Unter solcher Konstellation wäre mit dem Konzept des Therapie-Response nichts gewonnen, das mit der Ausgangsinformation der Patientinnen nicht schon bekannt wäre. D.h. an dieser Stelle ist zu fragen, ob

- der Einfluß von Therapie-Response auf die Erfolgswahrscheinlichkeit auch nach der Kontrolle anderer wichtiger Variablen noch bestehen bleibt und
- ob durch die Hinzunahme von Therapie-Response zu den Prädiktoren vom Beginn der Behandlung eine Verbesserung der Vorhersage über den Therapieerfolg erreicht werden kann (vgl. Abschnitt 3.3.3).

Das Vorgehen geschieht dabei folgendermaßen: Zunächst wird mit Hilfe von schrittweiser logistischer Regression ein Ausgangsmodell *ohne* Therapie-Response zur Vorhersage des Therapieerfolgs gebildet. Für das schrittweise Hinzufügen (forward selection) relevanter Prädiktoren in das Ausgangsmodell wurde als Kriterium ein Signifikanzniveau von $p < 0.1$ festgelegt. Zu diesem Modell wird in einem zweiten Schritt die jeweilige Response-Operationalisierung als Kovariate hinzugefügt.

Die Bewertung des Einflusses von Therapie-Response erfolgt an zwei Kriterien: zum einen anhand des bedingten Odds-Ratios, der den Unterschied in der Erfolgswahrscheinlichkeit zwischen Respondern und Non-Respondern angibt und zum anderen aufgrund der Anzahl der durch das Modell richtig klassifizierten Patientinnen. Anzeichen für einen positiven Zusammenhang zwischen Therapie-Response und Therapieerfolg liegen dann vor, wenn der bedingte Odds-Ratio für Response größer als eins ist und die Anzahl der aufgrund der Kovariaten korrekt klassifizierten Patientinnen im Modell *mit* Therapie-Response höher ist als im Modell ohne Response.

Liste der Prädiktoren:

Tabelle 10 zeigt die Liste der klinisch relevanten Kovariaten. Sie ist das Ergebnis eines umfangreicheren Auswahlprozesses, mit dem Ziel, die für Eßstörungen relevanten Faktoren zu berücksichtigen und gleichzeitig das Ausmaß an Kollinearität gering zu halten. Die einzelnen Prädiktoren (ohne Therapie-Response) korrelieren höchstens zu 0.40 miteinander. Sie beinhalten sowohl Patienteneigenschaften, Behandlungsparameter (Dauer und Intensität der Therapie) als auch Klinikeigenschaften. Die Intensität der Behandlung wurde in zwei Variablen festgehalten: zum einen die Anzahl der Stunden einer Patientin in psychotherapeutischer Einzel- oder Gruppensitzungen ('Standard') und zum anderen die Anzahl der Stunden in anderen, die Behandlung ergänzenden Maßnahmen wie Bäder und Massagen, Sport, Mal- und Ergotherapie etc... Dabei gehen nur die Informationen aus Woche 3 und 4 der Behandlung ein, unter der Annahme, daß das in diesen Wochen angebotene Programm nach den Aufnahmeverfahren weitgehend konstant bleibt.

Die Ausgangsmodelle für AN und BN:

Die Modellierung des Therapieerfolges an dieser Stelle hat nicht zum Ziel, die Relation zwischen Aufwand und Ergebnis eingehend zu untersuchen - dies ist bereits an anderer Stelle geschehen (Kächele 2000); im Zentrum des Interesses steht die statistische Kontrolle möglicher Einflüsse, um den Effekt des Therapie-Response besser abschätzen zu

Tabelle 10: *Prädiktoren für Therapieerfolg*

Prädiktor	Operationalisierung
Patientenvariablen	
Motivation	0 = nicht motiviert 1 = motiviert
Alter	Jahre
Beeinträchtigung durch Eßstörung	Patienten-Rating 1 =niedrig bis 5 =hoch
Untergewicht	%-Anteil des aufgrund der Norm erwarteten BMI
Eßstörungssymptomatik	EDI, Summe der Subskalen 1,2,3
Persönlichkeit (1)	FPI-Extremitäts-Score: Quadratsumme der Abwei- chungen von Normwerten über alle Skalen hinweg
Persönlichkeit (2)	Narzissmus-Inventar: Summe aller Items
Vorhandensein von AN bzw. BN Symptomen	0 = keine 1 = nicht alle Kriterien erfüllt 2 = alle Kriterien erfüllt
Arbeitsunfähigkeit	Anzahl der Tage im letzten Jahr*
Multi-Impulsivität	Rating aus Angaben zu Drogenkonsum, Aggressivität und Impulsivität
Krankheitsdauer	Jahre*
Mißbrauchserfahrungen	0 = kein Mißbrauch bekannt 0.5 = Verdacht auf Mißbrauch 1 = Mißbrauch vorhanden
Psychische Symptome	SCL-90-R Gesamt-Index (GSI)
Beziehung zu Familie	Subskala 'Socioeconomic State' aus Morgan-Russell Skalen
Vorbehandlung	Anzahl ambulant oder stationär (keine bis 3 oder mehr)
niedrige Anfangssymptome	Score der Kurzbeurteilung zu Behandlungsbeginn (≤ 1.25)
Behandlungsvariablen	
Therapiedauer	in Monaten*
Therapie-Intensität (Standard)	aus Woche 3 und 4 der Behandlung
Therapie-Intensität (Zusatz)	aus Woche 3 und 4 der Behandlung
Klinikvariablen	
Klinikategorie	1 = kein spezielles Eßstörungsprogramm 2 = spezielles Eßstörungsprogramm 3 = eßstörungsspezifische Station 4 = eßstörungsspezifische Klinik

* wegen schiefer Verteilung logarithmiert

können. Tabelle 11 zeigt das Ausgangsmodell für Anorexie mit Therapieerfolg aus der Patientenperspektive als abhängiger Variable ohne Therapie-Response als Prädiktor. Sie gibt die relevanten Prädiktoren und deren Zusammenhang zum Therapieergeb-

Tabelle 11: *Modellierung des Therapieerfolgs, Anorexie (Ausgangsmodell)*

Variable	Koeffizient	se	Wald- χ^2	p	Odds-Ratio	Einheit
Behandlungsdauer	3.03	0.63	23.38	0.0001	– *	
Untergewicht	0.14	0.03	15.54	0.0001	3.5 (1.9, 6.4)	9
Wunschgewicht	0.08	0.04	4.08	0.0434	2.0 (1.0, 3.9)	9
Motivation	1.03	0.54	3.65	0.0560	2.8 (1.0, 8.0)	1
EDI	-0.06	0.02	5.22	0.0223	0.7 (0.6, 1.0)	5
SCL90 GSI	-0.88	0.45	3.81	0.0510	0.6 (0.4, 1.0)	0.5
privat vs. sonst.	-0.35	0.62	0.31	0.5778	0.7 (0.2, 2.4)	1
BfA vs. sonst.	-0.76	1.23	0.38	0.5374	0.5 (0.0, 5.2)	1
DAK vs. sonst.	-0.85	1.16	0.53	0.4657	0.4 (0.0, 4.2)	1
AOK vs. sonst.	1.85	0.73	6.43	0.0112	6.3 (1.5, 26.4)	1

$R^2 = 0.46$, $-2\text{LogLikelihood} = 119.7$

* wegen der Log-Transformation ist hier der Odds-Ratio nicht sinnvoll

nis wieder. Die Koeffizienten geben über die Richtung des Zusammenhanges Auskunft und aus den Odds-Ratios in der Verbindung mit der angegebenen Einheit läßt sich die Größe abschätzen. Die Zusammenhänge hier sind wie folgt: je schwerer das Eingangsgewicht und Wunschgewicht und je länger die Behandlungsdauer, desto eher war ein Therapieerfolg wahrscheinlich. Eine gute Behandlungsmotivation trug ebenfalls zu einem besseren Therapieergebnis bei. Mehr Krankheitswert auf den Skalen des EDI und des SCL-90 war mit schlechteren Erfolgsaussichten verbunden. Patientinnen mit einem um neun %-Punkte höheren Eingangsgewicht als andere Patientinnen hatten eine um 3.4 mal höhere Wahrscheinlichkeit auf Therapieerfolg; motivierte Patientinnen eine 2.7 mal höhere als nicht motivierte Patientinnen (bei gleichen Werten auf den anderen Variablen). Patientinnen mit einem um 0.5 Punkte höheren SCL-90 Wert hatten dagegen eine um fast die Hälfte (Odds-Ratio = 0.6) reduzierte Erfolgswahrscheinlichkeit. Die unteren vier Zeilen der Tabellen geben die Unterschiede der einzelnen Versicherungen im Vergleich zu ‘sonstigen’ Versicherungen an (Dummy-Kodierung). Die Varianzaufklärung liegt bei 46%, was auf eine gute Vorhersagbarkeit des Ergebnisses aufgrund der Prädiktoren deutet.

Das Ausgangsmodell für Bulimie ist in der nächsten Tabelle aufgeführt (Tabelle 12). Hier konnte nur für die Variablen ‘Behandlungsdauer’, ‘Klinikategorie’ und ‘anfängliche Symptomschwere’ ein Zusammenhang mit dem Therapieergebnis gefunden werden. Eine längere Behandlung war mit einer geringeren Erfolgswahrscheinlichkeit verbunden. Alle Patientinnen in Kliniken mit einem spezialisierten Programm zur Behandlung

Tabelle 12: Modellierung des Therapieerfolgs, Bulimie (Ausgangsmodell)

Variable	Koeffizient	se	Wald- χ^2	p	Odds-Ratio
Intercept	1.06	0.81	1.72	0.1895	–
Behandlungsdauer	-0.60	0.29	4.20	0.0405	– *
unspez. vs. spez. Programm	0.69	0.30	5.41	0.0200	2.0 (1.1,3.5)
unspez. vs spez. Abteilung	1.20	0.35	11.58	0.0007	3.3 (1.6,6.6)
unspez. vs spez. Klinik	1.12	0.29	15.33	0.0001	3.1 (1.8,5.4)
anfängliche Symptomschwere	-0.56	0.17	10.60	0.0011	0.6 (0.4,0.8)

$R^2 = 0.08$, $-2\text{LogLikelihood} = 628.5$

* wegen der log-Transformation ist hier der Odds-Ratio nicht sinnvoll

von Eßstörungen hatten eine 2-3 mal so große Wahrscheinlichkeit auf Therapieerfolg wie Patienten in unspezifischem Behandlungssetting. Je niedriger die Symptomschwere (laut Kurzbeurteilung) zu Behandlungsbeginn war, desto höher war auch die Chance auf Therapieerfolg. Die Varianzaufklärung ist mit 8% recht gering. Die Modelle für die Erfolgsoperationalisierung aus Therapeutesicht unterscheiden sich nur geringfügig und werden daher nicht im einzelnen berichtet.

Der ‘Effekt’ der Behandlungsdauer in beiden Modellen erfordert eine ergänzende Bemerkung: aufgrund des quasi-experimentellen Designs kann von *keinem* ursächlichen Einfluß der Behandlungsdauer auf das Therapieergebnis gesprochen werden. Die Beendigung der Therapie ist zum Teil auch vom erreichten Krankheitsstatus abhängig, etwa wenn aufgrund guter Fortschritte eine Patientin schon nach kurzer Dauer entlassen werden kann oder wegen langsamen Fortschritts die Behandlung verlängert wird. Die obenstehenden Zahlen beschreiben daher lediglich einen beobachteten Zusammenhang, dessen ursächliches Zustandekommen ungewiß bleibt. Wie schon erwähnt, dienen die Modelle hier auch nicht der inhaltlichen Erklärung von Effekten, sondern der statistischen Kontrolle von in dieser Stichprobe relevanten Kovariaten von Therapieerfolg.

Zusammenfassung:

Therapieerfolg zu Behandlungsende ließ sich mit Hilfe der zu Beginn der Behandlung verfügbaren Information bei AN besser vorhersagen als bei BN. Mit einer höheren Erfolgswahrscheinlichkeit waren für AN folgende Variablen verbunden: höheres Eingangs- und Wunschgewicht, gute Behandlungsmotivation und längere Behandlungsdauer; außerdem niedrigere Symptombelastung in der SCL-90 und dem EDI. Insgesamt hatten also die leichter erkrankten Patientinnen eine bessere Chance auf Therapieerfolg. Der Anteil aufgeklärter Varianz lag bei 46%.

Das Behandlungsergebnis der Bulimien ließ sich weniger gut vorhersagen (aufgeklärte Varianz: 8%). Auch hier hatten leichter erkrankte Patientinnen eine bessere Chance auf Therapieerfolg; zusätzlich gab es deutliche Vorteile für Patientinnen, die in einer spezialisierten Einrichtung behandelt wurden. Längere Behandlungsdauern tendierten zu niedrigeren Erfolgsraten.

Schritt 2: Vorhersage des Therapieerfolgs aufgrund der Eingangsinformation und des Therapie-Response :

Zu diesen Ausgangsmodellen wurden nun die einzelnen Response-Operationalisierungen zugefügt, um zu überprüfen, ob der Zusammenhang zwischen Therapie-Response und Therapieergebnis auch nach der Kontrolle relevanter Kovariaten erhalten bleibt.

Die nachfolgenden zwei Tabellen geben die Modelle der Vorhersage des Therapieerfolgs aus der Patientenperspektive mit Hilfe der Eingangsinformation und zusätzlich des Therapie-Response (hier: Verbesserung der Symptomatik) wieder. Ein Vergleich

Tabelle 13:

Modellierung des Therapieerfolgs, Anorexie (zuzüglich Therapie-Response)

Variable	Koeffizient	se	Wald- χ^2	p	Odds-Ratio	Einheit
Behandlungsdauer	3.01	0.64	22.18	0.0001	– *	
Untergewicht	0.14	0.04	15.91	0.0001	3.7 (1.9,7.0)	9
Wunschgewicht	0.08	0.04	4.31	0.0378	2.1 (1.0,4.2)	9
Motivation	1.03	0.55	3.53	0.0603	2.8 (1.0,8.2)	1
EDI	-0.05	0.02	4.74	0.0295	0.8 (0.6,1.0)	5
SCL90 GSI	-0.84	0.47	3.11	0.0778	0.6 (0.4,1.0)	0.5
privat vs. sonst.	-0.36	0.63	0.32	0.5723	0.8 (0.2,2.4)	1
BfA vs. sonst.	-0.65	1.25	0.27	0.6047	0.5 (0.0,6.0)	1
DAK vs. sonst.	-0.79	1.17	0.45	0.4999	0.5 (0.0,4.5)	1
AOK vs. sonst.	2.10	0.76	7.59	0.0059	6.8 (1.8,36.2)	1
Response	1.41	0.53	7.15	0.0075	4.1 (1.5,11.5)	1

$R^2 = 0.50$, $-2\text{LogLikelihood} = 112.4$

* wegen der log-Transformation ist hier der Odds-Ratio nicht sinnvoll

von Tabelle 13 mit Tabelle 11 zeigt, daß die Koeffizienten der Eingangsvariablen weitestgehend gleichgeblieben sind. Therapie-Response leistet einen bedeutsamen Beitrag zur Vorhersage des Therapieerfolgs: Das Modell mit Therapie-Response klärt 4% mehr Varianz auf und bildet die Daten besser ab, als das Ausgangsmodell (Anpassungstest⁸ über die Differenz des -2 Log Likelihood-Wert: $\Delta\chi^2 = 7.3$, $df = 1$, $p < 0.01$). Responder

⁸unterscheiden sich zwei Modelle um nur einen Prädiktor, wie es hier der Fall ist, so kommt dieser Modelltest dem Signifikanztest dieses Prädiktors gleich (vgl. mit der Spalte ‘Wald- χ^2 ’ in der Tabelle).

haben nach Kontrolle bedeutsamer Eingangsvariablen eine ca. 4 mal so hohe Chance auf Therapieerfolg als Non-Responder.

Tabelle 14 gibt die Ergebnisse des Modells zur Vorhersage von Therapieerfolg für Bulimie aus der Patientenperspektive wieder, das sowohl relevante Eingangsinformationen als auch Therapie-Response (hier auch die Verbesserung der Symptomatik) miteinbezieht: Ähnlich wie schon bei der Anorexie unterscheiden sich die Koeffizienten der Ein-

Tabelle 14:

Modellierung des Therapieerfolgs, Bulimie (zuzüglich Therapie-Response)

Variable	Koeffizient	se	Wald- χ^2	p	Odds-Ratio
unspez. vs. spez. Programm	0.71	0.30	5.44	0.0197	2.0 (1.1,3.7)
unspez. vs spez. Abteilung	0.85	0.33	6.43	0.0112	2.3 (1.2,4.5)
unspez. vs spez. Klinik	0.66	0.31	4.67	0.0307	1.9 (1.1,3.5)
anfängliche Symptomschwere	-1.04	0.20	28.45	0.0001	0.4 (0.2,0.5)
Response	1.62	0.23	47.68	0.0001	5.1 (3.2,8.0)

$R^2 = 0.19$, $-2\text{LogLikelihood} = 581.4$

gangsbedingungen (vgl. Tabelle 12 und 14) nur geringfügig. Therapie-Response stellt bei den Bulimien auch nach der Kontrolle relevanter Eingangsvariablen einen wichtigen Prädiktor des Therapieerfolgs dar. Patienten, die nach vier Wochen eine deutliche Verbesserung der Symptomatik aufweisen, haben eine fünf mal so hohe Wahrscheinlichkeit auf Therapieerfolg als Non-Responder. Obwohl das Modell mit Response mehr als doppelt so viel Varianz aufklärt (19%) als das Ausgangsmodell (8%), klärt es deutlich weniger Varianz auf als das Modell der Anorexien.

Bedingte Odds-Ratios: Um nicht die Tabellen der Analysen aller Operationalisierungen in Relation zu Erfolg auf Therapeuten- und Patientenperspektive darstellen zu müssen, sind die Ergebnisse in den nachfolgenden Grafiken zusammengefaßt. Wie im vorigen Abschnitt demonstriert, wurde für alle Operationalisierungen ein Modell gerechnet, welches zu dem Ausgangsmodell die jeweilige Operationalisierung des Therapie-Response hinzunahm. Die Koeffizienten und Standardfehler der Ausgangsmodelle veränderten sich durch die Hinzunahme von Therapie-Response nicht wesentlich, allerdings bei AN in stärkerem Maße als bei BN. Jeder Punkt in der Grafik repräsentiert folglich den bedingten Odds-Ratio des Therapie-Response aus einem separaten Modell.

Abbildungen 6 und 7 geben die Odds-Ratios der einzelnen Operationalisierungen und deren 95%-Konfidenzintervalle unter Kontrolle der Kovariaten wieder. Auch hier sind alle Odds-Ratios – sowohl bei Therapieerfolg aus Patienten- als auch aus Therapeutenperspektive – größer als eins und belegen damit die höhere Wahrscheinlichkeit der

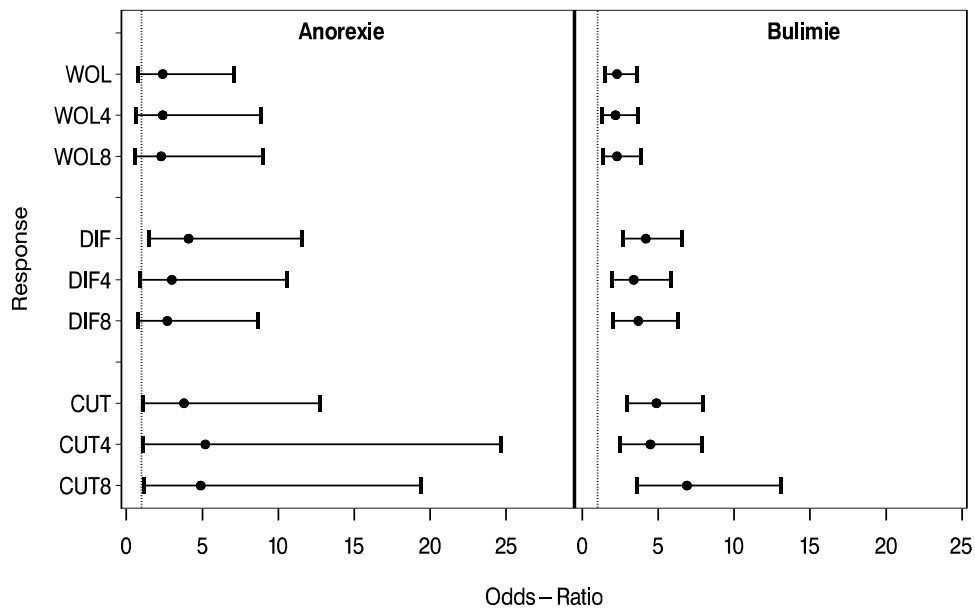


Abbildung 6: Odds-Ratios unter Kontrolle relevanter Faktoren (Patientenperspektive)

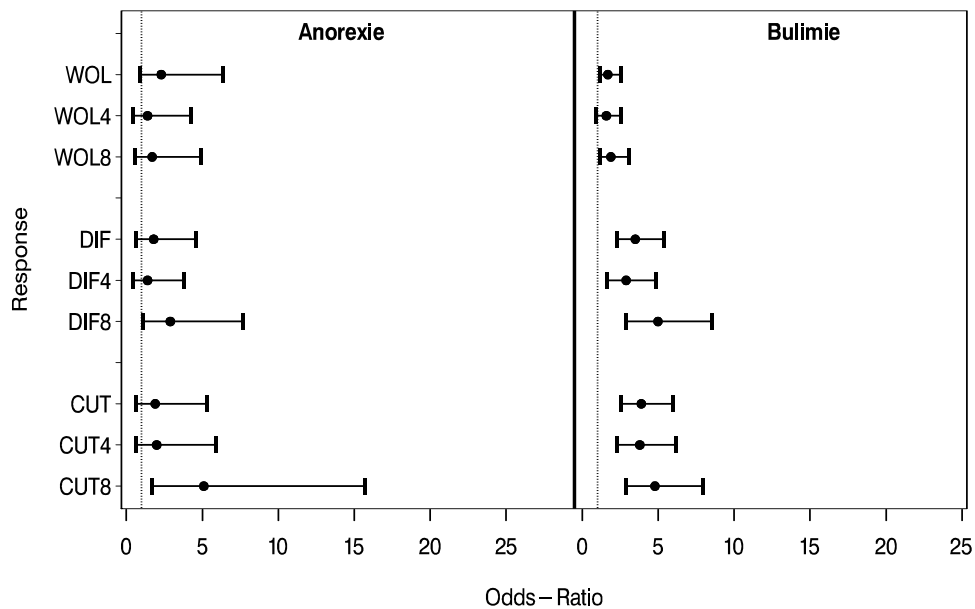


Abbildung 7: Odds-Ratios unter Kontrolle relevanter Faktoren (Therapeutenperspektive)

Responder auf Therapieerfolg. Bei Bulimie sind die bedingten Odds-Ratios im Vergleich zu den unbedingten fast unverändert. Dies gilt besonders im Zusammenhang mit Erfolg aus der Patientenperspektive, aber auch bei Erfolg aus Therapeutenperspektive sind die Odds-Ratios nur wenig niedriger.

Bei Anorexie liegen die bedingten Odds-Ratios auch noch auf etwa der gleichen Höhe wie die unbedingten, allerdings sind bei der Patientenperspektive die Konfidenzintervalle deutlich breiter. Sie sind ein Hinweis auf Interkorrelationen zwischen Therapie-Response und den klinischen Eingangsvariablen, d.h. daß die klinische Eingangsinformation u.U. eine gewisse Vorhersage des Therapie-Response zuläßt (siehe Abschnitt 8). Bei der Therapeutenperspektive ist die Vergrößerung der Konfidenzintervalle nicht zu beobachten, jedoch ist nun für die moderate Symptomatik (CUT) der Zusammenhang zum Ergebnis etwas niedriger, so daß hier nur bei CUT8 noch ein sicherer Zusammenhang feststellbar ist.

Der Erhebungszeitpunkt von Therapie-Response scheint auch hier keine ausgeprägte Rolle zu spielen: lediglich bei der Therapeutenperspektive deutet sich ein bescheidener Vorteil des Responses nach acht Wochen an – allerdings sind die Unterschiede nicht ausgeprägt genug, um von durchschlagender Bedeutung zu sein.

Anteil korrekter Klassifikationen: Einen weiteren Hinweis auf den Zusammenhang zwischen Therapie-Response und Therapieergebnis erhält man über den Vergleich der aufgrund verschiedener Modelle erzielten Anteile korrekter Klassifikationen. Die Tabellen 15 und 16 geben eine Übersicht über die Anteile der aufgrund des Ausgangsmodells ('ohne Response') korrekt klassifizierten Patientinnen und dem Anteil derjenigen Modelle, in die Therapie-Response als zusätzlicher Prädiktor eingesetzt wurde. Den Anteil

Tabelle 15:

*%-Anteil der richtig klassifizierten Patientinnen
und R^2*

Response		Anorexie		Bulimie	
		%	R^2	%	R^2
ohne Response	P*	91.7	45.6	70.7	8.0
	T*	90.7	43.6	63.8	12.5
Wellbeing	P	91.7	47.1	71.7	11.6
	T	90.3	45.0	64.3	13.1
Verringerung der Sympt.	P	91.6	49.9	71.9	19.4
	T	90.3	44.7	67.1	20.5
Moderate Symptome	P	92.0	48.5	70.6	18.0
	T	90.3	44.8	68.2	21.3

* P=Patienten-, T=Therapeutenperspektive
mind. 6 Wochen Behandlungsdauer

an korrekten Klassifikationen erhält man, in dem alle Patientinnen mit einer bestimmten (aus dem Modell errechneten) Erfolgswahrscheinlichkeit, etwa ≥ 0.5 , als 'Erfolg' wertet. So lassen sich vorhergesagte und tatsächliche 'Erfolge' und 'Mißerfolge' vergleichen. Zusätzlich ist in den Tabellen für alle Modelle noch der Anteil an aufgeklärter

Varianz (R^2) angegeben.

Tabelle 15 zeigt die Ergebnisse für Response nach vier Wochen, bei Patientinnen mit mindestens sechs Wochen Behandlung. Über 90% der AN-Patientinnen werden im Ausgangsmodell korrekt klassifiziert; das Hinzufügen von Therapie-Response erbringt keinen weiteren Zugewinn an korrekten Klassifikationen und keinen nennenswerten Zugewinn an erklärter Varianz. Bei Bulimie werden ohne Response 64% bis 70% richtig klassifiziert; diese Werte können durch die Hinzunahme von Therapie-Response noch leicht gebessert werden auf 68% bis 72%. Der Anteil an erklärter Varianz steigt in allen Operationalisierungen an: er verdoppelt sich für die 'Verringerung der Symptome' und die 'moderate Symptomatik' (Patientenperspektive).

Ganz ähnlich fallen die Ergebnisse für den Vergleich von Therapie-Response nach vier bzw. acht Wochen aus (Tabelle 16): bei Anorexie kann Therapie-Response unabhängig

Tabelle 16:
%-Anteil der richtig klassifizierten Patientinnen nach
vier bzw. acht Wochen

Response		Anorexie		Bulimie	
		%	R ²	%	R ²
<i>ohne Response</i>					
	P	92.9	46.3	71.9	7.3
	T	88.3	37.6	65.3	14.8
<i>Wellbeing</i>					
nach 4 Wochen	P*	92.8	47.5	73.0	13.5
nach 8 Wochen	P	92.6	47.1	72.3	12.3
nach 4 Wochen	T*	87.8	39.5	65.9	17.9
nach 8 Wochen	T	87.9	39.3	66.9	17.9
<i>Verringerung der Sympt.</i>					
nach 4 Wochen	P	91.3	49.0	73.8	17.8
nach 8 Wochen	P	92.2	48.6	73.0	19.4
nach 4 Wochen	T	87.2	39.8	68.3	21.1
nach 8 Wochen	T	88.5	44.0	71.5	27.3
<i>Moderate Symptome</i>					
nach 4 Wochen	P	91.8	50.2	71.9	16.0
nach 8 Wochen	P	92.7	50.7	71.9	21.2
nach 4 Wochen	T	87.2	40.1	68.3	22.6
nach 8 Wochen	T	88.0	45.3	71.2	26.2

* P=Patienten-, T=Therapeutenperspektive

mind. 10 Wochen Behandlungsdauer

vom Erfassungszeitpunkt keinen wesentlichen Zugewinn mehr erzielen. Lediglich in Relation zum Erfolg aus der Therapeutenperspektive vergrößert sich etwas der Anteil an erklärter Varianz. Bei Bulimie erhöht sich die Rate der korrekten Klassifikationen um

1 bis 2% für die Operationalisierungen ‘Verringerung der Symptome’ und ‘Moderate Symptome’. Das R^2 steigt bei Hinzunahme von Response an, besonders bei den symptombezogenen Operationalisierungen. Der Zeitpunkt der Erfassung spielt nur bei der Therapeutenperspektive eine Rolle: hier sind die Raten bei Response nach acht Wochen etwas höher als für Response nach vier Wochen. Allerdings ist der Unterschied nicht sehr groß.

Zusammenfassung:

Der positive Zusammenhang zwischen Therapie-Response und Therapieerfolg bleibt auch nach der statistischen Kontrolle anderer Parameter erhalten. Der Schwächste Zusammenhang besteht für die Operationalisierung des Therapie-Response über das ‘momentane Wohlbefinden (Wellbeing)’; hier liegen die bedingten Odds-Ratios zwischen 1 und 2.5. Am stärksten ist er für die Verringerung der Symptome nach vier Wochen bzw. das Vorliegen moderater Symptome nach vier Wochen; hier liegen die Odds-Ratios zwischen 3 und 6. Patientinnen mit Responder nach diesen beiden Kriterien, hatten folglich eine 3 bis 6 mal höhere Wahrscheinlichkeit auf Therapieerfolg am Ende der Behandlung als Patientinnen ohne Response.

Sowohl Zeitpunkt der Erfassung von Therapie-Response als auch die Perspektive des Erfolgskriteriums am Ende der Behandlung spielen keine bedeutende Rolle. Der Anteil an Respondern nach vier Wochen auf Response nach acht Wochen nimmt zwar etwas zu, jedoch verändert dies im wesentlichen nicht den Zusammenhang. Die Erfassung von Therapie-Response nach acht Wochen bringt damit keinen *Zugewinn* für die Vorhersage des Therapieergebnisses.

Aufgrund der geringeren Stichprobengröße und der unausgeglicheneren Erfolgs-/Mißerfolgsrate von 10% zu 90% sind bei Anorexie die Ergebnisse unsicherer – eine bessere Differenzierung der Erfolgsrate ist auch durch die Hinzunahme von Therapie-Response kaum noch möglich. Bei Bulimie ist mit Hilfe der Eingangsinformation der Patientinnen und Therapie-Response eine leichte Verbesserung der Vorhersage des Therapieergebnisses möglich als im Vergleich zur Vorhersage mit der Eingangsinformation alleine.

8 Exkurs I: Kann Therapie-Response vorhergesagt werden?

Eine weitere wichtige Frage ist, ob die zu Beginn der Behandlung bekannten Informationen über den Patienten eine Prognose bzgl. des Therapie-Response zulassen. Aus den bisherigen Informationen wissen wir, daß Therapie-Response auf symptomatischer Ebene mit einer erhöhten Wahrscheinlichkeit auf ein erfolgreiches Behandlungsergebnis einhergeht; und ebenso daß die Hinzunahme von Therapie-Response zur Vorhersage des

Therapieergebnisses bei AN keine nennenswerte Verbesserung mit sich bringt, während dies bei BN der Fall ist. Das legt die Vermutung nahe, daß bei AN eine Korrelation zwischen Therapie-Response und den Eingangsinformationen besteht, während dies bei BN in erheblich geringeren Maße der Fall ist.

Um dies genauer zu überprüfen wurden weitere logistische Regressionen mit schrittweiser Hinzunahme der Prädiktoren durchgeführt; *abhängige Variable* waren nun jeweils die verschiedenen Operationalisierungen des Therapie-Response (nach vier Wochen), *unabhängige Variablen* die in Tabelle 10 (S. 57) aufgeführten Eigenschaften. Dabei fallen die zu Beginn der Behandlung nicht bekannten Behandlungsparameter (Dauer und Intensität) als Prädiktoren weg; die Klinikzugehörigkeit wurde (zunächst) ebenfalls weggelassen, da sie keine Patienten- sondern eine Klinikeigenschaft darstellt.

Die Zusammenhänge der Prädiktoren mit dem Therapie-Response sind in Tabelle 17 aufgeführt.

Tabelle 17: *multiple Korrelation der Eingangsvariablen mit Therapie-Response*

Response Operationalisierung	AN	BN
Wellbeing	0.09	0.09
Verringerung der Symptomatik	0.01	0.09*
Moderate Symptomatik	0.46	0.17

Response nach 4 Wochen

* siehe Anmerkung im Text

Es wird deutlich, daß die klinische Eingangsinformation nur recht spärlich mit Therapie-Response zusammenhängt. Lediglich die Response-Operationalisierung der ‘moderaten Symptomatik’ bei Anorexie zeigt einen größeren Zusammenhang, alle anderen Operationalisierungen kommen über einen Zusammenhang von 0.2 nicht hinaus. Die Koeffizienten des Modells zur Vorhersage der ‘moderaten Symptomatik’ bei AN (Tabelle 18) verraten, welche Patienten am ehesten Therapie-Response zeigen: es sind dies in erster Linie diejenigen, welche schon mit geringer Eßstörungssymptomatik die Behandlung beginnen. Je höher das tatsächliche Gewicht und das Wunschgewicht und je niedriger die Eßstörungssymptome laut EDI-Subskalen 1 bis 3, desto höher die Chance auf Therapie-Response (hier: moderate Symptomatik nach vier Wochen). An den standardisierten Koeffizienten läßt sich die relative Größe der Effekte ablesen: Untergewicht und EDI-Skalen haben den stärksten Einfluß.

Dieses Ergebnis weist darauf hin, daß bei Anorexie lediglich solche Patientinnen Therapie-Response zeigen, die auch schon bei Beginn der Behandlung relativ milde Symptome aufwiesen. Dieses Ergebnis wirft nochmal von anderer Perspektive Licht auf die langsamere Veränderung der Symptome bei AN im Vergleich zu BN: wie schon

Tabelle 18: Vorhersage des Therapie-Response am Beispiel der ‘moderaten Symptomatik’ (Anorexie)

Variable	Koeffizient	se	Wald- χ^2	p	stand. Koeff.
Intercept	-18.18	3.03	35.94	0.0001	–
Alter	0.07	0.03	6.65	0.0099	0.23
Untergewicht	0.13	0.02	32.86	0.0001	0.66
Wunschgewicht	0.08	0.03	9.83	0.0017	0.36
EDI Subskalen 1-3	-0.07	0.01	20.22	0.0001	-0.50

$R^2 = 40.6$

in Abschnitt 6.4.3 erwähnt, verändern sich die Symptome bei AN innerhalb der ersten vier Wochen nicht so rasch, daß nur wenige Patientinnen mit starken Symptomen schon nach 4 Wochen eine Symptomsschwere (laut Kurzbeurteilung) von 1.25 oder weniger erreichen.

Die Schwere der Anfangssymptomatik ist jedoch auch ein entscheidender Faktor für Response bei Bulimie: Die ‘Verringerung der Symptome nach vier Wochen’ erfordert ein Mindestmaß an Symptomen (laut Kurzbeurteilungsscore mindestens 1.5) zu Beginn der Behandlung. Daher können bei Bulimie per definitionem nur solche Patientinnen nach dieser Operationalisierung Response zeigen, welche zu Behandlungsbeginn mindestens einen Symptomwert von 1.5 oder höher aufwiesen. Aus diesem Grund enthält das entsprechende Modell in Tabelle 17 den Prädiktor ‘gering ausgeprägte Symptomatik’ nicht. Der geringe Zusammenhang der übrigen Variablen zum Therapie-Response darf daher nicht darüber hinwegtäuschen, daß nur Patientinnen mit einem Anfangs-Score von mindesten 1.5 Response zeigen können.

Untersucht man bei Bulimie zusätzlich den Zusammenhang zwischen Therapie-Response und den Klinikategorien, so stellt sich heraus, daß sie zu einem nicht unerheblichen Teil mit Therapie-Response verknüpft sind. Dazu wurden zu den o.g. Modellen noch jeweils die Klinikategorien als Prädiktor hinzugenommen. Sie zeigten einen deutlichen Zusammenhang mit den symptomorientierten Response-Operationalisierungen, d.h. sie wurden beim schrittweisen Hinzufügen der Variablen als bedeutsame Variablen ausgewählt. Die Klinikategorien können einen weiteren guten Teil der Variabilität aufklären; so steigt der Anteil aufgeklärter Varianz bei der Operationalisierung ‘moderate Symptomatik’ von 17.4% auf 30.7%. In eßstörungsspezifischen Kliniken war der Anteil an Patientinnen mit Therapie-Response deutlich höher als in anderen Klinikategorien. Dieser Befund steht in einem gewissen Gegensatz zu den Ergebnissen von Olmsted et al. (1996), denen es nicht gelang mit Hilfe der zu Behandlungsbeginn verfügbaren Information langsame, schnelle und Non-Responder vorherzusagen. Das mag darauf zurückzuführen sein, daß deren Studie mit einem stark ausgewähltem und damit recht homogenen Klientel durchgeführt wurde, während die vorliegende Untersuchung sehr unterschiedliche Patientinnen enthält.

Im Vergleich mit den Tabellen 13 und 14 fällt auf, daß bei der Modellierung des Therapie-Response teilweise die gleichen Variablen einen substantiellen Beitrag leisten, welche dies auch bei der Vorhersage des Therapieerfolges tun: bei AN ist dies in erster Linie die Schwere der Anfangssymptomatik, bei den Bulimien die Klinik-kategorien. Diese Variablenkomplexe und auch Therapie-Response 'konkurrieren' folglich miteinander bei der Vorhersage des Therapieerfolges.

Dies macht die stark vergrößerten Konfidenzintervalle des Zusammenhangs zwischen Response (moderate Symptomatik nach vier Wochen) und Erfolg bei der Anorexie verständlich (vgl. Abbildung 6): relativ gering ausgeprägte Symptome zu Behandlungsbeginn hängen folglich *sowohl* mit Therapie-Response *als auch* mit Therapieerfolg am Ende der Behandlung zusammen. Responder, die nach vier Wochen nur noch moderate Symptomatik aufweisen, haben also eine höhere Wahrscheinlichkeit auf Therapieerfolg – allerdings handelt es sich bei ihnen zu einem guten Teil um diejenigen, die auch schon 4 Wochen vorher niedrige Symptomatik hatten. Etwa 20% der AN-Patientinnen hatten zu Beginn der Behandlung eine relativ niedrige Symptomatik; nach vier Wochen kamen weitere 10% hinzu, so daß zwei Drittel der Responder schon zu Beginn der Behandlung nur wenig Symptome aufwiesen. Diese 10% hatten eine ähnlich hohe Erfolgsrate (25%) als die Patientinnen mit eingangs niedrigen Symptomen (23%).

Bei der Bulimie spielt die anfängliche Symptomschwere auch eine entscheidende Rolle bei der Vorhersage des Therapieerfolgs: hier konkurrieren Anfangsstatus und Therapie-Response allerdings nicht so stark miteinander wie bei der Anorexie, Tabelle 14 zeigt, daß es zwei Gruppen von Patientinnen mit größerer Chance auf Therapieerfolg gibt: dies sind zum einen die, die nach vier Wochen nurmehr moderate Symptome aufweisen und zu anderen die, bei denen dies schon zu Behandlungsbeginn der Fall war. 10% (n=56) der Bulimien hatten zu Beginn der Behandlung einen Score von 1.25 und niedriger, ihre Erfolgsrate lag bei 43% (n=24). Im Gegensatz zu AN haben bei BN nach vier Wochen deutlich mehr Patientinnen moderate Symptomatik, nämlich 58% – der Anteil derjenigen mit anfänglich niedrigen Symptomen liegt hier nur bei 13%. Bei AN sind somit die Patientinnen mit anfangs geringerer Symptomatik und die Responder zu 2 Drittel identisch, bei Bulimie dagegen nicht.

Die 90% (n=484) der BN-Patientinnen, die mit stärkerer Symptomatik starteten hatten eine Erfolgsrate von 29% (n=138). Hatten diese Patientinnen Therapie-Response (moderate Symptome in der vierten Woche, n=272) so lag ihre Erfolgsrate bei 42% (n=113) ähnlich hoch wie bei Patientinnen mit anfangs niedriger Eingangssymptomatik; die Erfolgsrate von Patientinnen mit hoher Eingangssymptomatik ohne Therapie-Response (n=212) hatten dagegen eine Erfolgsrate von lediglich 12% (n=25).

Bei Response als 'Verringerung der Symptome' werden im Zusammenhang mit geringer Anfangssymptomatik (bei Bulimie) zwei disjunkte Gruppen mit erhöhter Erfolgswahrscheinlichkeit erfaßt; da Responder hier mit einer Mindestsymptomatik von 1.5 Punk-

ten starten müssen, überschneiden sie sich nicht mit der Gruppe der Patientinnen mit niedrigen Anfangssymptomen.

Zusammenfassung:

Therapie-Response hing in dieser Stichprobe mit den Eingangsbedingungen der Patientinnen zusammen; beide Faktoren sind darüberhinaus mit einer erhöhten Erfolgswahrscheinlichkeit verbunden. Dies gilt insbesondere für die symptombezogenen Response-Operationalisierungen, während dies für Response als Verbesserung des Wohlbefindens nicht der Fall war. Sowohl Patientinnen mit niedriger Anfangssymptomatik, als auch Patientinnen mit moderater Symptomatik nach vier Wochen haben eine höhere Wahrscheinlichkeit für Therapieerfolg. AN-Patientinnen verändern ihre Symptome wesentlich langsamer als BN-Patientinnen; daher ist bei ihnen der Zusammenhang zwischen Anfangsschwere und Therapie-Response ausgeprägter als bei BN. Wer bei Anorexie nach vier Wochen moderate Symptomatik zeigt, war zu Behandlungsbeginn auch nicht stark beeinträchtigt; bei den Bulimien dagegen können die Responder auch von sehr starker Anfangssymptomatik her kommen.

Definiert man Therapie-Response als Verringerung der Symptome, so stellt sich bei AN kein Zusammenhang mit der Anfangssymptomatik heraus, da für Response nur eine Verringerung von nur 0.5 Punkten erforderlich ist. Bei BN dagegen besteht dieser Zusammenhang, da per definitionem nur Patientinnen mit einem Anfangsscore von mindestens 1.5 Response zeigen können.

Aus diesem Exkurs bleibt festzuhalten, daß zur Abschätzung der Erfolgswahrscheinlichkeit sowohl die Schwere der Anfangssymptomatik, als auch Therapie-Response zu berücksichtigen sind. Beide Faktoren tragen zur Vorhersage des Therapieerfolgs bei. Redundanzen ergeben sich nur bei AN hinsichtlich der Response-Operationalisierung als 'moderate Symptomatik nach vier Wochen'.

9 Exkurs II: Therapie-Response und längerfristiger Erfolg

Ein weiterer Baustein in der Überprüfung der Validität von Therapie-Response liegt in dem Zusammenhang mit dem langfristigen Therapieerfolg. Wenn man aus dem Vorliegen von Therapie-Response nach vier bzw. acht Wochen auf eine erhöhte Wahrscheinlichkeit für einen günstigen Zustand auch längere Zeit nach der Behandlung schließen kann, so untermauerte dies die Bedeutsamkeit des Therapie-Response nachhaltig. Das Design der MZ-ESS Studie (siehe S. 36) erlaubt den Vergleich der Erfolgsraten von Respondern und Non-Respondern nach 2.5 Jahren. Parallel zu den in Abschnitt 7.2 durchgeführten Analysen sollen im folgenden mit Hilfe von logistischer Regression die Erfolgswahrscheinlichkeiten nach 2.5 Jahren untersucht werden. Der Vergleich von Re-

spondern und Non-Respondern erfolgt dabei – unter statistischer Kontrolle relevanter Einflußfaktoren – über die bedingten Odds-Ratios.

Das Vorgehen ist genau das gleiche wie in den vorangegangenen Analysen: Für alle Patientinnen, von denen Informationen zum Katamnesezeitpunkt vorliegen, wurde nach den gleichen Kriterien wie in Tabelle 3 (S. 43) Therapieerfolg operationalisiert. Dieser wird als abhängige Variable in die logistische Regression eingesetzt. Unabhängige Variablen sind die in Tabelle 10 (S. 57) aufgelisteten Prädiktoren zuzüglich der jeweiligen Operationalisierung von Therapie-Response. Die Variablenauswahl der einzelnen Modelle erfolgte wieder über schrittweises Hinzufügen relevanter Variablen (Einschlußkriterium $p \leq 0.1$), wobei Therapie-Response unabhängig von der Stärke des Zusammenhangs zum Therapieerfolg immer ins Modell aufgenommen wurde. Für jede Response-Operationalisierung liegt damit ein separates Modell vor, in dem der bedingte Odds-Ratio des Therapie-Response den Unterschied der Erfolgswahrscheinlichkeiten zwischen Respondern und Non-Respondern – unter Kontrolle anderer relevanter Einflußfaktoren – angibt. Die Odds-Ratios (mit 95% Konfidenzintervallen) der einzelnen Modelle wurden daraufhin wieder in einer Grafik zusammengefaßt. Wie schon vorher wird Anorexie und Bulimie getrennt analysiert.

Von den insgesamt 870 Patientinnen dieser Arbeit liegen von 67.7% (n=589) Informationen zum Katamnesezeitpunkt vor (AN=211, BN=378). Die Quote der erreichten Patientinnen unterscheidet sich zwischen beiden Diagnosen nicht. Für die Untersuchung des Response nach acht Wochen bleiben bei AN noch 138 (65%) bei BN noch 250 (66%) Patientinnen mit einer Behandlungsdauer von mindestens zehn Wochen übrig. Die zur Katamnese erreichten Patientinnen unterschieden sich nur unwesentlich von den Patienten die nicht mehr erreicht werden konnten, so daß Selektionseffekte weitgehend ausgeschlossen werden können (vgl. Kächele 2000) Abbildung 8 zeigt die bedingten Odds-Ratios der verschiedenen Response-Operationalisierungen. Die Beschriftung und Lesart der Grafik geschieht auf die gleiche Weise wie bei Abbildung 6 und 7 (S. 62 und 62) in Abschnitt 7.2.

Der eindeutig positive Zusammenhang zwischen Response und Therapieerfolg bei Anorexie ist geblieben, bei Bulimie dagegen nicht mehr durchgängig präsent. Bei AN sind fast alle Ratios deutlich größer als eins. Am auffälligsten ist der Zusammenhang zwischen dem ‘verbesserten Well-being’ (WOL) und Therapieerfolg: hier liegen die Odds-Ratios zwischen 2.5 und 7, wobei die Konfidenzintervalle die eins nicht einschließen. Zusätzlich spielt sowohl die Dauer der Behandlung als auch der Zeitpunkt der Erfassung von Therapie-Response eine Rolle: Berücksichtigt man nur die Patientinnen mit einer Behandlungsdauer von mindesten zehn Wochen (WOL4) ist der Zusammenhang stärker; bei diesen länger behandelten Patientinnen ist die Wahrscheinlichkeit auf ein positives Therapieergebnis nochmals höher, wenn Response nach acht Wochen erfaßt wird (WOL8). Die Ergebnisse dieses Modells sind in Tabelle 19 näher erläutert.

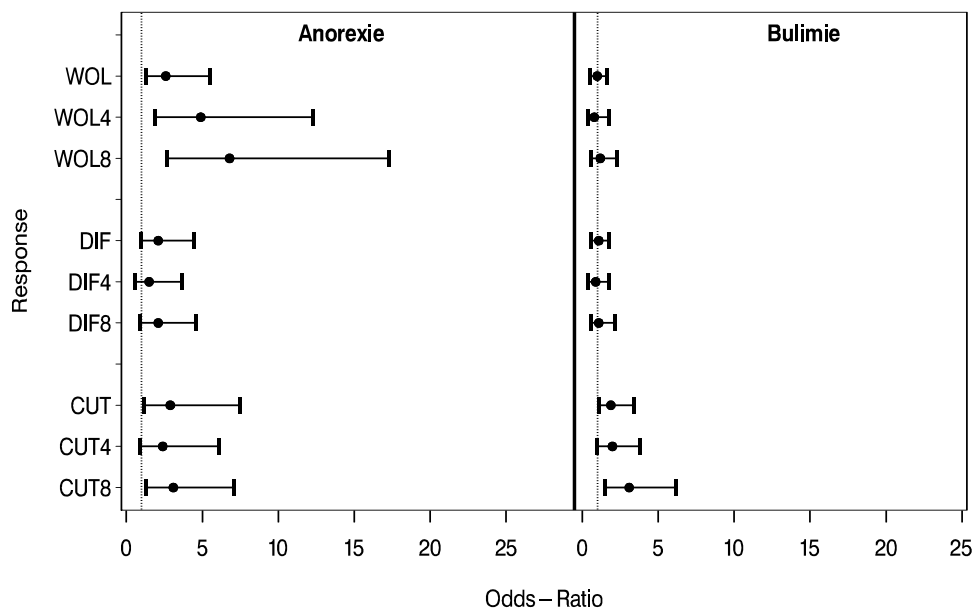


Abbildung 8: Katamnese: Odds-Ratios unter Kontrolle relevanter Faktoren (Patientenperspektive)

Die Odds-Ratios der beiden symptomorientierten Response-Operationalisierungen liegen im Bereich zwischen 1.5 und 3. Bei der ‘Verringerung der Symptome’ (DIF) schließen die Konfidenzintervalle die eins mit ein und lassen somit den Zusammenhang unsicher erscheinen; bei der ‘moderaten Symptomatik’ ist der Zusammenhang sicherer, da die Konfidenzintervalle die eins einschließen. Bei beiden spielen Behandlungsdauer und Zeitpunkt der Erfassung von Response keine entscheidende Rolle.

Bei Bulimie ergibt sich ein anderes Bild: Sowohl für das ‘verbesserte Well-Being’ (WOL) als auch für die ‘Verringerung der Symptomatik’ (DIF) kann kein Zusammenhang zwischen Response und Therapieergebnis mehr festgestellt werden: alle Odds-Ratios liegen hier ganz in der Nähe der eins. Nur bei der ‘moderaten Symptomatik’ (CUT) läßt sich eine zwei- bis dreifach höhere Wahrscheinlichkeit der Responder auf Therapieerfolg im Vergleich zu den Non-Respondern feststellen.

Der Zusammenhang zu Therapieerfolg aus der Therapeutenperspektive zeigt insgesamt etwas niedrigere Odds-Ratios, ansonsten aber ein ganz ähnliches Muster; daher wird hier auf eine weitere Grafik verzichtet. Einzig bei der ‘moderaten Symptomatik’ (CUT) ergeben sich leichte Veränderungen: Bei Anorexie besteht durch die insgesamt niedrigeren Odds-Ratios kaum noch ein Zusammenhang mit Therapieerfolg; bei der Bulimie ist dieser Zusammenhang dagegen etwas stärker als in Relation zur Patientenperspektive.

Die Tabelle 19 gibt die Koeffizienten der Modelle zur Vorhersage des Therapieer-

folgs von Anorexie und Bulimie wieder. Beim Modell für Anorexie wurde aufgrund

Tabelle 19: *Vorhersage des Therapieerfolgs zum Katamnesezeitpunkt*

Anorexie: Response: 'Verbesserung des Well-beings'						
Variable	Koeffizient	se	Wald- χ^2	p	stand. Koeff.	
Intercept	-6.15	2.89	4.51	0.0336	-	
Response (WOL)	-5.20	2.70	3.69	0.0545	-1.23	
Beh.-Dauer	-0.69	0.46	2.27	0.1317	-0.17	
Narzißmus-Inventar	-0.09	0.04	4.86	0.0274	-0.21	
Erkrankungsdauer	-0.12	0.04	9.02	0.0027	-0.35	
Untergewicht	0.06	0.02	7.94	0.0048	0.29	
Wunschgewicht	0.06	0.02	7.74	0.0054	0.29	
Response \times Beh.-Dauer	2.48	1.08	5.23	0.0222	1.48	
$R^2 = 30.4$						
Bulimie: Response: 'moderate Symptomatik'						
Variable	Koeffizient	se	Wald- χ^2	p	stand. Koeff.	
Intercept	-0.89	0.30	8.71	0.0032	-	
Response (CUT)	0.66	0.28	5.58	0.0181	0.18	
1 Vorbehandlung	-0.13	0.40	0.10	0.7532	-0.02	
2 Vorbehandlungen	-0.90	0.32	7.59	0.0059	-0.25	
3 und mehr Vorbeh.	-0.06	0.38	0.03	0.8713	-0.01	
anorekt. Symptome	-0.69	0.36	3.66	0.0558	-0.16	
multi-impulsives Verh.	-0.18	0.11	2.98	0.0841	-0.13	
$R^2 = 10.4$						

Abbildung 8 noch die Interaktion zwischen Therapie-Response und Behandlungsdauer hinzugenommen. Die Ergebnisse zeigen, daß höhere Werte im Narzißmus-Inventar und eine längere Behandlungsdauer mit einer geringeren Wahrscheinlichkeit auf Therapieerfolg einhergehen, während ein höheres Ausgangsgewicht und Wunschgewicht erhöhte Erfolgchancen bedeuten. Die Einzeleffekte des Response und der Therapiedauer lassen sich wegen der vorhandenen Interaktion nicht alleine aufgrund der Zahlen in der Tabelle interpretieren. Sie sind je nach Behandlungsdauer und Therapie-Response unterschiedlich: Mit dem Anstieg der Behandlungsdauer vergrößert sich auch der Unterschied zwischen Respondern und Non-Respondern. Bei einer Dauer von acht Wochen besteht praktisch kein Unterschied zwischen beiden, bei längeren Behandlungen haben die Responder eine höhere Wahrscheinlichkeit auf Therapieerfolg als Non-Responder. Die Differenz zwischen beiden kann wieder als Odds-Ratio ausgedrückt werden ⁹: Der

⁹dazu muß die Differenz der Logits beider Gruppen exponenziert werden: $\text{Odds-Ratio} = e^{\text{Differenz der Logits}}$. Die Logits sind die der logistischen Regression zugrunde liegende Rechen- einheit und geben das Verhältnis von erfolgreichen und nicht erfolgreichen Patientinnen an.

Odds-Ratio zwischen Respondern und Non-Respondern beträgt bei acht Wochen 0.96, bei zwölf Wochen 2.62 und bei 16 und 20 Wochen 5.34, bzw. 9.29.

Bei Bulimie ist nur Therapie-Response (CUT) mit einer höheren Erfolgswahrscheinlichkeit verbunden, das Vorliegen von zusätzlicher anorektischer Symptomatik und von multi-impulsivem Verhalten reduziert die Erfolgchancen. Die Anzahl der vorhergehenden Behandlungen ging als dummy-kodierte Variable ein: hier besteht nur ein Unterschied zwischen den Patientinnen, die 2 vorige Behandlungen hatten und denen ohne vorige Behandlung: Patientinnen mit 2 Vorbehandlungen hatten eine niedrigere Chance auf Therapieerfolg.

Zusammenfassung:

Patientinnen mit Therapie-Response werden auch 2.5 Jahre nach der Behandlung häufiger als Therapieerfolg eingestuft. Für die Anorexien gilt dieser Zusammenhang insgesamt bei allen Operationalisierungen, am deutlichsten überraschenderweise für die 'Verbesserung des Well-being' (Odds-Ratios zwischen 2.5 und 7), obwohl hier in Relation zum Behandlungsergebnis bei *Therapieende* der Zusammenhang niedriger war. Für länger behandelte Patientinnen mit Therapie-Response sowie bei der Erhebung des Response nach acht Wochen verstärkt sich hier der Effekt des Therapie-Response. Am geringsten ist der Zusammenhang zwischen Response und Erfolg bei der 'Verringerung der Symptome' – immerhin haben auch hier die Responder eine bis zu doppelt so hohe Wahrscheinlichkeit auf Therapieerfolg.

Bei Bulimie konnte eine erhöhte Erfolgchance der Responder nach 2.5 Jahren nur bei der Operationalisierung über die 'moderate Symptomatik' festgestellt werden. Sie ist hier immer noch doppelt so hoch wie bei Patientinnen ohne Therapie-Response. Bei den anderen Operationalisierungen besteht der Vorteil der Responder gegenüber den Non-Respondern nicht mehr.

10 Therapie-Response und Therapieverlauf

Die bisherigen Analysen demonstrieren den relativ hohen Zusammenhang zwischen Krankheitsstatus nach vier bis acht Wochen (Therapie-Response) und dem Ergebnis der Therapie zum Abschluß der Behandlung (Therapieerfolg). Dabei ist bisher nur der Verlauf der ersten vier bzw. acht Wochen untersucht worden. Da jedoch für jede Patientin Angaben über die gesamte Behandlungslänge hinweg vorliegen, ermöglicht das die Untersuchung der Symptomverläufe bis zum Behandlungsende. Die Stärke des Zusammenhangs zwischen Therapie-Response und Therapieerfolg, läßt vermuten, daß sich das nach dem Therapie-Response erreichte Symptomniveau beim einem Großteil der Patientinnen auch im weiteren Verlauf hält oder noch weiter verbessert. So müßten sich Unterschiede in der Verlaufsgestalt zwischen Respondern und Non-Respondern

erkennen lassen.

10.1 Der Verlauf auf Gruppenebene

Die verschiedenen Arten von Therapie-Response implizieren unterschiedliche zu erwartende Verlaufsformen. Abhängig von der Ausgangsschwere und dem Status nach vier Wochen sind folgenden Verläufe zu erwarten:

Tabelle 20: *Erwartete Symptomentwicklung der Responder und Non-Responder*

Moderate Symptomatik nach 4 bzw. 8 Wochen:

Responder

- starke Anfangssymptome: schnelle Reduktion auf ein moderates Niveau
- moderate bis leichte Anfangssymptome: gleichbleiben oder leichte Reduktion der Symptome

Non-Responder

- starke Anfangssymptome: gleichmäßig hohe Symptomatik über die erste Phase der Therapie hinweg
- leichte Anfangssymptome: Verschlechterung des Status

Verringerung der Symptome innerhalb von 4 bzw. 8 Wochen:

Responder

- starke Anfangssymptome: Reduktion der Symptome auf ein moderates Niveau
- moderate bis leichte Anfangssymptome: Reduktion der Symptome auf ein niedriges Niveau

Non-Responder

- starke Anfangssymptome: gleichmäßig hohe Symptomatik über die erste Phase der Therapie hinweg, bzw. leichte Reduktion
- leichte Anfangssymptome: Gleichbleiben bzw. Verschlechterung des Status

Mit diesen Punkten ist allerdings nur die erste Phase des Verlaufes beschrieben – wie sich der Krankheitsstatus *nach* den vier bzw. acht Wochen weiterentwickelt, ist offen. Immerhin wissen wir aus Abschnitt 7, daß die Responder im Vergleich zu den Non-Respondern mit einer 3 - 5 mal so hohen Wahrscheinlichkeit am Ende Therapieerfolg aufweisen – der Großteil von ihnen müßte daher das zum Therapie-Response erreichte Symptommiveau halten bzw. weiter verbessern. Umgekehrt ist es bei den Non-Respondern wahrscheinlicher, daß sie auch in der nachfolgenden Zeit nicht genügend Verbesserungen aufweisen, um am Ende als ‘Erfolg’ eingestuft zu werden – entsprechend sind bei den meisten von ihnen im weiteren Verlauf keine deutlichen Verbesserungen mehr zu erwarten.

Gleichzeitig wissen wir jedoch aus der Tabelle 9 (S. 53), daß es durchaus Patientinnen gibt, die trotz Ansprechen auf die Therapie am Ende nicht erfolgreich abschneiden und ebenso, daß Patientinnen ohne Therapie-Response am Ende doch noch Therapieerfolg aufweisen können. Von der Perspektive eines positiven bzw. negativen Therapieergebnisses her schränken sich die Möglichkeiten des Fortgangs weiter ein: unter der Annahme, daß die Therapie zu einem positiven Endergebnis kommt, sollten Therapie-Responder einen Verlauf zeigen, der durch eine starke Symptomreduktion innerhalb der ersten Wochen und eine schwächere Reduktion bzw. ein Gleichbleiben danach gekennzeichnet ist; unter der Annahme eines negativen Ausgangs ist mit einem u-förmigen Verlauf zu rechnen: die Symptome der Responder sollten zunächst stark zurückgehen, danach jedoch wieder ansteigen.

Für Non-Responder unter der Annahme eines positiven Ergebnisses wäre eine geringe Reduktion oder ein Gleichbleiben der Symptome zu Beginn des Verlaufs mit später einsetzender Reduktion zu erwarten; unter der Annahme eines negativen Therapieergebnisses sollten sich die Symptome der Non-Responder weder in den ersten vier bis acht Wochen der Therapie noch in der darauffolgenden Zeit deutlich reduzieren - im wesentlichen also auf hohem Niveau konstant bleiben.

In diesem Abschnitt wird der Gruppenverlauf dieser vier Bedingungen exploriert und dargestellt. Die Bedingungen entsprechen den Kreuztabellen in Abschnitt 7.2: Patientinnen mit und ohne Therapie-Response und solche, mit und ohne Therapieerfolg bei Behandlungsende. Die hier gewählte Operationalisierung für Therapie-Response ist die 'Verringerung der Symptome' innerhalb der ersten 4 Wochen. Sie hat hinsichtlich der Verlaufsform der Symptome die klarsten Implikationen und soll deswegen hier genauer dargestellt werden. Die Abbildungen 9 und 10 zeigen die Verteilungen der Symptomschwere-Scores von Respondern und Non-Respondern (aufgeteilt nach positivem und negativem Therapieergebnis) pro Monat. Die Y-Achse bezeichnet die Symptomschwere (je höher der Score, desto mehr Symptome), während die X-Achse die Monate grob einteilt. Der erste Boxplot zeigt die Verteilung der Scores zum Aufnahmezeitpunkt, der nächste Boxplot die Verteilung im darauffolgenden Monat usw... Die Mediane der Verteilungen sind durch eine Linie verbunden; die gestrichelte Linie stellt die o.g. Grenze von 1.25 Punkten dar, ab der keine akute Symptomatik mehr vorliegt. Die graue Box markiert den Bereich, in dem 50 Prozent Beobachtungen um den Median liegen. Über jeder einzelnen Grafik steht die Gruppeneinteilung und die Anzahl der Patienten, auf die diese Einteilung zutrifft. Die Bilder enthalten nur diejenigen Patientinnen mit bis zu 6 wiederholten Messungen, um die Grafik nicht durch weitere Boxplots der (wenigen) länger behandelten Patientinnen zu überfrachten¹⁰.

Bei der Anorexie (Abbildung 9) spiegeln sich die niedrigen Erfolgsraten zu Behand-

¹⁰Die Bilder der länger behandelten Patientinnen sind diesen sehr ähnlich.

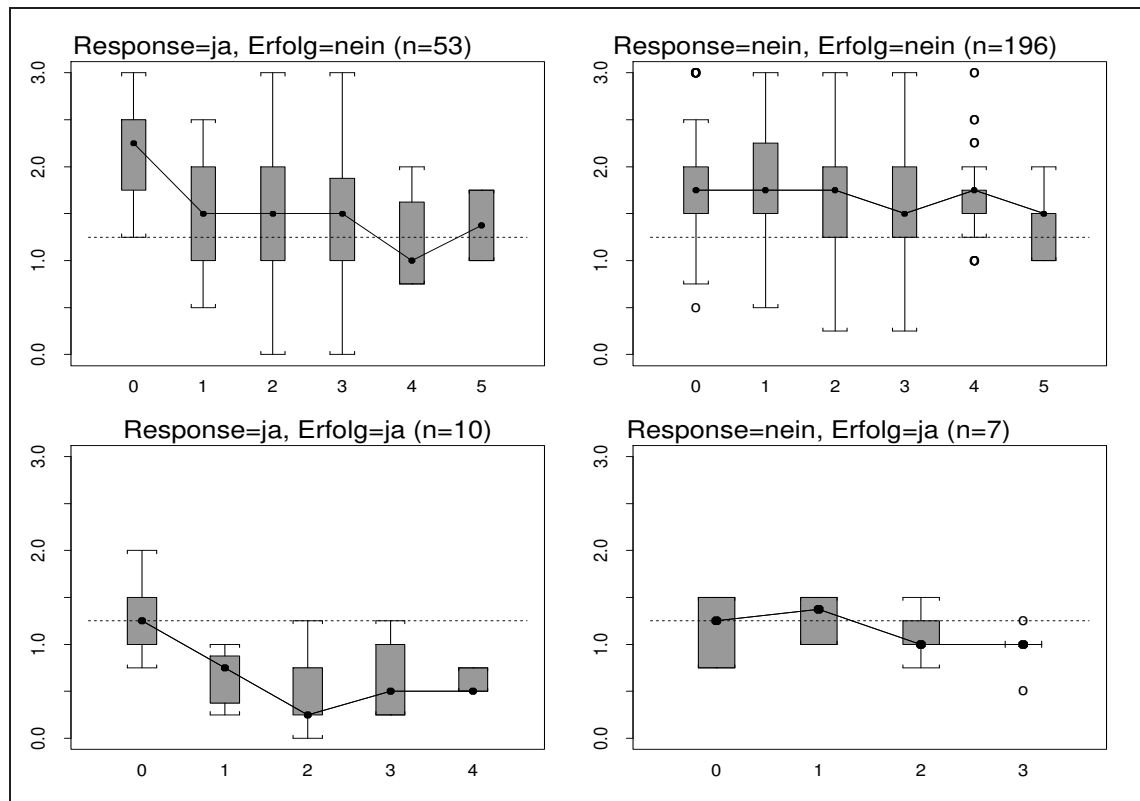


Abbildung 9: **Anorexie:** Verteilungen der Symptomschwere pro Monat (bis zu 6 Messungen)

lungsende wieder: nur 17 Patientinnen (6%) der hier dargestellten 266 Patientinnen wurden als ‘Therapieerfolg’ eingestuft. Die Responder (erste Spalte) reduzieren innerhalb des ersten Monats die Symptome und behalten die Verbesserung – als Gruppe betrachtet – im wesentlichen bei. Die erfolgreichen Responder (n=10, knapp 16% der Responder) starten dabei mit niedrigerem Anfangsniveau als die nicht erfolgreichen; ebenso zeigen sie auch nach dem Response eine weitere Verbesserung.

Die Non-Responder (zweite Spalte) bleiben in ihrer Verteilung der Scores auf durchschnittlich konstantem Niveau. Die Gruppe der Non-Responder, die am Ende erfolgreich abschnitten, besteht nur aus 7 Patientinnen (3% der Non-Responder); sie starteten ebenfalls schon mit relativ niedriger Symptomatik. Der Großteil der AN-Patientinnen gehört zu den Non-Respondern ohne Therapieerfolg; sie starten mit stärkerer Symptomatik und behalten sie in der Mehrzahl auch bei, so daß immer 75% der Patientinnen oder mehr über dem Wert von 1.25 liegen.

Auch hier spiegelt sich das Ergebnis von Abschnitt 7.2 wieder, daß bei der Anorexie der Erfolg zu einem großen Teil von der Schwere der Anfangssymptomatik abhängt; immerhin liegt die Erfolgsrate bei Patientinnen mit Therapie-Response bei 15% und ist damit mehr als doppelt so hoch als die Basisrate von hier 6%.

Auch die Verteilungen der Responder bei BN in Abbildung 10 (erste Spalte) zeigen eine Reduktion der Symptome zwischen Aufnahme und 4-Wochenstatus. Bei den

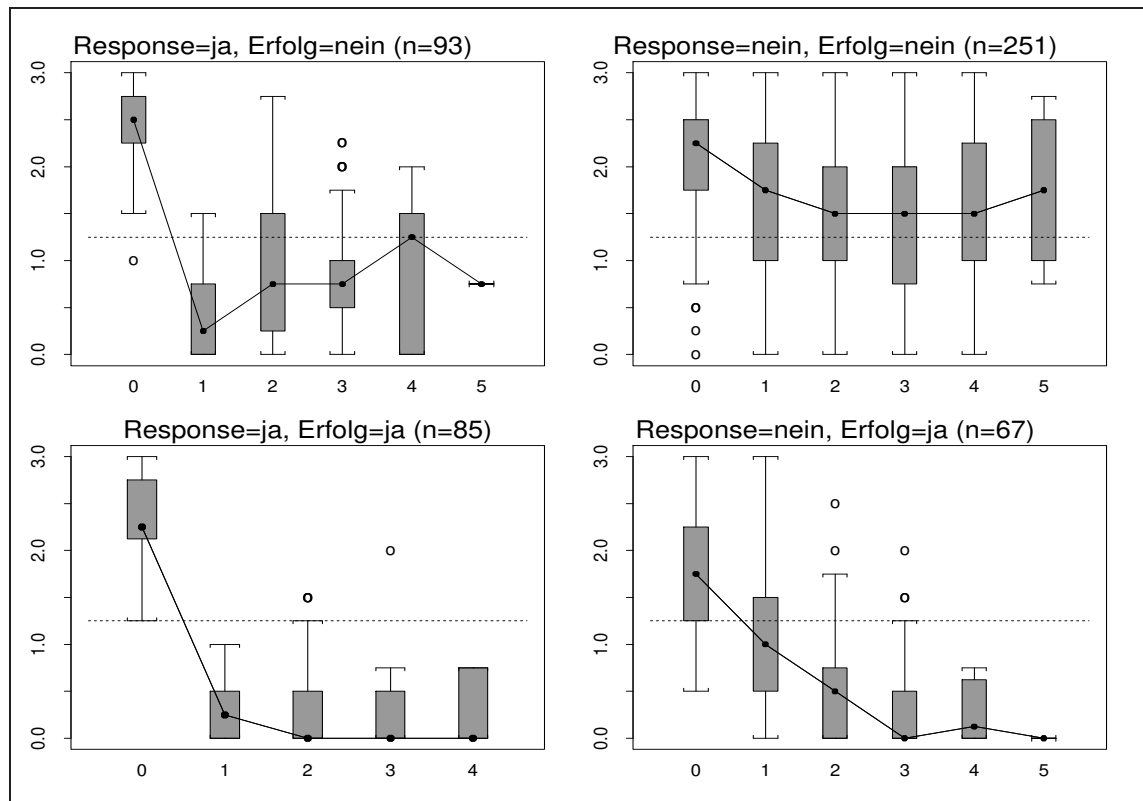


Abbildung 10: **Bulimie:** Verteilungen der Symptomschwere pro Monat (bis zu 6 Messungen)

am Ende nicht erfolgreich behandelten Patientinnen dieser Gruppe ($n=85$, knapp 48% der Responder) fällt auf, daß die anfängliche Verbesserung im Schnitt nicht gehalten werden kann und sich die Symptomwerte wieder erhöhen. Bei den erfolgreichen Respondern dagegen bleiben die Symptomwerte auf sehr niedrigem Niveau. Im Gegensatz zu den AN-Patientinnen ist das Anfangsniveau der erfolgreichen und nicht erfolgreichen Responder in etwa vergleichbar.

Auch bei BN hat nur ein relativ kleiner Teil der Non-Responder (zweite Spalte) am Ende ein positives Therapieergebnis ($n=67$, 21%); sie reduzieren die Symptome anfangs im Durchschnitt langsamer und erreichen erst nach drei Monaten (vierter Boxplot) eine Verteilung, die den erfolgreichen Respondern nach vier Wochen vergleichbar ist. Die nicht erfolgreichen Non-Responder starten auf etwas höherem Niveau als die erfolgreichen; sie reduzieren die Symptome am langsamsten, aber nie liegt der Median unter der Grenze von 1.25 oder ist ein nennenswerter größerer Anteil der Scores unter 1. Die Fallzahlen illustrieren auch hier die höhere Erfolgsrate unter den Respondern (47.8%) im Vergleich zur Basisrate von 30.6%.

Während bei AN sich die Muster dieser vier Gruppen in erster Linie durch die Höhe der Anfangssymptomatik und weniger durch die weitere Verlaufsgestalt unterscheiden, scheint es bei BN vier in sich relativ homogene Gruppen von Verläufen zu geben: Zum einen Patientinnen, die ihre Symptome innerhalb der ersten vier Wochen rasch verlie-

ren und anschließend a) diesen Status halten können oder b) wieder mehr Symptome bekommen. Zum anderen Patientinnen, die die Symptome anfänglich nicht so stark reduzieren und c) auf konstant hohem Niveau bleiben oder d) sie anschließend langsam weiter reduzieren.

Allerdings stellen die Abbildungen 9 und 10 keine Zusammenfassung der *individuellen* Verläufe dar! Es ist also durchaus möglich, daß sich der Score einer einzelnen BN-Patientin (erfolgreich behandelt und Responder) gegen den allgemeinen Trend entwickelt, z.B. von 0.5 im zweiten Boxplot zurück auf 1.0 im dritten Boxplot.

10.2 Der Verlauf auf individueller Ebene

Einen guten Überblick über die individuelle Symptomentwicklung erhält man, in dem man Verläufe einzeln auf der gleichen Skala abträgt, wie dies in den Abbildungen 20 auf S. 74 demonstriert ist. Die Verläufe aller Patientinnen wurden – getrennt nach den vier Gruppen – auf diese Weise dargestellt und exploriert.

Die Exploration ergab, daß sich zwar auch innerhalb der vier Gruppen die einzelnen Verläufe unterscheiden, aber dennoch die Mehrzahl der Patientinnen ähnliche Verlaufscharakteristika aufweist. Die erfolgreichen Responder folgen fast alle dem erwarteten Verlauf: sie reduzieren anfänglich sehr schnell die Symptome und bleiben anschließend meist auf niedrigem Niveau; lediglich bei einigen längeren Behandlungen steigen die Symptome zwischenzeitlich noch einmal an. Bei den nicht erfolgreichen Respondern steigen erwartungsgemäß bei einem guten Teil Patientinnen nach dem Response die Symptome wieder an, aber es gibt auch einige Patientinnen, bei denen dies nicht der Fall ist.

Bei den Non-Responder ohne Therapieerfolg ist ein mehr oder weniger linearer Verlauf auf hohem Niveau ohne starken Rückgang der Symptome zu erwarten oder bei niedriger Ausgangssymptomatik ein Anstieg. Dies ist in der überwiegenden Mehrzahl der Verläufe der Fall. Die Symptomverläufe der erfolgreichen Non-Responder zeichnen sich durch ihre langsamere Geschwindigkeit der anfänglichen Symptomreduktion aus.

Zusammenfassung:

Die Darstellung der Verteilungen der Symptomschwere-Scores pro Monat wie auch der Einzelverläufe verdeutlicht, daß die Patientinnen innerhalb einer Gruppe zwar immer noch Unterschiede hinsichtlich ihrer Verlaufsform aufweisen, aber meist wichtige Charakteristika gemeinsam haben, die durch das Vorhandensein von Therapie-Response und/oder Therapieerfolg gegeben sind. Eine Repräsentation dieser Verlaufsgruppen als ‘Durchschnittsmodelle’, mit denen neu hinzukommende Patientinnen verglichen werden können, stellt dabei das Verbindungsglied zur Anwendung im Qualitätsmanagement dar; dies soll in den nächsten Abschnitten erläutert und illustriert werden.

TEIL III: THERAPIE-RESPONSE IN EINEM VERLAUFSORIENTIERTEN QM

11 Verlaufsorientiertes Qualitätsmanagement

Dieser Abschnitt beschäftigt sich mit der näheren Beschreibung eines verlaufsorientierten Qualitätsmanagements. Die in Abschnitt 2.1.1 formulierten Anforderungen an ein QM-System bleiben in einem verlaufsorientierten System bestehen, können allerdings in der konkreten Anwendung nicht einfach aus bestehenden Systemen übernommen werden. In einem ersten Teilabschnitt werden daher zunächst die Anforderungen für ein verlaufsorientiertes QM-System näher erläutert, um danach in einem zweiten Teilabschnitt das COMPASS-Modell von Howard, Brill, Lueger, O'Mahoney & Grant (1993) als ein verlaufsorientiertes QM-System genauer vorzustellen und anhand der im vorigen Teil genannten Kriterien zu bewerten.

11.1 Anforderungen an ein verlaufsorientiertes QM-System

In dem in Abschnitt 2.2.2 genannten Zitat von Schulte deutet sich die Stoßrichtung an, die ein verlaufsorientiertes QM haben sollte: es geht darum, den Verlauf der Therapie mit objektiven Instrumenten kontinuierlich zu erfassen. Um diese Information bestmöglich für die weiteren Behandlungsentscheidungen nutzen zu können, sollte sie prognostische Relevanz haben, d.h. Vorhersagen über den weiteren Verlauf oder das Therapieergebnis zulassen. Damit kann eine Bewertung des bisher Geschehenen und die Planung der weiteren Interventionen auf eine sicherere Grundlage gestellt werden. Wie dies im einzelnen geschehen kann, soll anhand der folgenden Punkte näher erläutert werden.

Zur Erinnerung seien zunächst nochmal die in Abschnitt 2.1.1 genannten allgemeinen Anforderungen an ein QM-System genannt:

- Erfassung der Therapieergebnisse auf individueller Ebene mit Hilfe von standardisierten psychometrischen Instrumenten zu mindestens zwei Meßzeitpunkten (prä und post)
- Bewertung der Ergebnisse auf dem Hintergrund von angemessenen Standards
- Rückmeldung der Ergebnisse an die behandelnden Personen
- leichte Integrierbarkeit des Systems in den therapeutischen Alltag

Was bedeutet dies für die Umsetzung in einem verlaufsorientierten QM-System? Die vier folgenden Abschnitte führen die wichtigsten Auswirkungen auf einzelne Teile eines solchen System auf.

11.1.1 Erfassung der Therapieergebnisse auf individueller Ebene

In einem verlaufsorientierten QM-System geht es nicht nur darum, das Endergebnis zu erfassen, sondern es muß während der Behandlung kontinuierlich der Status des Patienten erhoben werden, um den Verlauf als Ganzes darstellen zu können. Dies hat unmittelbare Konsequenzen für die Meßinstrumente und ihre Zusammenstellung: zunächst ist klar, daß eine Erhebung zu mehr als nur zwei Zeitpunkten erforderlich ist. Außerdem sollten Instrumente zum Einsatz kommen, die zur *Verlaufsdiagnostik* entwickelt wurden oder geeignet sind (vgl. Lambert, Burlingame, Umphress, Hansen, Vermeersch, Clouse & Yanchar 1996, Umphress, Lambert, Smart, Barlow & Clouse 1997, Nisslmüller 1999, Kordy, Hannover, Richard & Lambert 2000). Sie sollten nicht nur reliabel und valide, sondern zusätzlich noch veränderungssensitiv sein; darüberhinaus dürfen die Bögen nicht zu umfangreich ausfallen, damit die Patienten nicht jedesmal zu viele Fragen bearbeiten müssen und am Ende die Mitarbeit verweigern. Mehrfache Erhebungen stellen außerdem einen Zeit- und Kostenfaktor dar, der in der Planung Eingang finden muß. Soll ein bestehendes System zu einem verlaufsorientierten System erweitert werden, ist es für die Alltagspraxis nicht ratsam, einfach die bestehenden Fragebogenbatterien zu weiteren Zeitpunkten zu verteilen. Vielmehr lohnt es sich hier gezielt auszuwählen, welche Konstrukte über den Verlauf der Therapie hinweg zu erheben sind, um Aufwand und Kosten sowie die Belastung der Patienten in vertretbaren Grenzen zu halten.

11.1.2 Bewertung der Ergebnisse auf dem Hintergrund angemessener Standards

Um eine aussagekräftige Interpretation der Ergebnisse zu ermöglichen, sollten sie zu allgemein anerkannten Standards oder Normen in Beziehung gesetzt werden. Aus der allgemeinen Diagnostik werden hierzu i.d.R. Normstichproben herangezogen, um den Wert eines Einzelnen mit dem Durchschnitt zu vergleichen. Auch bei der Bewertung von Veränderungen die in der Psychotherapie stattfinden und mit Hilfe psychometrischer Verfahren erhoben werden, spielen diese Normen eine zentrale Rolle. Hier hat sich in der Psychotherapieforschung das Konzept der 'clinical significance' (vgl. Jacobson & Truax 1991) als Standard etabliert. Kennzeichnend für dieses Konzept ist die Einschätzung zweier Aspekte der Veränderung: zum einen die Größe der Veränderung und zum anderen die 'Lage' des (z.B. zur Entlassung) erreichten Wertes im Vergleich zu den Normwerten einer nicht-klinischen Population. Es ist beispielsweise möglich, daß sich ein Patient deutlich verbessert, aber der Wert bei Entlassung doch noch nicht

im Spektrum der Norm liegt. Wichtig sind dabei die klinischen Implikationen des Vergleichs mit den Normen: ein deutlich vom Mittelwert einer nicht-klinischen Normstichprobe abweichender Status, wird dabei als Hinweis auf Behandlungsbedarf angesehen.

Wie kann eine geeignete Referenz für eine *Symptomentwicklung* aussehen? Welche Implikationen für die Behandlung entstehen aus einem Vergleich mit einer 'Verlaufsnorm' oder 'Verlaufsreferenz'?

Ein erster Anhaltspunkt könnte darin liegen, im Verlauf der Behandlung kontinuierlich zu beobachten, ob ein Status erreicht wird, der im Normbereich liegt. Das entspricht einem Vergleich des aktuellen Wertes eines Patienten mit der o.g. (statischen) Ergebnissnorm. Er beantwortet die Frage danach, ob sich der Status in einem Bereich befindet (und ggf. stabilisiert), der dem einer nicht-klinischen Population entspricht und daher keine weitere Behandlung notwendig ist. Bei starkem Beeinträchtigungsgrad des Patienten, wird darin allerdings nur sehr wenig Informationsgehalt liegen, da der aktuelle Wert u.U. so weit von der Norm entfernt ist, daß es offensichtlich noch lange dauert, bis die Norm erreicht wird.

Im Hinblick auf die individuelle Behandlungsplanung ist die Frage viel relevanter, ob sich ein Patient so entwickelt, daß ein positiver Abschluß erwartet werden kann oder ob der Verlauf Anzeichen für eine unerwünschte Entwicklung aufweist. Ähnlich dem Vergleich eines einzigen Wertes mit dem Mittelwert und der Streuung einer 'gesunden' Stichprobe müßte für diesen Zweck der aktuelle Verlauf eines Patienten mit dem durchschnittlichen Verlauf und der Streuung *positiv endender Behandlungen* verglichen werden. Damit könnte zu jedem Zeitpunkt ermittelt werden, ob der Status eines Patienten im Rahmen dessen liegt, wie es bei erfolgreich Behandelten zum jeweiligen Zeitpunkt der Fall war. Diese Art von Referenzverlauf erforderte damit eine Normstichprobe an Verläufen und eine geeignete Technik, um den durchschnittlichen Verlauf und seine Streuung darzustellen; dies wird beispielhaft in Abschnitt 12.1 erläutert werden.

11.1.3 Rückmeldung der Ergebnisse an die behandelnden Personen

Zentrales Element eines QM-Systems ist die Rückmeldung der Ergebnisse an diejenigen Personen, die die Behandlung durchführen (Lutz 1997); fehlt dieses Element, besteht die Gefahr, daß die Informationen zwar gesammelt, aber nicht ausgewertet und 'vor Ort' also bei den Personen, die direkt an der Behandlung beteiligt sind, in konkrete Handlungen umgesetzt werden. Die Darstellung der Ergebnisse auf dieser Ebene ist sehr wichtig, da sie Transparenz nach innen schafft und die Motivation der Mitarbeiter erhöht (Kordy & Hannover 1998).

Die Rückmeldung der individuellen Ergebnisse zum Behandlungsende muß ebenfalls verschiedenen Anforderungen gerecht werden: sie sollte kurz sein, die wichtigsten Informationen über die Behandlungsmodalitäten sowie die eingetretenen Veränderungen

– möglichst in leicht erfaßbarer grafischer Form – enthalten. Im Stuttgart-Heidelberger Modell geschieht dies beispielsweise auf einem Rückmeldeblatt mit Hilfe von Dot-Charts und Tabellen, das Berner Modell (Grawe & Braun 1994) benutzt dafür mehrseitige ‘Effekt-’ und ‘Zustandskonfigurationen’.

Die Rückmeldung eines Verlaufes bis zum aktuellen Zeitpunkt hat den Vorteil, daß nicht alle Aspekte der Prä-Post-Bewertung enthalten sein müssen, sondern nur die Informationen, die tatsächlich im Verlauf erhoben werden. Sie erfordert eine Grafik mit dem bisherigen Verlauf seit Beginn der Behandlung und dem Referenzverlauf; dieser sollte mit einer Grenze versehen sein, die die Variabilität des Referenzverlaufes angibt.

11.1.4 Leichte Integrierbarkeit des Systems in den therapeutischen Alltag

Dieser Punkt ist für die praktische Anwendbarkeit eines Systems von entscheidender Bedeutung. Wenn die erforderlichen Abläufe zur Durchführung von QM (Verteilen, Ausfüllen und Einsammeln der Fragebögen, Auswertung und Verwaltung der Informationen, etc...) den Ablauf der Behandlung behindert oder sich als zu großer organisatorischer Aufwand herausstellt, wird die Motivation an der Teilnahme und dadurch die Glaubwürdigkeit der Angaben sinken, bzw. wird ein großer Teil die Teilnahme verweigern.

In einem verlaufsorientierten System erhöhen sich der organisatorische Aufwand, durch die regelmäßige Vergabe von Fragebögen während der Therapie. Sollen die erhobenen Informationen für die Planung der Behandlung genutzt werden, muß die Rückmeldung an den Therapeuten rasch erfolgen, um die Planung auf den aktuellsten Zustand des Patienten abzustimmen. Dazu ist eine computergestützte Erfassung, Auswertung und Darstellung der Ergebnisse unumgänglich. Der Prozeß vom Austeilen des Fragebogens bis zur Rückmeldung sollte nicht länger als ein paar Tage (bzw. höchstens eine Woche in ambulantem Behandlungssetting) dauern, um die Behandlung gegebenenfalls noch umstellen zu können.

11.1.5 Indikatoren zur Prognose des weiteren Verlaufes

Unter der Perspektive der Nutzbarkeit der Verlaufsinformation im Prozeß der Behandlungsplanung kommt bei verlaufsorientierten Systemen eine Anforderung an das QM-System hinzu, die es bei einem Prä-Post-Vergleich nicht gibt: es sollte valide Indikatoren zur Prognose des weiteren Verlaufs oder des Therapieergebnisses enthalten. Die Beobachtung des Verlaufes bietet die Möglichkeit zu zwei Aussagen. Zum einen liefert sie Information darüber, ob sich die Therapie gerade günstig oder ungünstig entwickelt; zum anderen bietet sich durch die Rückmeldung der Ergebnisse auch die Gelegenheit, eine

Prognose für diesen Patienten zu erstellen und diese Information in Behandlungsentscheidungen einfließen zu lassen. Dafür sollte die Prognose jedoch entsprechend valide sein.

11.2 Kritische Würdigung des COMPASS QM-Systems

Das bisher am meisten verlaufsorientierte QM-System ist das COMPASS-System, das von Howard, Brill, Lueger, O'Mahoney & Grant (1993) und ausführlicher von Lutz, Martinovich & Howard (1999) vorgestellt wurde. In diesem Abschnitt wird es in den Grundzügen beschrieben; dabei sollen die gerade aufgezählten Anforderungen eines verlaufsorientierten QM-Systems an dieses System angelegt und bewertet werden.

Das COMPASS-Programm wird von einem amerikanischen Versicherungsunternehmen (COMPASS) zum Behandlungsmanagement ambulanter Psychotherapien genutzt. Es stützt sich auf das Phasenmodell der Psychotherapie (Howard, Lueger, Maling & Martinovich 1993, vgl. Abschnitt 3.2.2). Über den Zeitraum der Therapie wird in nicht näher vorgeschriebenen (unregelmäßigen) Abständen der Zustand eines Patienten erhoben. Dazu entwickelten die Autoren ein Inventar, das die Konstrukte des Phasenmodells erfaßt: momentanes Wohlbefinden, Symptomatik und allgemeines Funktionsniveau. Die psychometrischen Eigenschaften des Inventars sind gut und erlauben eine Differenzierung zwischen Patienten und Nicht-Patienten (vgl. Sperry, Brill, Howard & Grissom 1996). Die Werte der drei Dimensionen werden zu einem 'Mental Health Index' (MHI) zusammengefaßt, der den allgemeinen Status des Patienten widerspiegelt; sie können aber auch getrennt untersucht werden. Das Ziel dieses Systems besteht darin, den Status der Patienten während der gesamten Therapiedauer zu erheben ('monitoring') und ihn ggf. an die Therapeuten zurückzumelden, um anstehende Behandlungsentscheidungen zu unterstützen.

Das Phasenmodell liefert Anhaltspunkte zur Formulierung von Erwartungen, die bei der Rückmeldung der Verläufe genutzt werden (Lueger 1993): Gemäß des Modells sollten nach vier Sitzungen das Wohlbefinden bzw. nach zwölf Sitzungen die Symptome Besserung aufweisen. Ist dies beispielsweise nach spätestens 6 bzw. 17 Sitzungen nicht der Fall, so wird dies dem Therapeuten berichtet und er wird angehalten, sein therapeutisches Vorgehen neu zu reflektieren.

Die Aussagekraft dieser Erwartungen bezüglich des weiteren Therapieverlaufs sind z.Tl. schon empirisch untersucht worden (Lueger 1998). Darin hatten Patienten mit einem verbesserten Wohlbefinden zur vierten Sitzung eine Erfolgsrate von 36% im Vergleich zu 18% die keine Verbesserung aufwiesen. Patienten, die nach zwölf Sitzungen eine gebesserte Symptomatik aufwiesen, hatten eine Erfolgsrate von 50% im Vergleich zu nur 7%, die zu diesem Zeitpunkt noch keine

symptomatische Besserung zeigten¹¹. Auch wenn Lueger (1998) die unterschiedlichen Eingangsbedingungen der Patienten nicht kontrollierte und es unklar bleibt, wie lange bzw. kurz vor dem Ende der Therapie der Response erfaßt wurde, ist auch diese Untersuchung ein Beleg dafür, daß die frühe Entwicklung in der Therapie mit dem endgültigen Therapieerfolg zusammenhängt. Gleichzeitig wird auch nochmal deutlich, daß es durchaus auch einige Fälle gibt, die keinen Response zeigen und trotzdem am Ende erfolgreich abschneiden,

Lutz, Martinovich & Howard (1999) präsentieren eine weitere Entwicklung des Systems: sie beschreiben, wie der aktuelle Verlauf eines Patienten mit dem aufgrund der Eingangsbedingungen ‘erwarteten’ Verlauf verglichen wird. Dieser erwartete Verlauf wird mit Hilfe von individuellen Wachstumskurven aus einer umfangreichen Stichprobe (N=890) errechnet. Gemäß des Phasenmodells wird für alle Patienten ein logarithmischer Verlauf der Symptombesserung angenommen: zu Beginn der Therapie bessern sich die Symptome schneller als in späteren Abschnitten. Die Modellierung anhand der individuellen Verlaufskurven erlaubt die Vorhersage der individuellen Verlaufparameter aus den Eingangsinformationen des QM-Inventars (Einzelheiten siehe Abschnitt 12.1.2, S. 88). Der große Vorteil gegenüber dem bisherigen Prozedere liegt in dem besser auf die individuelle Eingangslage abgestimmten erwarteten Verlauf. Wurden zuvor, wie oben beschrieben, die Implikationen des Phasenmodells auf den Verlauf aller Patienten übertragen, so wird nun der Verlauf eines jeden Patienten mit dem aufgrund seiner individuellen Eingangsinformation vorausberechneten Verlauf verglichen, um einzuschätzen, ob die Therapie auch ‘auf dem richtigen Kurs ist’ (oder wie Lutz et al. (1999) es formulieren: ‘... to address the most clinically relevant question, “Is treatment working?” ’ (S.571) Für diesen vorausberechneten Verlauf wird eine Art Toleranzgrenze (‘failure boundary’) konstruiert. Die Patienten, bei denen die Symptomwerte während der Therapie unterhalb der Grenze und damit hinter den Erwartungen zurück bleiben, werden an die Therapeuten zurückgemeldet.

Legt man die im vorigen Abschnitt explizierten Anforderungen für ein verlaufsorientiertes QM and das COMPASS-System an, so kommt man zu folgenden Feststellungen:

1. Erfassung der Therapieergebnisse auf individueller Ebene:

Der Verlauf der symptomatischen Entwicklung wird in diesem Modell mit sorgfältig entwickelten und ausführlich getesteten Instrumenten durchgeführt.

¹¹Die Kreuztabellen in diesem Artikel entsprechen dem Vorgehen in dieser Arbeit (Abschnitt 7.2, Tabelle 9) und ermöglichen einen Vergleich der einfachen Odds-Ratios: der Odds-Ratio der Responder vs. Non-Responder im Bezug auf das ‘momentane Wohlbefinden’ beträgt 2.6 (1.3, 5.2). Der einfache Odds-Ratio von AN betrug 1.6 (0.7, 3.7) und von BN 2.5 (1.7, 3.8). Der Odds-Ratio der ‘symptomatischen Responder’ im Vergleich zu den Non-Respondern beträgt 14 (5, 38). Diese Werte sind nicht mehr mit denen der vorliegenden Untersuchung vergleichbar, da Response nach ca. 12 Wochen (Sitzungen) erhoben wurde.

Ein besonderer Vorteil liegt darin, daß die drei Konstrukte des Phasenmodells getrennt erhoben werden.

2. Bewertung der Ergebnisse auf dem Hintergrund angemessener Standards:

Bei der Konstruktion des Standards ('erwarteter Verlauf') ist die umfangreiche Stichprobe mit wiederholten Erhebungen pro Patient positiv zu bewerten.

Als problematisch ist jedoch die Annahme einer generell logarithmischen Verlaufsgestalt anzusehen. Sie leitet sich zwar aus den Ergebnissen der Dosis-Wirkungs-Relation ab (Howard et al. 1986), bezieht sich dort aber nicht auf die *individuellen* Verläufe sondern auf den Zusammenhang der über Studien hinweg aggregierten Erfolgsraten und deren Behandlungsdauer. Die nachfolgenden Untersuchungen mit Hilfe von individuellen Verlaufskurven setzen die logarithmische Verlaufsgestalt bereits voraus anstatt zu explorieren, zu welchem Grad diese Annahme tatsächlich zutrifft (vgl. Howard et al. 1996, Sperry et al. 1996, Lueger 1998). Auch die Explorationen der Verläufe der vorliegenden Arbeit weisen auf eine größere Heterogenität hinsichtlich der Verlaufsgestalten hin.

Ein weiterer Kritikpunkt besteht in der Zusammensetzung der Stichprobe, die zur Berechnung des erwarteten Verlaufs herangezogen wird: Lutz et al. (1999) berechnen den Referenzverlauf aus der *gesamten* Stichprobe; d.h. daß sowohl positiv als auch negativ verlaufende Therapien darin Eingang gefunden haben. Der Referenzverlauf besteht also aus dem geschätzten mittleren Verlauf aller Therapien. Im Hinblick auf die Frage nach dem Handlungsbedarf in einer Therapie, wäre ein Vergleich des aktuellen Verlaufes mit dem durchschnittlichen Verlauf der *erfolgreich behandelten Patienten* von größerer Aussagekraft, ob die Behandlung 'auf dem richtigen Kurs ist'.

3. Rückmeldung der Ergebnisse an die behandelnden Personen:

Das System sieht eine rasche Rückmeldung an die Therapeuten vor, um ggf. die therapeutischen Strategien den besonderen Erfordernissen der Patienten anzupassen. Dies geschieht mit Hilfe einer Grafik, die den aktuellen Verlauf eines Patienten mit dem erwarteten Verlauf (und dessen 'Toleranzgrenze') abbildet.

4. Leichte Integrierbarkeit des Systems in den therapeutischen Alltag:

Das Inventar des Systems ist zwar relativ umfangreich, allerdings werden die erhobenen Daten mit Scanner eingelesen und mit entsprechender Computerunterstützung ausgewertet und aufbereitet, so daß sich der Aufwand in vertretbaren Grenzen hält.

Dieses System bietet eine anschauliche Repräsentation des aktuellen Verlaufes im Vergleich zu dem insgesamt zu erwartenden Verlauf; weitgehend unklar ist jedoch noch, wann ein neuer Verlauf so weit von der Erwartung abweicht, daß die *Prognose* deutlich

ungünstiger wird. Lutz et al. (1999) berichten, daß die Besserungsraten von Patienten, die zweimal während der Therapie unter die Toleranzgrenze fallen eine schlechtere Aussicht auf Besserung hatten (36%), als Patienten, die die Grenze nicht unterschritten (65%). An dieser Stelle sind noch weitere Indikatoren denkbar, die eine Prognose ermöglichen. Aus der Perspektive dieser Arbeit liegt insbesondere der Einbezug von Therapie-Response in die Prognose nahe, da Response zum einen mit einer höheren Erfolgswahrscheinlichkeit und zum anderen Responder und Non-Responder hinsichtlich ihrer Verlaufsgestalt unterscheiden. Wie dies geschehen kann, soll im Abschnitt 13 untersucht werden.

Zusammenfassung:

Die Anforderungen an ein verlaufsorientiertes QM liegen insbesondere in der Verwendung kurzer, veränderungssensitiver und doch valider Instrumente, die mehrfache Erhebungen ohne all zu großen Mehraufwand ermöglichen. Besondere Beachtung gilt der Definition verlaufsorientierter Normen, die einen Vergleich des aktuellen Verlaufs mit dem Verlauf *positiv verlaufender* Therapien ermöglichen sollten. Aufgrund der komplexeren Information ist eine computergestützte Erfassung, Verwaltung und grafische Aufbereitung unumgänglich. Die Rückmeldung an die Therapeuten erfordert ein rasches Vorgehen und schnelle Kommunikationswege, damit Patient und Therapeut noch im Verlauf der Therapie davon profitieren können. Die kontinuierliche Verlaufserhebung und Rückmeldung während der Therapie ermöglicht darüber hinaus die Entwicklung von Prognose-Indikatoren, die für die Planung der weiteren Behandlung mit herangezogen werden können.

Das COMPASS-System bietet bisher das am meisten verlaufsorientierte QM: es arbeitet mit ökonomischen und validen Instrumenten und hat sich als in die Praxis integrierbar erwiesen. Es ist das erste System, das im Bereich Psychotherapie den Patientenzustand kontinuierlich erhebt und an die Therapeuten zurückmeldet. Bei der Definition des Referenzverlaufs ist jedoch die generelle Annahme eines logarithmischen individuellen Symptomverlaufes kritisch zu sehen, sowie die Tatsache, daß der Referenzverlauf nicht zwischen positiv und negativ verlaufenden Therapien unterscheidet, sondern beide vermischt. Prognose-Indikatoren sind noch nicht erschöpfend untersucht worden.

Der Vergleich zwischen aktuellem Verlauf und Referenzverlauf ermöglicht die Entwicklung von Regeln, anhand derer eine Prognose des weiteren Verlaufes möglich wird – Therapie-Response als prognostisch relevantes Ereignis kann hier Berücksichtigung finden.

12 Illustration der Anwendung von Therapie-Response im QM

Die Exploration der Gruppen- und Einzelverläufe ergab, daß sich durch die zwei Bedingungen Therapie-Response (ja/nein) und Therapieergebnis (positiv/negativ) verschiedene Verlaufstypen ergeben, die auf empirischer Ebene ihre Entsprechung finden. Diese Verlaufstypen können als Referenz dienen, um den aktuellen Verlauf einer Patientin einzuschätzen und zusammen mit der vorhandenen klinischen Information die weitere Behandlung zu planen.

Wie ein Vorgehen dabei aussehen kann, wird in diesem Abschnitt beschrieben. Zur Illustration werden an dieser Stelle (wegen der Stichprobengröße und des eindeutigeren Zusammenhangs zwischen Response und Erfolg) nur die bulimischen Patientinnen herangezogen; als Therapie-Response wurde wegen der deutlicheren Implikationen auf den Symptomverlauf die ‘Verringerung der Symptome nach vier Wochen’ gewählt.

12.1 Grundidee und Methodik

12.1.1 Grundidee

Ziel ist es, den Besserungsverlauf einer Patientin mit den Verläufen solcher Patientinnen zu vergleichen, die unter vergleichbaren Anfangsbedingungen zu einem positivem Therapieergebnis kamen. Dieser Referenzverlauf wird durch ein statistisches Modell, das auf einer repräsentativen Stichprobe beruhen sollte, aus den zu Beginn der Behandlung bekannten Eigenschaften der Patienten konstruiert. Sollte der aktuelle Verlauf hinter diesem ‘Standard’ zurückbleiben, so könnte dies ein Warnsignal für die an der Therapie Beteiligten darstellen, daß die Therapie nicht den gewünschten Verlauf nimmt und ggf. Handlungsbedarf besteht. Die Bewertung des Vergleiches geschieht dabei nicht automatisch in den Kategorien ‘positiv’ und ‘negativ’, sondern dient als Instrument zur Entdeckung (‘flagging’) *potentieller* Defizite, die jedoch als solche erst eindeutig identifiziert werden müssen, bevor man daran gehen kann, geeignete Gegenmaßnahmen zu ergreifen. Der Vergleich von aktuellem Verlauf und Referenz muß an den Therapeuten oder das Team zurückgemeldet werden, um diese Information mit den klinischen Eindrücken zu verbinden und ggf. geeignete Interventionen einzuleiten. Die Rückmeldekomponente ist unverzichtbar im Hinblick auf die Interpretation der durch das QM-System bereitgestellten Information: Die bereitgestellten Informationen per se enthalten nämlich keinerlei Hinweise darauf, *ob* ein Zurückbleiben hinter dem Standard tatsächlich ein Defizit darstellt, oder ob es sich um eine durchaus positiv zu bewertende Sensibilisierung (vgl. Kordy & Lutz 1995) handelt. Ebenso läßt sich aus den Informationen nicht automatisch ableiten, *welche* Interventionen geeignet wären,

um eine tatsächlich ungünstig verlaufende Therapie einem positiven Ergebnis näher zu bringen. Erst die Verbindung von klinischer Information mit der Information des QM-Systems durch die behandelnden Therapeuten stellt eine angemessene Bewertung der Ergebnisse sicher. Soweit ist diese Konzeption mit der Idee des outcome monitoring der Arbeitsgruppe in Chicago (Howard et al. 1996) identisch.

Wegen den unterschiedlichen Verlaufsformen von Respondern und Non-Respondern ist es sinnvoll und angemessen, eine Patientin, die nach 4 Wochen Response zeigt, mit all jenen zu vergleichen, die bei ähnlichen Anfangsbedingungen ebenfalls Therapie-Response aufwiesen; und umgekehrt Patientinnen ohne Response nach 4 Wochen nur mit solchen zu vergleichen, die während ihrer Behandlung auch keinen Response hatten. Dies hat den Vorteil, daß a) homogenere Gruppen miteinander verglichen werden können und b) gleichzeitig über das Konzept des Therapie-Response sicherere Aussagen über Erfolg oder Mißerfolg, wie sie in den vorigen Abschnitten belegt wurden, möglich sind. Allerdings kann die Vorhersage des Verlaufs wegen des Einbezugs von Therapie-Response nicht mehr direkt vom Behandlungsbeginn an erfolgen, sondern erst nach einem Zeitraum von vier Wochen, dafür aber mit einem Zugewinn an Informationsgehalt. Zum Behandlungsbeginn kann eine Vorhersage lediglich ‘unter der Annahme von Response nach vier Wochen’ (oder Non-Response) erfolgen – oder man konstruiert ein Modell, das Therapie-Response nicht berücksichtigt, dafür allerdings Risiken in sich birgt (s. nächster Abschnitt).

Der Vorteil eines solchen outcome monitoring liegt darin, daß die Ergebnisse den noch laufenden Behandlungen zu gute kommen können. Der hier vorgeschlagene Ansatz baut das System von Howard und Mitarbeitern dahingehend aus, daß *zusätzlich* die Symptomentwicklung der ersten vier Wochen berücksichtigt wird, da sie wichtige Information über das Therapieergebnis beinhaltet.

12.1.2 Methodik

Hierarchische Lineare Modelle. Die Referenzverläufe müssen auf der Grundlage einer Stichprobe ermittelt werden, die von den Patientinnen mehrfache Erhebungen zu verschiedenen Zeitpunkten enthält und so die Symptomentwicklung abbilden; im vorliegenden Fall sind dies die per Kurzbeurteilung erhobenen Therapieverläufe aus der Studie MZ-ESS. Zur Abbildung von Form und Variabilität des Referenzverlaufes müssen die einzelnen Verläufe zusammengefaßt werden. Dazu eignen sich Hierarchische Lineare Modelle (Bryk & Raudenbush 1992, Snijders & Bosker 1999, Goldstein 1995). Sie ermöglichen eine Zusammenfassung der individuellen Verläufe zu einem ‘durchschnittlichen Gruppenverlauf’ und darüber hinaus die Schätzung der interindividuellen Variabilität um diesen mittleren Verlauf herum. Dazu wird für alle Verläufe ein mathematisches Modell gewählt, das jede einzelne Patientin repräsentiert – im einfachsten

Fall ist dies eine Gerade (siehe Gleichung 1). Sie ist durch die beiden Parameter ‘Höhe des Startpunktes’ (π_1 , Intercept) und ‘Größe der Steigung’ (π_2) definiert.

$$y_{it} = \pi_{1i} + \pi_{2i}t + e_{it} \quad (1)$$

Der Variabilität zwischen den einzelnen Patientinnen wird dadurch Rechnung getragen, daß jede Patientin einen individuellen Intercept und eine individuelle Steigung haben kann. Für jede Patientin i wird aus den wiederholten Messungen y zu den Zeitpunkten t eine Gerade berechnet mit dem Intercept π_{1i} , der Steigung π_{2i} und den entsprechenden Residuen e_{it} . Der Anfangszeitpunkt ist i.d.R. mit 0 kodiert (Beginn der Behandlung in unserem Fall), d.h. daß zu diesem Zeitpunkt der Ausdruck $\pi_{2i}t$ 0 ergibt und π_{1i} die direkte Schätzung der anfänglichen Symptomschwere darstellt. Den ‘mittleren Gruppenverlauf’ kann man sich als den Mittelwert der individuellen Intercepts π_{1i} und Steigungen π_{2i} vorstellen.

Das Verfahren ist eine Erweiterung üblicher Regressionsverfahren und erlaubt durch Einsetzen geeigneter Prädiktoren die Berechnung der Durchschnittsverläufe für verschiedene Subgruppen und Ausgangsparameter. Es handelt sich hierbei um eine multiple Regression mit festen und zufälligen Effekten (vgl. Searle 1988, Searle, Casella & McCulloch 1992, Richard 1996). Die zufälligen Effekte entsprechen im vorliegenden Fall den individuellen Verlaufsparemtern, d.h. den einzelnen Intercepts und Steigungen, welche eine gemeinsame bivariate Verteilung bilden. Es gibt daher in diesem Verfahren 3 zu schätzende Verteilungsparameter: die Verteilungen der Intercepts σ_{int}^2 und Steigungen σ_{stieg}^2 , sowie die der Residuen $\sigma_{\epsilon t}^2$. Der Parameter σ_{int}^2 bezeichnet also die Variabilität der Intercepts; der Parameter σ_{stieg}^2 gibt die Variabilität der Steigungen an und $\sigma_{\epsilon t}^2$ die Variabilität der einzelnen Messungen um die geschätzten Verläufe. Die Variabilität dieser Verteilungen kann durch das Einsetzen von Prädiktoren weiter aufgeklärt werden. Es ist also möglich, die Höhe des Intercepts und die Stärke der Steigung mit anderen Variablen (z.B. den Anfangscharakteristika der Patientinnen) wie in einer üblichen Regression in Verbindung zu bringen. So kann man mit diesem Verfahren aus den Anfangsbedingungen der Patientinnen den Intercept und die Steigung einer Patientin vorhersagen, wie es für die vorgesehene Anwendung nötig ist (siehe das Beispiel weiter unten Abschnitt 12.2.2).

Allerdings wird dabei die Annahme gemacht, daß die Form der Symptomverbesserung sich bei allen Patientinnen im *gleichen* mathematischen Modell fassen läßt. Unter der Perspektive des Therapie-Response ist diese Annahme, wie in Abschnitt 10.2 gezeigt, unrealistisch: Bei Patienten mit Therapie-Response verringern sich die Symptome anfangs wesentlich stärker als bei Non-Respondern; nicht erfolgreiche Responder zeichnen sich im Gegensatz zu den erfolgreichen durch einen erneuten Anstieg der Symptome aus. Diese Verlaufseigenschaften erlauben es nicht, für die vier eingangs genannten Bedingungen ein einheitliches Modell zu unterstellen. Es muß also jeweils

ein separates Modell konstruiert werden, welches die in sich relativ homogenen Verläufe zusammenfaßt. Die verschiedenen Verlaufsformen lassen sich durch geeignete Transformationen der Zeitachse erreichen: so kann bei den Respondern z.B. die anfangs starke und später langsamer werdende Reduktion der Symptome durch ein logarithmisches Modell abgebildet werden; dies wird durch die Transformation der Zeitachse t gewährleistet: $t' = \ln(t + 1)$. Die Entscheidung darüber, welches Modell am besten paßt, kann anhand eines Anpassungstests (Littell, Milliken, Stroup & Wolfinger 1996) und an der Größe der Fehlervarianz $\sigma_{\epsilon t}^2$ entschieden werden. Eine geringere Fehlervarianz bedeutet, daß die einzelnen Meßwerte weniger stark um die Modellverläufe streuen und daher die Beobachtungen besser abbilden.

‘Bandbreite’ der Verläufe. Der Vergleich mit dem Referenzmodell dient zur Abklärung der Frage, inwieweit der beobachtete Verlauf einer in Behandlung befindlichen Patientin im Bereich dessen liegt, wie es bei günstigen Therapieverläufen zu beobachten ist, oder nicht. Anders ausgedrückt: ob der bis dahin beobachtete Verlauf zur Klasse ‘der sich günstig entwickelnden Therapien’ gehört. Damit wird es notwendig, einen Bereich festzulegen, innerhalb dessen die Zugehörigkeit zum jeweiligen Referenzmodell eher wahrscheinlich ist und außerhalb dessen eher unwahrscheinlich. Daher ist bei der Darstellung der Referenzverläufe insbesondere ihre Variabilität zu beachten. Der Durchschnittsverlauf eignet sich nicht als Grenze, da – unter der Annahme der Normalverteilung der individuellen Verläufe um den mittleren Gruppenverlauf herum – 50% der Verläufe unterhalb und oberhalb des Mittels liegen. Es ist deshalb sinnvoller, eine Grenze zu wählen, jenseits derer nur noch relativ wenige Beobachtungen liegen. Diese Grenze wird hier der Einfachheit halber bei einer Standardabweichung vom Mittelwert festgesetzt, was bedeutet, daß diesseits etwa 70% der Beobachtungen liegen und jenseits davon nur noch etwa 30%.

Um diese Grenze zu finden, muß die gesamte Residualvarianz eines Modells ermittelt werden. Sie läßt sich mit folgender Gleichung (siehe Snijders & Bosker 1999, Gleichung 5.5) berechnen.

$$\sigma_{total}^2 = \sigma_{int}^2 + 2t \text{ cov}(int, steig) + t^2 \sigma_{steig}^2 + \sigma_{\epsilon}^2 \quad (2)$$

Das entspricht der Summe der Varianzen der Intercepts, der Steigungen zum Zeitpunkt t , ihrer Kovarianz und der Variabilität der einzelnen Messungen um die individuellen Verläufe. Mit diesem Ausdruck kann man die Standardabweichung vom mittleren Gruppenverlauf zu jedem Zeitpunkt t errechnen. Zum Anfangszeitpunkt $t = 0$ reduziert er sich auf

$$\sigma_{total}^2 = \sigma_{int}^2 + \sigma_{\epsilon}^2 \quad (3)$$

Da in der Stichprobe dieser Arbeit zu den späteren Zeitpunkten immer weniger Beobachtungen vorhanden sind, kann Gleichung 2 nicht einfach zur Berechnung der

Residual-Varianz zu allen Zeitpunkten herangezogen werden; dies führte zu einer Berechnung der Residual-Varianz, welche die mit Zunahme der Behandlungsdauer kleiner werdende Stichprobe vernachlässigt. Daher wurde in diesem Fall in mehreren Durchgängen ‘scheibchenweise’ vorgegangen¹²: Die gesamte Residualvarianz wurde immer nach Gleichung 3 berechnet; dabei wurde die Zeitachse so transformiert, daß der Nullpunkt an unterschiedlichen Zeitpunkten lag¹³, nämlich jeweils im Abstand von zwei Wochen. In jedem Durchgang waren nur noch diejenigen Patientinnen eingeschlossen, deren Behandlungsdauer den jeweiligen Nullpunkt überschritt. Folglich stammen die Schätzungen der Residual-Varianz zu den verschiedenen Zeitpunkten immer nur aus den Verläufen, die zu dem jeweiligen Zeitpunkt auch noch tatsächlich Beobachtungen enthielten.

12.2 Konstruktion der Referenzmodelle

Im folgenden werden nun die Referenzmodelle beschrieben. Anschließend soll die Anwendung dieser Modelle anhand einiger Beispiele illustriert werden. Dazu bleiben pro Bedingung sechs zufällig ausgewählte Patientinnen aus der Konstruktion der Referenz ausgeschlossen; diese dienen dann als diejenigen Verläufe, die mit den Referenzverläufen verglichen werden.

12.2.1 Festlegung der Verlaufsformen der 4 Referenzmodelle

Der erste Schritt besteht in der Aufteilung der Stichprobe in die 4 eingangs erwähnten Gruppen: zum einen Responder (mit und ohne Therapieerfolg) und zum anderen Non-Responder (mit und ohne Therapieerfolg). Die Konstruktion der Referenz erfolgte nach dem gleichen Prinzip wie es von Howard et al. (1996) beschrieben wird. Der wesentliche Unterschied liegt darin, daß hier für jede Gruppe *ein separates Modell mit unterschiedlicher Verlaufsform* erstellt wird, welches auf der Grundlage der zu Beginn der Behandlung bekannten Informationen und dem Therapie-Response die Verläufe der Patienten modelliert. Für alle vier Gruppen mußte dafür ein geeignetes mathematisches Modell zur Abbildung der individuellen Verläufe gewählt werden. Folgende mathematischen Modelle standen zur Auswahl: eine lineare oder quadratische Funktion, sowie (auf der Grundlage verschiedener Zeiteinheiten) mehrere logarithmische Funktionen und deren quadrierte Formen. Die Wahl der Zeiteinheit (Tage, Wochen, Monate) spielt für die logarithmische Verlaufsform eine Rolle: je ‘kleiner’ die Einheit, desto stärker die

¹²Ich danke Prof. J. Singer von der Harvard University, Boston, für diese Anregung

¹³D.h. die Zeitpunkte einer Patientin waren beim ersten Durchgang in realer Zeit (in Wochen), also z.B. mit 0, 2, 4, 6, kodiert; beim zweiten Durchgang wurde der Nullpunkt um zwei Wochen nach hinten verschoben (Kodierung: -2, 0, 2, 6), beim dritten Durchgang um 4 Wochen (Kodierung: -4, -2, 0, 2). So kann die Variabilität um den Nullpunkt zu verschiedenen Zeitpunkten geschätzt werden.

anfängliche Krümmung. Der Verlauf einer quadrierten logarithmischen Funktion ist nicht mehr – wie die nicht transformierte quadratische Funktion – symmetrisch, sondern fällt bzw. steigt zu Beginn stark und steigt bzw. fällt anschließend in schwächerem Maße. Die Auswahl aus diesen Modellen erfolgte aufgrund des o.g. Anpassungstests. Die Anpassungsgüte der linearen und logarithmischen Verläufe kann aufgrund der gleichen Anzahl von Parametern direkt verglichen werden: je geringer der Werte, desto besser die Anpassung. Die quadratischen Verlaufsformen wurden mit dem besten Modell der linearen und logarithmischen Modelle mittels des χ^2 -Anpassungstests verglichen (siehe Tabelle 21).

Tabelle 21: *Modellauswahl der vier Verlaufsgruppen*

Verlaufsform	$-2LL^{**}$	σ_ϵ^2	χ^2	$-2LL^{**}$	σ_ϵ^2	χ^2
<i>nicht erfolgreiche Non-Responder</i>			<i>erfolgreiche Non-Responder</i>			
linear	2378.2	0.19	–	654.7	0.19	–
log.(Monate)	2311.5	0.18	–	649.4	0.22	–
log.(Wochen)	2307.8	0.19	–	664.2	0.25	–
log.(Tage)	2330.7	0.20	–	683.7	0.29	–
quadratisch	2307.2	0.17	0.6	629.7	0.19	19.6*
log-quadr.(Monate)	2278.8	0.17	28.9*	†	†	–
log-quadr.(Wochen)	2279.5	0.18	28.3*	†	†	–
log-quadr.(Tage)	2276.9	0.18	30.8*	632.1	0.21	17.3*
<i>nicht erfolgreiche Responder</i>			<i>erfolgreiche Responder</i>			
linear	963.7	0.52	–	767.7	0.43	–
log.(Monate)	877.8	0.41	–	651.3	0.30	–
log.(Wochen)	798.8	0.32	–	554.9	0.22	–
log.(Tage)	717.3	0.26	–	471.7	0.17	–
quadratisch	807.3	0.31	-90.0*	598.6	0.23	-126.9*
log-quadr.(Monate)	647.5	0.20	69.7*	470.8	0.19	0.8
log-quadr.(Wochen)	564.0	0.16	153.2*	423.8	0.15	47.8*
log-quadr.(Tage)	525.7	0.16	191.5*	402.0	0.15	69.7*

χ^2 -Test bei 3 Freiheitsgraden

† = nicht konvergiert, * = $p < 0.01$, ** = -2 Log-Likelihood

Abbildung 11 gibt die durchschnittlichen Verläufe und Verlaufsformen der vier Gruppen wieder. Die resultierenden Gruppenmodelle stimmen recht gut mit den Erwartungen überein, wie sie sich aus Abschnitt 10.2 ergaben. Die durchschnittlichen Verläufe der Responder spiegeln die starke anfängliche Symptomreduktion wieder, wobei die nicht erfolgreichen Responder nach dem Response wieder einen Anstieg der Symptome verzeichnen. Bei den Non-Responder ist zu erkennen, daß die nicht erfolgreichen Patiententinnen nur in den ersten zwei Monaten eine leichte Symptomreduktion auf-

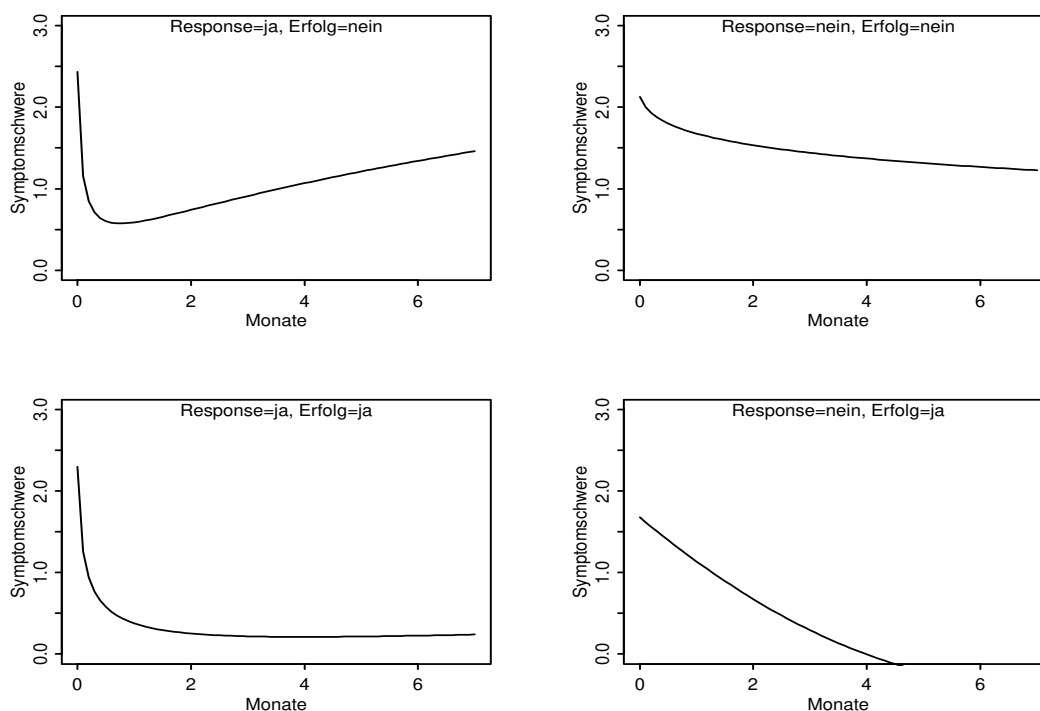


Abbildung 11: Durchschnittliche Verläufe der Responder und Non-Responder

weisen und sich danach kaum noch weiter verbessern und auf relativ hohem Niveau bleiben. Bei den erfolgreichen Non-Respondern hält die Symptomreduktion länger an. Sie haben daher im Durchschnitt nach vier bis fünf Monaten die Symptome verloren. Tabelle 22 zeigt die Funktionsgleichungen der vier Gruppen; Responder sowie nicht erfolgreiche Non-Responder können im gleichen mathematischen Modell beschrieben werden: ein quadratisches Modell auf der Basis der logarithmierten Behandlungsdauer (in Tagen), während die erfolgreichen Non-Responder eine ‘normale’ quadratische Funktion erfordern.

Tabelle 22: Die individuellen Modelle der 4 Referenzgruppen

Gruppe	Modell
Responder	$y_t = \pi_0 + \pi_1 \ln(\text{tage} + 1)_t + \pi_2 \ln(\text{tage} + 1)_t^2 + e_t$
nicht erfolgreiche Non-Responder	$y_t = \pi_0 + \pi_1 \ln(\text{tage} + 1)_t + \pi_2 \ln(\text{tage} + 1)_t^2 + e_t$
erfolgreiche Non-Responder	$y_t = \pi_0 + \pi_1 (\text{monate})_t + \pi_2 (\text{monate})_t^2 + e_t$

$y =$ Kurzbeurteilungsscore, $t =$ Zeitpunkt

12.2.2 Vorhersage des Verlaufs aufgrund der Anfangsinformation

Es ist wichtig zu beachten, daß die hier abgebildeten Verläufe nur einen Mittelwert darstellen, während die individuellen Verläufe in einer gewissen Bandbreite um diese Kurven streuen. Aufgrund der quadratischen Form ist nun jeder Verlauf durch 3 Parameter gekennzeichnet: Intercept, Steigung und Beschleunigung. Durch die Hinzunahme relevanter Variablen kann die Variabilität der individuellen Verlaufparameter erklärt und so die Modelle der Verläufe präzisiert werden. Folglich wurden für alle vier Gruppen separate Modelle konstruiert. Dazu wurde wieder das Verfahren des schrittweisen Hinzufügens relevanter Prädiktoren ('Stepwise Regression', Einschlusskriterium: $p \leq 0.1$) herangezogen. Zur Auswahl der Variablen dienten alle in Tabelle 10 (S. 57) aufgeführten Prädiktoren (ausschließlich der zu Beginn der Behandlung unbekanntem Faktoren, also Behandlungsdauer und -intensität). Statt der Variable 'niedrige Anfangssymptome' wird nun der Symptomschwere-Score aus den Kurzbeurteilungen zu Behandlungsbeginn verwendet.

Tabelle 23 zeigt beispielhaft die Ergebnisse für die Gruppe der nicht erfolgreichen Non-Responder. Der obere Teil gibt die Variabilität der zufälligen Effekte (σ_{int}^2 , σ_{steig}^2 , σ_{besch}^2 und σ_{ϵ}^2) wieder, der untere die Regressionsgewichte der durch die schrittweise Selektion erhaltenen Variablen und ihren Effekt auf die individuellen Verlaufparameter. Die Variabilität der Intercepts σ_{int}^2 wird notwendigerweise durch den Anfangsstatus bei Beginn der Behandlung gänzlich aufgeklärt, d.h. die individuellen Intercepts variieren in Abhängigkeit des Symptomschwere-Scores zum Behandlungsbeginn ('non-randomly varying intercepts', vgl. Bryk & Raudenbush 1992). Die Variabilität der Steigungen beträgt 0.09, die der Beschleunigungen 0.006. Der Signifikanztest deutet an, daß sich diese Schätzungen deutlich von 0 unterscheiden. Die Variabilität in den Verlaufparametern könnte also noch durch weitere, hier nicht erhobene Kovariaten aufgeklärt werden. Steigungen und Beschleunigungen korrelieren sehr stark zu -0.93. Die Koeffizienten der festen Effekte geben die Richtung des Einflusses auf die Verlaufparameter an: Die anfängliche Höhe des geschätzten Verlaufes variiert, wie schon erwähnt, ausschließlich mit der Stärke der Anfangssymptomatik; der Koeffizient von 0.99 belegt, daß der Anfangsstatus fast ohne Fehler modelliert wird. Die Steigungen und Beschleunigungen hängen ebenfalls von der anfänglichen Symptomschwere und weiter vom Wunschgewicht, dem Alter, der Erkrankungsdauer, dem SCL-90 GSI Scores und der Zugehörigkeit zur Klinikategorie ab. So haben beispielsweise ältere Patientinnen mit einem relativ hohen Wunschgewicht ein stärkeres Gefälle, aber (aufgrund der hohen Korrelation) gleichzeitig eine größere Beschleunigung, die die Reduktion der Symptome wieder 'abbremst'. Die Klinikategorien ergeben einen signifikanten Haupteffekt, während die Effekte der einzelnen Kategorien zur Kategorie vier (Dummy-Kodierung) keine bedeutsamen Unterschiede aufweisen.

Tabelle 23: Modell der nicht erfolgreichen Responder

zufällige Effekte						
Varianzen		Koeff.	se	Z	p > Z	
Anfangsstatus	σ_{int}^2	–	–	–	–	
Steigung	σ_{steig}^2	0.096	0.026	3.67	0.0002	
Beschleunigung	σ_{besch}^2	0.006	0.001	3.80	0.0001	
Fehler	σ_{ϵ}^2	0.122	0.006	19.64	0.0001	
feste Effekte						
Parameter		Koeff.	se	df	T	Pr > T
<i>Effekte auf den Anfangsstatus</i>						
anfängliche Symptomschwere		0.998	0.034	776	29.17	0.0001
<i>Effekte auf die Steigung</i>						
	Intercept	0.421	0.185	278	2.27	0.0238
	Wunschgewicht	-0.002	0.0007	776	-2.60	0.0095
	Alter	-0.012	0.006	776	-1.86	0.0640
	ln(Erkrankungsdauer)	-0.077	0.058	776	-1.32	0.1867
	SCL-90 GSI	0.096	0.049	776	1.96	0.0509
	Klinikategorie 1	0.091	0.105	776	0.87	0.3862
	Klinikategorie 2	-0.136	0.110	776	-1.23	0.2176
	Klinikategorie 3	0.132	0.106	776	1.24	0.2153
<i>Effekte auf die Beschleunigung</i>						
	Intercept	-0.049	0.045	278	-1.09	0.2784
	anfängliche Symptomschwere	-0.037	0.004	776	-9.19	0.0001
	Alter	0.003	0.001	776	1.80	0.0717
	ln(Erkrankungsdauer)	0.019	0.014	776	1.31	0.1892
	SCL-90 GSI	-0.018	0.012	776	-1.48	0.1397
	Klinikategorie 1	0.013	0.026	776	0.48	0.6301
	Klinikategorie 2	0.052	0.027	776	1.87	0.0615
	Klinikategorie 3	0.003	0.026	776	0.12	0.9053

Durch Einsetzen der Koeffizienten in die Regressionsgleichung kann nun – wie bei einer normalen Regression – ein Referenzmodell berechnet werden, wie es sich aus den Daten der nicht erfolgreichen Non-Responder ergibt. Für eine 20 Jahre alte Patientin mit einer anfänglichen Symptomschwere von 2.5 laut Kurzbeurteilung, einem Wunschgewicht von 76% des erwarteten BMI, einem SCL-90 GSI-Wert von 1.6 und einer Erkrankungsdauer von 3.23 Jahren, die in einer Klinik ohne eßstörungsspezifischem Therapieprogramm behandelt wird, errechnet sich die Symptomschwere folglich aus den Gleichungen:

Gleichung (4):

$$0.998 * 2.5 = 2.49 \quad (\text{zum Aufnahmezeitpunkt})$$

Gleichung (5):

$$\begin{aligned} &2.49 \\ &+0.421*t \quad -0.002*76*t \quad -0.012*20*t \quad -0.077*\ln(3.23+1)*t \quad +0.096*1.6*t \quad +0.091*t \\ &-0.049*t^2 \quad -0.037*2.5*t^2 \quad +0.003*20*t^2 \quad +0.019*\ln(3.23+1)*t^2 \quad -0.018*1.6*t^2 \quad +0.013*t^2 \\ &= 2.24 \quad (\text{nach 28 Tagen}) \end{aligned}$$

Wobei in Gleichung 5 der erste Werte aus Gleichung 4 stammt und t für die logarithmierte Anzahl von Tagen seit Behandlungsbeginn steht. Die Grenzen, die die Variabilität dieses vorhergesagten Verlaufs markieren, ergeben sich aus der im zweiwöchigen Abstand berechneten Residualvarianz dieses Modells (siehe Abschnitt 12.1.2, Gleichung 3): die obere und untere Grenze ist zum jeweiligen Zeitpunkt eine Standardabweichung höher bzw. niedriger als der vorhergesagte Wert.

Die Angaben dieser neuen Patientin können ebenso in das Modell eingesetzt werden, wie es sich aus den *erfolgreichen* Non-Respondern herleitet, auch wenn dieser Verlauf weniger wahrscheinlich ist. Der Vergleich mit beiden Referenzverläufen wird im nächsten Abschnitt demonstriert.

12.3 Vergleich von ‘neuen’ Verläufen mit den Referenzmodellen

In diesem Abschnitt soll demonstriert werden, wie die Referenzverläufe im Rahmen eines Qualitätsmanagements angewendet werden können. Nehmen wir also an, es gäbe eine Klinik, in der die symptomatische Entwicklung der Patientinnen mit Eßstörungen kontinuierlich mit Hilfe der Kurzbeurteilungen erhoben und mit den entsprechenden Referenzverläufen verglichen werden. Dieser Vergleich geschieht während die betreffenden Patientinnen noch in Behandlung sind und soll die anstehenden Behandlungsentscheidungen unterstützen.

Alle hier dargestellten Verläufe stammen von den 24 Patientinnen, die nicht in die Modellierung der Referenzverläufe miteinbezogen wurde (siehe Abschnitt 12.2); damit handelt es sich bei den Referenzverläufen um echte Vorhersagen der symptomatischen Entwicklung aus den Anfangsinformationen der Patientinnen.

12.3.1 Beispiel 1: Responder mit Therapieerfolg

Abbildung 12 zeigt den aktuellen Verlauf einer Patientin zu vier verschiedenen Zeitpunkten im Vergleich zu den Referenzverläufen. Die ausgewählte Person gehört zu der Gruppe von Patientinnen, die nach vier Wochen auf die Therapie angesprochen haben

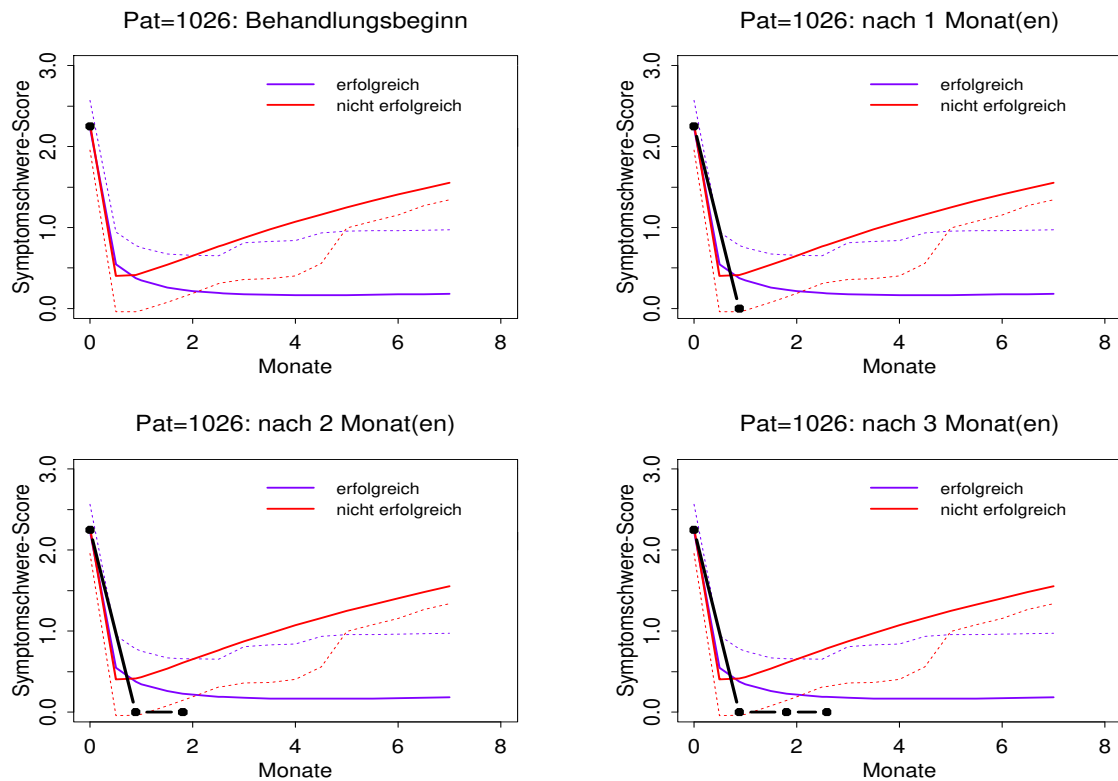


Abbildung 12: Aktueller Verlauf vs. Referenz (1)

und am Ende der Behandlung als Erfolg gewertet wurden. Jedes Kästchen zeigt den Verlauf nach weiteren vier Wochen (schwarze Linie). Die farbigen Linien stellen die durchschnittlichen Verläufe von Patientinnen mit gleichen Anfangsvariablen dar, die ebenfalls Therapie-Response aufwiesen: die blauen Linien stammen von Patientinnen, die am Ende erfolgreich abgeschnitten hatten und die roten Linien von solchen, die bei Behandlungsende keinen Therapieerfolg hatten. Die unterbrochenen Linien markieren die obere bzw. untere Grenze der Referenzverläufe. Beim Referenzverlauf erfolgreicher Patientinnen interessiert die obere Grenze der Bandbreite mit den höchsten Scores; beim Referenzverlauf nicht erfolgreicher Patientinnen ist umgekehrt die untere Grenze der Bandbreite mit den niedrigsten Symptomschwere-Scores von Interesse. Mit diesen Grenzen werden zwei sich teilweise überlappende Räume eröffnet, in denen sich Symptomschwere-Scores der eher erfolgreichen bzw. nicht erfolgreichen bewegen. Je weiter diese beiden Grenzen auseinanderliegen, desto sicherer lassen sich erfolgreiche von den nicht erfolgreichen Patientinnen unterscheiden.

Wie man sieht, verlaufen die beiden Referenzlinien im ersten Monat fast gleich – wie es für Patientinnen mit Therapie-Response auch zu erwarten ist. Zwischen Monat 1 und 2 beginnen sie sich auseinanderzuentwickeln. Die Bandbreite beider Referenzen überlappen sich teilweise und haben erst nach fünf Monaten keine Überlappungen mehr.

Nach vier Wochen (Bild rechts oben) hat die Patientin ihre Symptome fast ganz verloren, und liegt sogar unterhalb des Mittelwertes erfolgreicher Verläufe (blaue durchgezogene Linie); gleichzeitig liegt sie aber immer noch im Bereich einer Standardabweichung derjenigen Verläufe, die am Ende keinen Therapieerfolg aufweisen (unterbrochene rote Linie). D.h. auch bei Verläufen mit letztlich negativem Ausgang ist bei Respondern anfänglich ein vollständiger Verlust der Symptome möglich, ob er sich jedoch auf Dauer hält, muß sich erst noch herausstellen. Die Wahrscheinlichkeit, daß diese Patientin erfolgreich abschneiden wird, ist aufgrund der Ergebnisse in Abschnitt 7 jedenfalls höher, als die Wahrscheinlichkeit die Therapie ohne Erfolg zu beenden.

Nach fast zwei Monaten Behandlung (Bild links unten) ist die Patientin immer noch symptomfrei. Im Vergleich zu den erfolgreichen Verläufen liegt sie immer noch besser als der Durchschnitt, im Vergleich zu den nicht erfolgreichen Verläufen hat sie gerade den Bereich von knapp über einer Standardabweichung vom Mittelwert verlassen; sie ist damit in einem Bereich, wo sich weniger als 30% der letztlich nicht erfolgreichen Patientinnen bewegen. Nehmen wir an, der behandelnde Therapeut bekäme diese Grafik zum Verlauf dieser Patientin im Rahmen des Qualitätsmanagements der Klinik zurückgemeldet. Die Information der Grafik trägt die Botschaft, daß diese Patientin zu den Respondern gehört und damit eine erhöhte Wahrscheinlichkeit für Therapieerfolg besteht. Nach zwei Monaten Behandlung ist die nach vier Wochen eingetretene Symptomfreiheit stabil geblieben. Die Patientin liegt in ihrer bisherigen Symptomentwicklung besser als der Durchschnitt vergleichbarer erfolgreich verlaufender Therapien und hat gerade den Bereich verlassen, in dem sich die überwiegende Mehrheit der nicht erfolgreich verlaufenden Therapien befinden. An dieser Stelle können sich für die weitere Behandlung wichtige Frage stellen: Wie schätzt der Therapeut die bisherige Entwicklung ein? Gibt es Anzeichen dafür, daß trotz dieser positiven Entwicklung mit einem raschen Rückfall zu rechnen ist (z.B. hat die Patientin nur deshalb keine Symptome mehr, weil sie unter strenger Kontrolle steht - oder sind die Verbesserungen ohne Kontrolle eingetreten)? Sind die bisherigen Veränderungen aller Voraussicht nach stabil genug, daß eine weitere stationäre Behandlung nicht mehr angezeigt ist und die Patientin in weitere ambulante Therapie vermittelt werden kann? Diese Fragen lassen sich *nicht allein* aus den Informationen des Qualitätsmanagement beantworten, sondern bedürfen der Einschätzung der behandelnden Therapeuten unter der Berücksichtigung weiterer klinischer Informationen. Insofern stellt die Rückmeldung des Verlaufes nur *einen* Baustein für die weitere Behandlungsplanung dar – der Vorteil liegt in der systematischen Nutzung des bisherigen Wissens über Therapieverläufe.

Die Grafik des Verlaufes dieser Patientin nach fast drei Monaten (rechts unten) – für diese Patientin war das der Entlassungszeitpunkt – zeigt, daß sie auch nach dieser Zeit immer noch symptomfrei ist. Sie liegt nun in einem Bereich, in dem nicht erfolgreiche Patientinnen kaum noch zu finden sind.

12.3.2 Beispiel 2: Non-Responder mit Therapieerfolg

In Abbildung 13 ist der Verlauf einer Patientin ohne Therapie-Response dargestellt, die bei Behandlungsende als 'erfolgreich' eingestuft wurde. Die Referenzverläufe un-

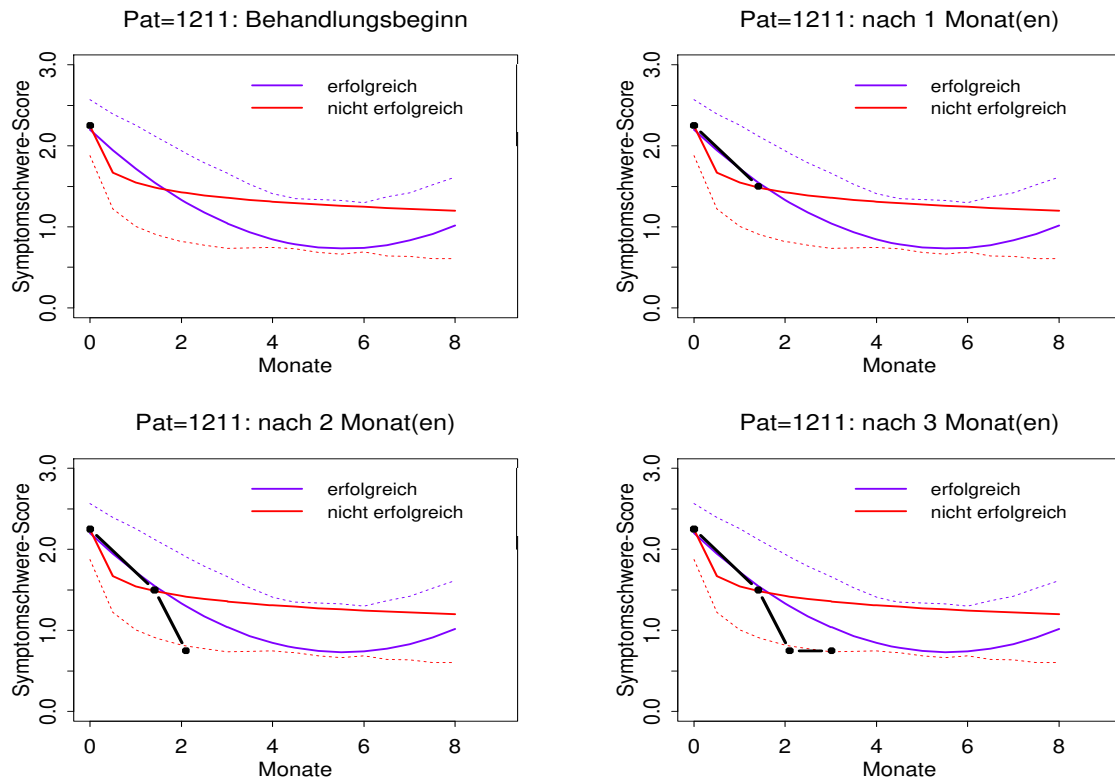


Abbildung 13: Aktueller Verlauf vs. Referenz (2)

terscheiden sich stark hinsichtlich des Gefälles: die nicht erfolgreichen verringern sich innerhalb der ersten zwei Monate um ca. 0.5 Punkte und anschließend nur noch wenig; die erfolgreichen haben über die angezeigten acht Monate hinweg ein deutliches Gefälle, aber erreichen erst nach drei Monaten das Ausmaß an Verringerung, daß Responder schon nach vier Wochen aufweisen. Die Bandbreite der Referenzverläufe weist bis acht Monate nach Behandlungsbeginn sich überlappende Bereiche auf.

Die Patientin startet auf recht hohem Niveau (2.25) und reduziert innerhalb von vier Wochen ihre Symptome auf einen Kurzbeurteilungs-Score von 1.5 (Bild rechts oben). Sie folgt damit in etwa dem durchschnittlichen Verlauf vergleichbarer Patientinnen mit erfolgreichem Therapieende; zu diesem Zeitpunkt haben jedoch auch Patientinnen mit negativem Therapieausgang häufig eine ähnlich hohe Symptomatik. Nach acht Wochen (Bild links unten) ist der Score weiter auf etwa 0.75 gesunken. Dieser Wert ist zu diesem Zeitpunkt für Patientinnen, die die Therapie nicht erfolgreich beenden wenig wahrscheinlich und liegt auch unterhalb des Mittelwertes der erfolgreichen Patientinnen. Bekäme ein Therapeut nach zwei Monaten Behandlung dieses Bild als Rückmel-

dung, so hätte er die bisherige erfolgreiche Entwicklung einzuschätzen. Als zusätzliche Information wäre bekannt, daß Non-Responder eher eine höhere Wahrscheinlichkeit auf einen negativen Therapieausgang haben. Zu diesem Zeitpunkt stellt sich, wie bereits bei voriger Patientin, die Frage, ob das erreichte Niveau erhalten bleibt oder nicht. Die Rückmeldung enthält hier eine gegenläufige Botschaft: a) die Patientin hat aufgrund des Non-Response ein höheres Risiko für einen negativen Therapieausgang, so daß ein Wiederanstieg der Symptome nicht sehr überraschend wäre und b) der bisherige Verlauf zeigt eine (im Vergleich zu den Respondern relativ langsame) kontinuierliche Verbesserung auf ein Niveau, das bei nicht erfolgreichen Patientinnen recht selten erreicht wird. In Verbindung mit den klinischen Eindrücken aus der Therapie und den aktuellen Rahmenbedingungen der Patientin, kann der Therapeut entscheiden, welche Maßnahmen für eine weitere Stabilisierung des Zustandes geeignet erscheinen. Die höhere Wahrscheinlichkeit für negativen Therapieausgang der Non-Responder legt es nahe, trotz des bisher positiven Verlaufes auf Anzeichen einer etwaigen Verschlechterung zu achten und ggf. gegenzusteuern.

Nach weiteren vier Wochen (Bild rechts unten) hat sich der erreichte Zustand als stabil erwiesen und die Patientin wurde mit leichter Restsymptomatik entlassen. Der Wiederanstieg des Referenzverlaufes der erfolgreichen Patientinnen ab Monat 6 stellt hier ein Artefakt dar, da nur recht wenige Patientinnen mit so langer Beobachtungsdauer in der Stichprobe enthalten sind.

12.3.3 Beispiel 3: Responder ohne Therapieerfolg

In Abbildung 14 ist der Verlauf einer Patientin zu sehen, die nach vier Wochen auf die Therapie anspricht, aber bei Behandlungsende als 'nicht erfolgreich' eingestuft wurde. Hier fällt auf, daß sich die Referenzverläufe schon sehr bald nach dem Response in gegensätzliche Richtungen entwickeln. Bereits nach zwei Monaten überlappen sie ihre Bandbreiten nicht mehr. Es scheint so, als ob sich bei Patientinnen mit der vorliegenden Ausgangssituation recht schnell entscheidet, ob die Behandlung erfolgreich verläuft oder nicht. Begann die Patientin mit sehr starken Symptomen, so sind sie nach vier Wochen fast vollständig verschwunden. Eine Rückmeldung zu diesem Zeitpunkt enthielte eine recht positive Prognose: da die Patientin zu den Respondern gehört, ist es wahrscheinlicher, daß sie dem grünen Referenzverlauf der erfolgreich behandelten Patientinnen folgen wird; zusätzlich befindet sie sich bereits zu diesem Zeitpunkt auf einem so niedrigen Niveau, das für nicht erfolgreich abschneidende Patientinnen relativ unwahrscheinlich ist – die Behandlung verläuft nach diesen Angaben bisher sehr erfolgversprechend. Wiederum wäre es hier die Aufgabe des therapeutischen Teams, diese Angaben mit dem klinischen Eindruck zu verbinden und daraus Schlüsse für die weitere Behandlungsplanung zu ziehen.

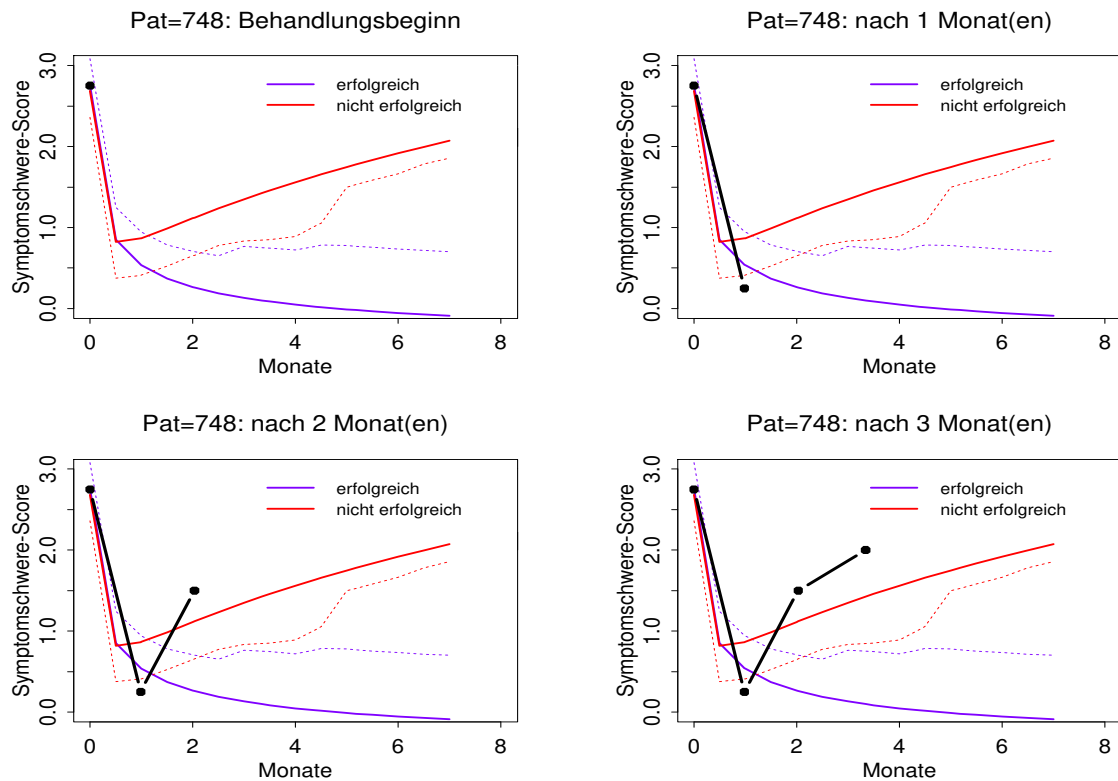


Abbildung 14: Aktueller Verlauf vs. Referenz (3)

Nach einem weiteren Monat hat sich der Zustand der Patientin jedoch wieder erheblich verschlechtert. Nun befindet sich die Symptomsschwere wieder im Bereich der eher erfolglos verlaufenden Behandlungen und bereits weit außerhalb des Bereichs der erfolgreich verlaufenden. Der Verlauf beginnt zunehmend dem der nicht erfolgreichen Behandlungen zu ähneln, trotz der günstigeren Prognose als Responder.

Nach drei Monaten haben die Symptome erneut etwas zugenommen, sodaß der Wert inzwischen sehr weit vom Bereich der erfolgreichen Verläufe entfernt ist. An dieser Stelle scheint die Symptomsschwere deutlich auf einen ungünstigen Behandlungsverlauf hinzuweisen. Kommt der Therapeut in seiner Einschätzung ebenfalls zu dem Ergebnis, daß die Behandlung nicht wie gewünscht verläuft, kann die Rückmeldung des bisherigen Verlaufes die Einschätzung des behandelnden Therapeuten untermauern und dazu anregen, das bisherige Vorgehen zu überprüfen, nach möglichen Ursachen für diese negative Entwicklung zu suchen und ggf. Veränderungen in der Behandlungsstrategie zu veranlassen. Das Resultat muß dabei keineswegs die Beendigung der Therapie sein, sondern z.B. die Zuweisung zu einer anderen Behandlungsgruppe, falls sich etwa herausstellt, daß sich die Patientin in zu starke Konflikte mit anderen Gruppenmitgliedern verwickelt und daraufhin zurückgezogen hat.

12.3.4 Beispiel 4: Non-Responder ohne Therapieerfolg

In Abbildung 15 ist der Verlauf einer Patientin zu sehen, die nach vier Wochen noch nicht auf die Therapie anspricht und bei Behandlungsende als 'nicht erfolgreich' eingestuft wurde. Auch bei dieser Patientin unterscheiden sich die beiden Referenzverläufe

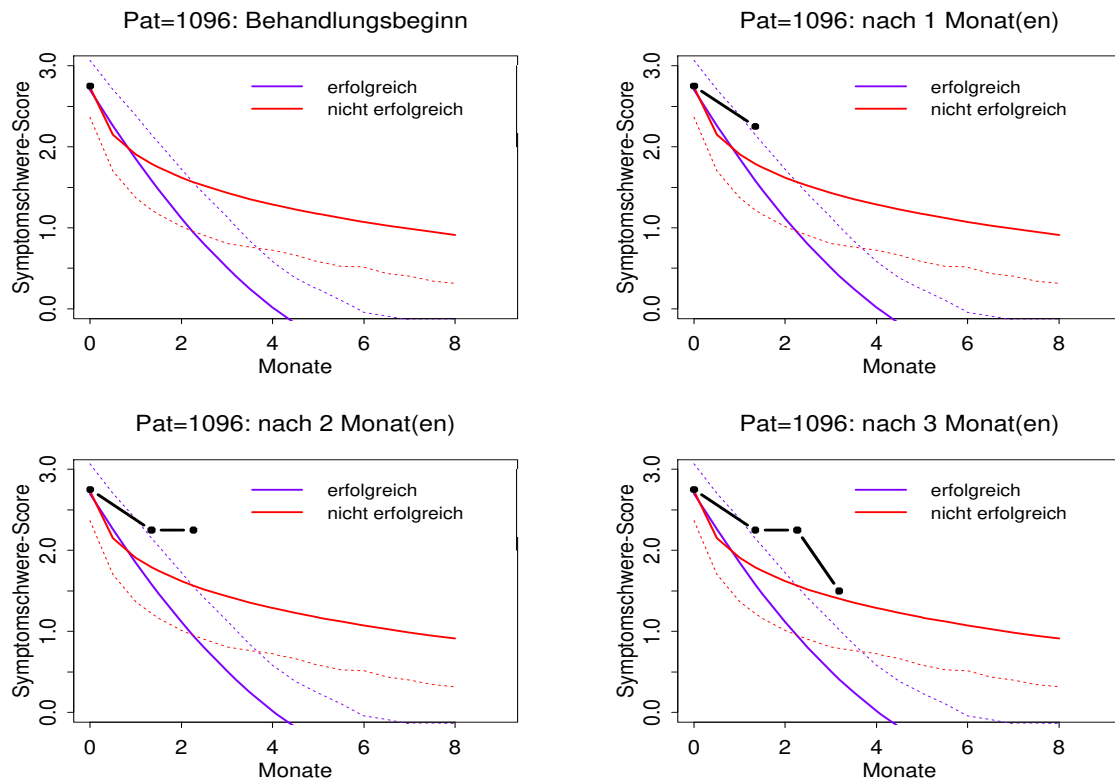


Abbildung 15: Aktueller Verlauf vs. Referenz (4)

deutlich: auch wenn nach vier Wochen kein Therapie-Response vorliegt, sind erfolgreich behandelte, vergleichbare Patientinnen im Durchschnitt nach vier Monaten symptomfrei. Patientinnen mit ungünstigem Therapieverlauf verlieren innerhalb der ersten zwei Monate zwar auch durchschnittlich 0.75 Punkte, aber nach diesem Zeitpunkt hält dieser Trend nicht an, sondern verlangsamt sich und bleibt auf relativ hohem Niveau. Die Patientin in diesem Beispiel verbessert sich innerhalb von zwei Monaten nur wenig und bleibt oberhalb des Mittelwertes der nicht erfolgreichen Verläufe und hat sich auch schon aus dem Bereich der erfolgreichen Verläufe herausbewegt; nach etwas mehr als zwei Monaten liegt der Symptomschwere-Score weit außerhalb des Bereiches, in dem sich die meisten erfolgreich verlaufenden Patientinnen bewegen. Auch nach einer weiteren Reduktion nach drei Monaten Aufenthalt in der Klinik verläßt sie nicht den Bereich der ungünstig verlaufenden Therapien und erreicht trotz weiterer Verbesserung auch nicht den Bereich der erfolgreichen Verläufe. An dieser Stelle sieht die Prognose der Patientin eher ungünstig aus: Als Non-Responder gehört sie zu der Gruppe mit

ungünstigeren Erfolgsaussichten und selbst nach drei Monaten Behandlung sind nur mäßige Fortschritte in der Symptomreduktion zu verzeichnen und scheint damit diese Prognose zu bestätigen. Nun liegt es am therapeutischen Team, ihr Wissen über diese Patientin mit dieser Information zu verknüpfen und abzuschätzen, ob sie sich evtl. doch langsam weiter verbessern wird oder ob es andere Faktoren gibt, die insgesamt eher für eine negative Prognose sprechen. Einigt sich das Team auf eine negative Prognose, so stünde als nächstes die Frage im Raum, welche Veränderungen in der Behandlung eine weitere Verbesserung ihres Zustandes ermöglichen.

12.3.5 Beispiel 5: unklare Klassifizierung

Die bisherigen Beispiele mögen den Eindruck geben, die Modelle erlauben nach gewisser Zeit eine recht eindeutige Klassifikation in erfolgreich und nicht erfolgreich verlaufende Entwicklungen. Abbildung 16 gibt ein Beispiel für eine uneindeutige Situation: die

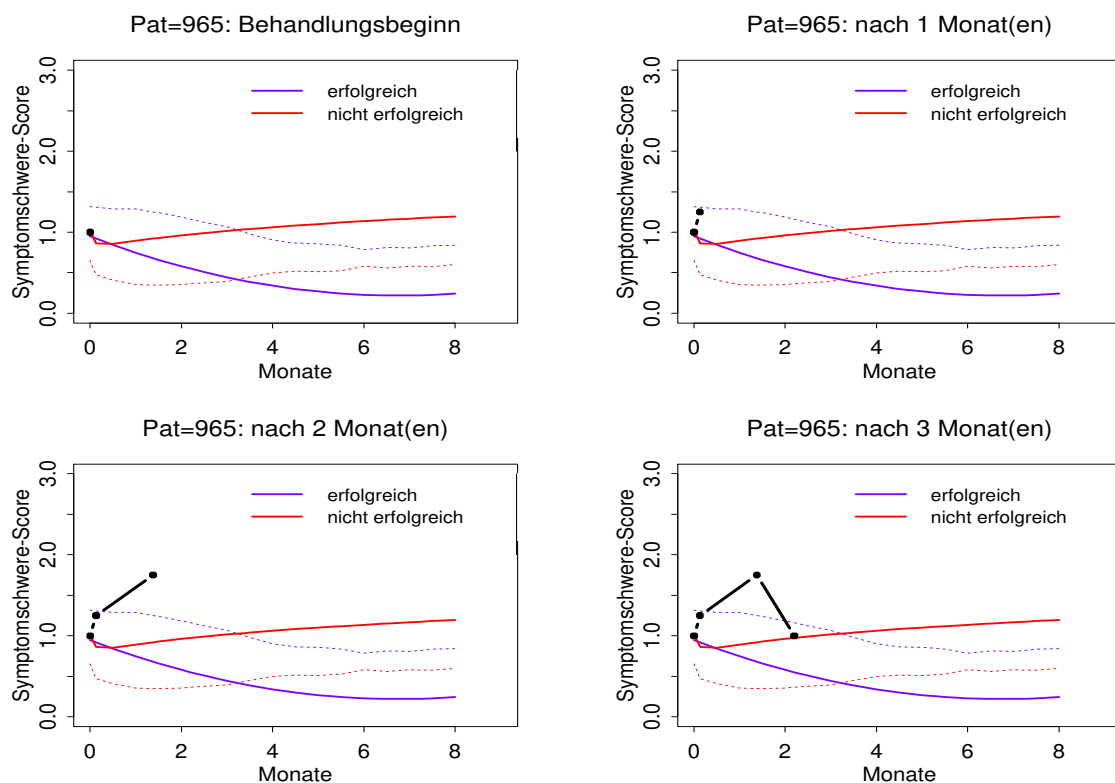


Abbildung 16: Aktueller Verlauf vs. Referenz (5)

Referenzen der erfolgreichen und nicht erfolgreichen Verläufe unterscheiden sich nur hinsichtlich des allgemeinen Trends, überlappen sich jedoch über den gesamten Zeitraum. Offensichtlich wurden vergleichbare Patientinnen, die auf diesem relativ niedrigen Niveau starteten und es auch beibehielten bzw. noch etwas weiter reduzierten,

mal als Erfolg und mal als nicht erfolgreich bewertet. Dies mag mit dem Erfolgskriterium zusammenhängen, das ja nicht aus den Kurzbeurteilungen selbst stammt (vgl. Abschnitt 6.2). Hier können geringfügige Unterschiede zwischen den Angaben in der Kurzbeurteilung und den Entlassungsfragebögen – z.B. etwas höhere Symptome in den Entlassungsfragebögen – zu einer negativen Erfolgsbeurteilung geführt haben auch wenn in der Kurzbeurteilung ein recht niedriger Score vorliegt. Die Interpretation des Vergleichs zwischen tatsächlichem Verlauf und Referenz bleibt daher vage: die Prognose aufgrund des nicht vorhandenen Therapie-Response ist weniger günstig – allerdings konnte die Patientin wegen des geringen Eingangsniveaus auch keinen Response zeigen. Die niedrigen Symptome weisen auf eine eher geringe Beeinträchtigung der Patientin hin, was wiederum ein eher günstiger Faktor ist. Die Patientin verschlechtert sich etwas innerhalb der ersten zwei Wochen, liegt aber immer noch innerhalb Bandbreite der günstigen Verläufe. Nach zwei Monaten hat sich der Zustand weiter verschlechtert. Nun liegt der Score deutlich außerhalb der günstigen Verläufe. Hier ist die Frage zu stellen, wie diese Entwicklung zu bewerten ist. Als Non-Responder mit relativ geringer Anfangssymptomatik ist die Prognose weder deutlich positiv noch deutlich negativ; im Verlauf hat der Zustand zunehmend verschlechtert, so daß hier anscheinend eine eher negativ zu bewertende Entwicklung stattgefunden hat.

Für eine empirisch unterstützte Prognose der Verläufe ist an dieser Stelle insbesondere von Interesse, welche prognostische Bedeutung den Bandbreiten zukommt. Hat eine Patientin tatsächlich eine geringere Chance auf Therapieerfolg, wenn ihr Symptomsschwere-Score nach vier oder acht Wochen außerhalb der Bandbreite erfolgreicher Verläufe liegt? Oder ist umgekehrt ein Wert innerhalb dieses Bereichs ein sicheres Zeichen für einen positives Therapieergebnis? Dieser Frage wird im folgenden Abschnitt, der den letzten Schritt in der Analyse von Therapie-Response darstellt, nachgegangen.

13 Die Bandbreite positiver Verläufe als Prognose-Indikator

13.1 Der Prognose-Indikator als zu überprüfende Hypothese

Bisher wissen wir, daß Therapie-Response mit einer im Vergleich zu Non-Respondern 2 bis 3-fach erhöhten Wahrscheinlichkeit auf Therapieerfolg verknüpft ist und daß sich die Verlaufsgestalten der Responder und Non-Responder unterscheiden. Der Vergleich des aktuellen Verlaufs mit den Referenzen ermöglicht eine optische Abschätzung, inwiefern der aktuelle Verlauf eher dem Muster der erfolgreichen oder nicht erfolgreichen

Therapien folgt. Die Konstruktion der Referenzverläufe und ihrer Bandbreite ist allerdings recht aufwendig und wirft die Frage auf, ob sich dieser Aufwand lohnt. Was bedeutet es, wenn sich die Symptomschwere einer Patientin zu einem gegebenen Zeitpunkt außerhalb des Bereiches der erfolgreichen Verläufe befindet? Hat eine solche Patientin tatsächlich geringere Erfolgschancen? Es geht im folgenden also darum, die Bedeutsamkeit der Bandbreiten zu untersuchen und zu klären, ob mit ihrer Hilfe die Prognose des Therapieerfolgs oder des weiteren Verlaufs präzisierbar ist.

Mehrere Varianten sind denkbar, wie solche Prognose-Indikatoren aussehen könnten. Aus den Abbildungen 12 bis 16 lassen sich die Ansätze für solche Prognose-Indikatoren ableiten und als Hypothese formulieren. Diese müssen dann wieder empirisch überprüft werden.

In diesem Abschnitt soll folgender Prognose-Indikator überprüft werden: Als Indikator dient ein Vergleich der beobachteten Symptomschwere nach 4 bzw. 8 Wochen mit der oberen Grenze der Bandbreite erfolgreicher Verläufe. Es wird registriert, ob sich der beobachtete Wert *innerhalb* oder *außerhalb* der Bandbreite erfolgreicher Verläufe befindet. Die unterschiedliche Verlaufsform von Respondern und Non-Respondern wird dabei berücksichtigt. Die entsprechende Hypothese lautet:

Patientinnen, deren Symptomschwere nach 4 bzw. 8 Wochen außerhalb der Bandbreite des positiven Referenzverlaufes liegt, haben eine geringere Chance auf Therapieerfolg am Ende der Behandlung, als Patientinnen, deren Wert innerhalb der Toleranzgrenze liegt.

Ein möglicher Vorteil eines solchen Indikators liegt darin, daß er (im Gegensatz zu der Variable Response/Non-Response) sowohl die unterschiedlichen Verlaufsgestalten von Respondern und Non-Respondern als auch deren Variabilität berücksichtigt. Mit dem Prognose-Indikator wird also eine weitere Bedingung eingeführt, von der wir annehmen, daß sie mit dem Therapieerfolg zusammenhängt. Nur wenn sich dieser Zusammenhang als gegeben und prognostisch relevant herausstellt, lohnt es sich, die Bandbreiten der Referenzverläufe zur Abschätzung des Therapieerfolgs heranzuziehen und in das QM-System zu integrieren.

13.2 Vorgehen und Ergebnisse der Analysen mit dem Prognose-Indikator

13.2.1 Vorgehen

Ähnlich wie Abschnitt 7.2.2 (S. 55ff.) wird mit Hilfe von logistischer Regression der Therapieerfolg am Ende der Behandlung modelliert und die dort beschriebenen Modelle

um die Variable des Prognose-Indikators erweitert. Ist die beobachtete Symptomschwere nach 4 Wochen außerhalb der Bandbreite, nimmt der Indikator den Wert 1 an, liegt sie innerhalb, bekommt er den Wert 0. Das Interesse liegt nun auf der *Vorhersage* der Wahrscheinlichkeit für Therapieerfolg, daher werden nun bei der Auswahl der Variablen im Modell diejenigen weggelassen, die bei Beginn der Behandlung nicht bekannt sind (Behandlungsdauer und -intensität). Entsprechend wird z.B. das Modell aus Tabelle 14 (S. 61) um die Variable des Prognose-Indikators erweitert. Um das Zusammenspiel der Variablen genauer zu untersuchen, wurden zusätzlich die Interaktionen des Indikators mit den anderen Variablen im Modell getestet und nur dann beibehalten, wenn sie einen substantiellen Beitrag zum Modell leisteten.

Der folgende Abschnitt enthält die Ergebnisse dreier Modelle, die die Bedeutsamkeit der Bandbreite zu zwei Zeitpunkten (nach 4 und 8 Wochen) untersuchen. Für das 4-Wochen Intervall wird dabei wieder die Stichprobe der Patientinnen mit mindestens 6 Wochen Behandlungsdauer herangezogen; für die Stichprobe mit einer Behandlungsdauer von mindestens 10 Wochen werden beide Intervalle untersucht.

13.2.2 Ergebnisse

Tabelle 24 gibt die Ergebnisse der Bulimie-Patientinnen mit mindestens 6 Wochen Behandlungsdauer wieder. Bei Hinzunahme des Prognose-Indikators zu dem Modell aus

Tabelle 24: *Vorhersage Therapieerfolg zuzüglich Prognose-Indikator*

Variable	Koeffizient	se	Wald- χ^2	p	Odds-Ratio
Intercept	1.35	0.45	9.01	0.0027	–
anfängliche Symptomschwere	-1.23	0.22	30.13	0.0001	–*
Symptomredukt. nach 4 Wochen	1.54	0.23	44.42	0.0001	4.7 (3.0,7.3)
Prognose-Indikator	-4.50	1.40	10.40	0.0013	–*
Symptomschwere X Indikator	1.50	0.58	6.64	0.0100	–*

$R^2 = 0.23$, $-2 \text{ Log Likelihood} = 558.6$, Behandlungsdauer: mindestens 6 Wochen

* wegen der Interaktion ist hier der Odds-Ratio nicht sinnvoll

Tabelle 14 verlor die Variable ‘Kliniktyp’ ihre Bedeutsamkeit ($\chi^2 = 5.12$, $df = 3$) und wurde daher aus dem Modell entfernt. Die Koeffizienten der anfänglichen Symptomschwere und des Therapie-Response bleiben im Vergleich zu Tabelle 14 im wesentlichen unverändert. Der Prognose-Indikator stellt sich sowohl als Haupteffekt, als auch in Interaktion mit der anfänglichen Symptomschwere als bedeutsam heraus.

Um die Art und Größenordnung dieses Effektes näher zu beschreiben, sind die Zusammenhänge der Variablen in Abbildung 17 veranschaulicht. Es wird deutlich, daß

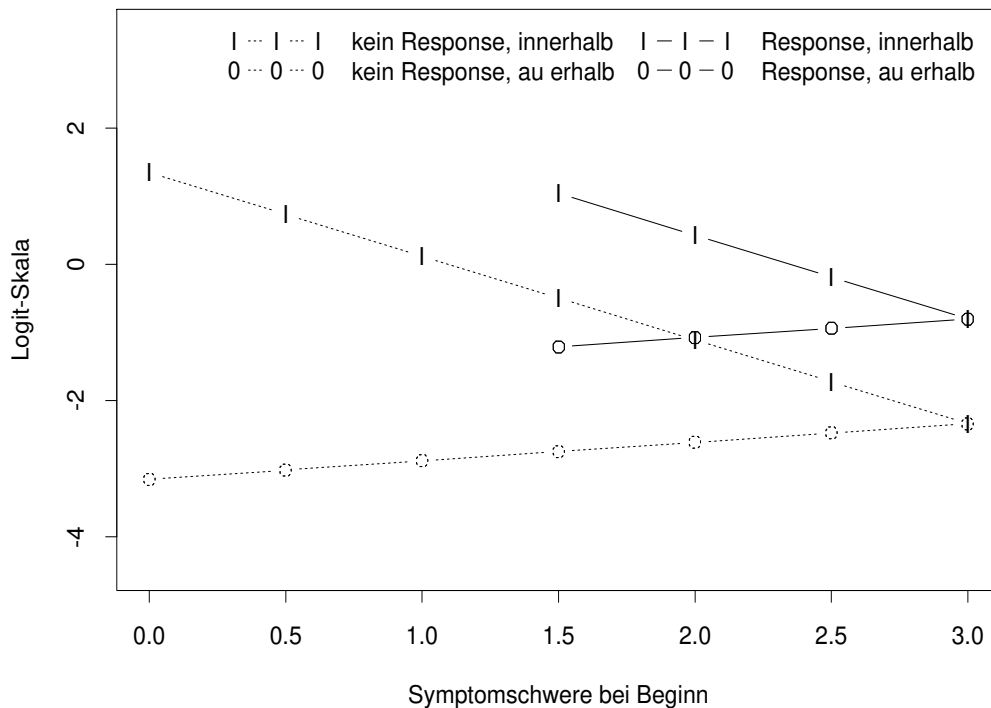


Abbildung 17: Relation zwischen anfänglicher Symptomschwere, Therapie-Response und Prognose-Indikator in Bezug auf die Erfolgswahrscheinlichkeit; mehr als 6 Wochen Behandlung, (s. Tab. 24)

Patientinnen, deren Symptomschwere nach 4 Wochen innerhalb der Bandbreite erfolgreicher Verläufe liegt, auch eine höhere Wahrscheinlichkeit für Therapieerfolg haben. Die Größe dieses Unterschieds hängt jedoch von der Symptomschwere bei Beginn der Behandlung ab: je niedriger die anfängliche Symptomschwere, desto größer der Unterschied. Starten Patientinnen mit voller Symptomatik, so besteht kein Unterschied zwischen beiden Gruppen. Starten sie bei 2, so beträgt der Unterschied 1.5 Logits, was einem Odds-Ratio von 4.5 entspricht¹⁴; Patientinnen mit einer anfänglichen Symptomschwere von 2, deren Symptomschwere nach 4 Wochen innerhalb der Bandbreite erfolgreicher Verläufe liegt, hätten also eine mehr als 4 mal so hohe Wahrscheinlichkeit auf Therapieerfolg als Patientinnen, deren Symptomwert nach 4 Wochen außerhalb der Bandbreite erfolgreicher Verläufe liegt. Dies gilt für Responder und Non-Responder in gleichem Maße; allerdings haben die Non-Responder insgesamt eine niedrigere Erfolgswahrscheinlichkeit. Die Kombination von Therapie-Response und Prognose-Indikator macht deutlich, welche Gruppen die höchste und welche die niedrigste Erfolgswahrscheinlichkeit haben: bei gleicher anfänglicher Symptomschwere – nämlich 2 – haben Patientinnen mit Therapie-Response, deren Symptomwert nach 4 Wochen innerhalb

¹⁴Der Odds-Ratio berechnet sich aus der exponentierten Differenz der Logits: $\exp(\text{Differenz der Logits})$

der Bandbreite erfolgreicher Verläufe liegt, die besten Erfolgschancen; die ungünstigsten Chancen haben Patientinnen ohne Therapie-Response, deren Symptomwerte nach 8 Wochen außerhalb der Bandbreite erfolgreicher Verläufe befindet. Der Unterschied zwischen diesen beiden Gruppen beträgt 3 Logits (Odds-Ratio = 20), was einer 20-fach höheren Erfolgswahrscheinlichkeit entspricht.

Da nur Patientinnen mit einer anfänglichen Symptomschwere von mindestens 1.5 zu den Respondern gehören können, reichen die Linien der Responder nur von 1.5 bis 3. Rund 80% der Patientinnen haben eine anfängliche Symptomschwere von 1.5 und mehr, daher sind die Unterschiede im Symptomschwerebereich 0 bis <1.5 mit größerer Unsicherheit behaftet und sollten nicht als exakte Schätzung interpretiert werden.

Die beiden Modelle der Patientinnen mit einer Behandlungsdauer von mindestens 10 Wochen kommen zu sehr ähnlichen Resultaten: Neben der anfänglichen Symptom-

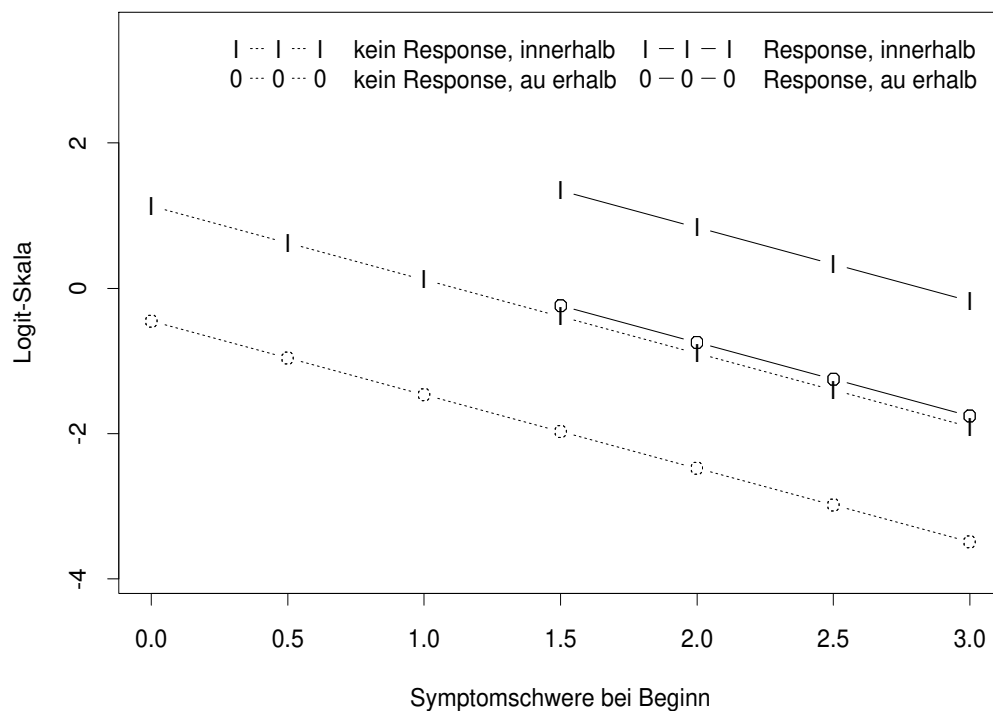


Abbildung 18: Relation zwischen anfänglicher Symptomschwere, Therapie-Response und Prognose-Indikator in Bezug auf die Erfolgswahrscheinlichkeit; mehr als 10 Wochen Behandlung

schwere, Therapie-Response und dem Prognose-Indikator, kommt keiner weiteren Variable noch Bedeutsamkeit zu. Der Unterschied der Erfolgswahrscheinlichkeit zwischen Patientinnen, deren Symptomwerte sich nach 4 Wochen innerhalb bzw. außerhalb der Bandbreite erfolgreicher Verläufe befinden, liegt in der gleichen Größenordnung: die Odds-Ratios liegen zwischen 4 und 5.

Für den Prognose-Indikator nach 8 Wochen fällt die Interaktion zwischen der anfänglichen Symptomschwere und dem Indikator weg, d.h. hier ist der Unterschied zwischen Patientinnen mit Symptomwerten innerhalb bzw. außerhalb der Bandbreite unabhängig von der anfänglichen Symptomschwere. (siehe Abbildung 18). Es fällt auf, daß Responder, die nach 8 Wochen außerhalb der Bandbreite erfolgreicher Verläufe liegen, eine in etwa gleich hohe Erfolgswahrscheinlichkeit haben wie Non-Responder, die innerhalb dieses Bereiches liegen. Es läßt sich erkennen, daß mit dem Prognose-Indikator eine Bedingung eingeführt wird, die eine deutliche Unterscheidung von Gruppen mit besonders hohen bzw. besonders niedrigen Erfolgchancen ermöglicht. Der Unterschied zwischen Non-Respondern (außerhalb der Bandbreite) und Respondern (innerhalb der Bandbreite) beträgt 3.4 Logits, das entspricht einem Odds-Ratio von 27.

Die Güte dieser Modelle läßt sich wiederum über die Anzahl korrekter Klassifikationen und über den Anteil erklärter Varianz abschätzen. Tabelle 25 zeigt in welchem Umfang sich die Vorhersage des Therapieerfolgs verändert, wenn die bedeutsamen Variablen sukzessive dem Modell hinzugefügt werden. Zusätzlich werden hier Sensitivität und Spezifität berichtet, um die Güte der Klassifikationen besser zu beurteilen. Die Sensitivität gibt von den tatsächlich erfolgreichen Patientinnen den Anteil der durch das Modell korrekt als 'Erfolg' klassifizierten Patientinnen an. Die Spezifität gibt von den tatsächlich nicht erfolgreichen Patientinnen den Anteil der korrekt als Mißerfolg klassifizierten an. 'Falsch positiv' und 'falsch negativ' geben entsprechend den Anteil der jeweils falsch klassifizierten Patientinnen wieder.

Tabelle 25: *Klassifikationsraten der Vorhersagemodelle*

Variable	%		Spezif.	falsch	falsch	R ²
	korrekt	Sensit.		pos.	neg.	
<i>Behandlungsdauer: mindestens 6 Wochen</i>						
anfängliche Symptomschwere	70.0	0	100	0	30	2.3
Symptomredukt. (4 Wochen)	73.0	38.3	87.8	42.6	23.1	17.2
Prognose-Indikator (4 Wochen)	74.6	43.1	88.0	39.5	21.6	23.0
<i>Behandlungsdauer: mindestens 10 Wochen</i>						
anfängliche Symptomschwere	71.4	0	100	0	28.6	3.6
Symptomredukt. (4 Wochen)	75.2	44.8	87.4	41.3	20.2	16.2
Prognose-Indikator (4 Wochen)	76.4	46.2	88.5	38.5	19.6	22.4
Prognose-Indikator (8 Wochen)*	77.1	39.8	91.9	33.9	20.6	27.8

* anstatt Prognose-Indikator nach 4 Wochen

In beiden Modellen nimmt die Anzahl an korrekten Klassifikationen und der Anteil

an aufgeklärter Varianz mit der Hinzunahme weiterer Variablen zu. Auffällig ist, daß die anfängliche Symptomschwere alleine im Vergleich zur Basisrate (30% erfolgreich, 70% nicht erfolgreich) keinen Beitrag zur Verbesserung der Klassifikation bringt. Erst die Hinzunahme von Therapie-Response verbessert die Rate; der Prognose-Indikator erreicht nochmals eine geringe Verbesserung der Vorhersage. Den Modellen mit allen drei Variablen gelingt eine gute Klassifikation der Patientinnen, die nicht erfolgreich abschneiden: die Spezifität liegt zwischen 88% und 91%. Die erfolgreichen Patientinnen werden weniger gut klassifiziert (die Sensitivität liegt zwischen 40% und 46%).

Der Vergleich zwischen dem Prognose-Indikator nach 4 Wochen oder nach 8 Wochen zeigt, daß auch nach 8 Wochen ein geringer Zugewinn an Information besteht: Die Klassifikationsrate ist nur geringfügig höher. Allerdings nimmt der Anteil an fälschlicherweise als ‘Erfolg’ klassifizierten Patientinnen ab, wie auch der Anteil an korrekt als ‘Erfolg’ eingestuft, während die Spezifität noch weiter zunimmt. Dieses Modell erfaßt folglich die nicht erfolgreichen Patientinnen etwas sicherer, als das Modell mit dem Prognose-Indikator nach 4 Wochen.

Diese Ergebnisse belegen einen deutlichen Zusammenhang zwischen dem Prognose-Indikator und dem Therapieerfolg. Patientinnen, deren Symptomschwere nach 4 Wochen innerhalb der Bandbreite der erfolgreichen Verläufe liegt, haben eine bessere Aussicht auf Therapieerfolg als Patientinnen, die außerhalb dieser Bandbreite liegen. Dieser Effekt gilt sowohl für Responder als auch für Non-Responder, wobei Non-Responder allgemein eine niedrigere Erfolgswahrscheinlichkeit aufweisen. Die Informationen aus dem Therapieverlauf (Therapie-Response und Prognose-Indikator) zeigen einen höheren Zusammenhang zum Therapieergebnis, als die Informationen vom Beginn der Behandlung, was sich darin niederschlägt, daß die Variablen vom Therapiebeginn ihren Einfluß verlieren, wenn die Variablen aus dem Therapieverlauf hinzugenommen werden – lediglich die anfängliche Symptomschwere behält ihren Einfluß.

Vergleicht man den Symptomwert nach 8 Wochen Behandlung mit der Bandbreite der erfolgreichen Verläufe und bildet den Prognose-Indikator, wird die Gesamtanzahl an korrekten Vorhersagen nicht wesentlich verändert – allerdings erhöht sich die Anzahl korrekt als Mißerfolg klassifizierter Patientinnen, während die Treffsicherheit für positive Verläufe sinkt. Unter dem Gesichtspunkt der Prognose ist aber durchaus einem Modell den Vorzug zu geben, welches dazu neigt, eher viele ‘Warnsignale’ zu produzieren, um zu vermeiden, daß eine größere Anzahl von Verläufen mit negativer Prognose übersehen wird. Von daher kann der Prognose-Indikator sowohl nach 4 als auch nach 8 Wochen eingesetzt werden. Nach 8 Wochen können die Patientinnen mit negativer Prognose etwas sicherer klassifiziert werden als nach 4 Wochen.

Diese Befunde sprechen für die Nützlichkeit der Bandbreiten und eines daraus abgeleiteten Indikators zur Prognose des weiteren Therapieverlaufs. Das legt nahe, solche Prognosen in einem verlaufsorientierten QM-System, das auf entsprechende Verläufe

zurückgreifen kann, zu integrieren, da sowohl Therapie-Response als auch der Vergleich der aktuellen Symptomatik mit den Referenzverläufen ('Prognose-Indikator') für den weiteren Verlauf bedeutsam sind. Diese Informationen können bei anstehenden Behandlungsentscheidungen den Prozeß der Entscheidungsfindung bereichern.

Gleichzeitig wird auch deutlich, daß sich die vorhersagten Klassifikationen über Erfolg und Mißerfolg durch solch ein Modell zwar verbessern lassen, aber nicht in einem solchen Umfang, daß eine nahezu fehlerlose Vorhersage möglich wäre. Die vorliegende Untersuchung demonstriert die *Möglichkeit* zur Verbesserung der Prognosen, aber aufgrund des Designs der Studie und der eingesetzten Instrumente sind die Ergebnisse nicht sofort in die Praxis zu übertragen. Z. B. bedürfen diese Ergebnisse noch einer Kreuzvalidierung an einer neuen Stichprobe, um die Genauigkeit des Modells an neuen Patientinnen zu überprüfen, die nicht an der Modellbildung beteiligt waren (siehe Abschnitt 14.1 für weitere Punkte).

13.3 Anwendungsbeispiele

Zur Verdeutlichung der Anwendung der Bandbreite als Prognose-Indikator sei nochmals auf den Symptomverlauf der Patientin aus Abbildung 14 (S.101) zurückgegriffen: Die Patientin startete mit fast extremer Symptomatik. Da sie nach 4 Wochen auf die Therapie angesprochen hat, sind ihre Chancen auf Therapieerfolg ca. 4-5 mal so hoch wie bei Patientinnen gleicher Symptomschwere, die keinen Therapie-Response zeigen. Sie befindet sich nach 4 Wochen innerhalb der Bandbreite erfolgreicher Verläufe, was jedoch aufgrund der starken Anfangssymptomatik keine wesentliche Verbesserung der Erfolgchancen bedeutet (vgl. Abbildung 17). Die aufgrund der Anfangsschwere, des Therapie-Response und des Prognose-Indikators vorhergesagte Erfolgswahrscheinlichkeit beträgt 0.37 und liegt damit eher auf der Seite des Mißerfolgs, obwohl die Patientin zu den Respondern gehört und sich bisher positiv entwickelt hat. Auch hier bestünde die Aufgabe des therapeutischen Teams darin, diese Angaben mit den klinischen Angaben und Eindrücken zu verbinden. So könnte bei dieser Patientin die enorme Verbesserung 'zu schnell' eingetreten sein, so daß im Team doch eine erhöhte Rückfallgefahr – ggf. untermauert durch Beobachtungen aus dem Klinikalltag – gesehen wird. Diese Interpretation würde durch die eher negative Prognose gestützt. So könnte sich das Team darauf verständigen, die Behandlung so weiterlaufen zu lassen wie bisher, um festzustellen, ob die Verbesserung stabil bleibt.

Nach weiteren 4 Wochen war der Zustand der Patientin wieder so weit verschlechtert, daß die Symptomschwere außerhalb der Bandbreite erfolgreicher Verläufe lag. Dies ist zu diesem Zeitpunkt ein negativer Prognose-Indikator. Die Erfolgswahrscheinlichkeit wird nun auf 0.18 geschätzt, also niedriger als 4 Wochen zuvor. Angesichts des Rückfalls und einer eher ungünstigen Prognose, stellt sich die Frage, ob das Behand-

lungsprogramm einer Veränderung bedarf. Vielleicht stellte das Programm zu hohe Anforderungen an die Patientin, die sie nicht durchhalten konnte. Solche und viele andere mögliche Gründe können diskutiert und unter der empirisch gestützten Vorgabe einer eher ungünstigen Prognose in Behandlungsentscheidungen münden, die eine für diese Patientin wirksamere Behandlungsstrategie einleiten.

Einen ähnlich steilen Verlauf nahm die Patientin aus Abbildung 12 (S. 97), der sich nach 8 Wochen jedoch deutlich anders entwickelte, als bei der gerade eben erwähnten Patientin: Sie begann auf etwas weniger hohem Niveau, zeigte nach 4 Wochen Therapie-Response und lag innerhalb der Bandbreite erfolgreicher Verläufe. Patientinnen mit ihrem Anfangsschweregrad haben eine etwas bessere Prognose, wenn sie sich nach 4 Wochen im Bereich der erfolgreichen Verläufe befinden (vgl. Abbildung 17) als Patientinnen, die sich außerhalb dieses Bereiches bewegen. Die geschätzte Erfolgswahrscheinlichkeit lag zu diesem Zeitpunkt bei 0.53 und würde hier schon eher auf einen positiven Therapieausgang deuten. Wieder wäre das behandelnde Team aufgefordert, diese Information mit den klinischen Eindrücken und Erfahrungen zu verbinden und zu bewerten.

Nach 8 Wochen befindet sich der Symptomschwere-Score immer noch im Bereich der erfolgreichen Verläufe. Damit gehört diese Patientin zu der Gruppe mit der besten Prognose: sie hat auf die Behandlung angesprochen und befindet sich im Bereich der erfolgreichen Verläufe; die geschätzte Erfolgswahrscheinlichkeit beträgt nun 0.64. Man könnte sagen, das Bild verdichtet sich, daß diese Patientin die Behandlung erfolgreich abschließen wird. Es ist allerdings auch bekannt, daß dieses Modell eine relativ niedrige Sensitivität für Therapieerfolg hat (39.8) – mehr als die Hälfte der tatsächlich erfolgreichen Patientinnen wurden im Modell nicht als solche erkannt. Dies verdeutlicht nochmals, daß das Modell *alleine* keine hinreichenden Vorhersagen zuläßt. Die ergänzende Einschätzung und Bewertung anhand der klinischen Erfahrungen bleibt notwendiger Bestandteil einer Entscheidungsfindung für die weitere Behandlung.

Zusammenfassung:

Die Bandbreiten der im vorigen Abschnitt konstruierten Referenzverläufe, haben prognostische Relevanz für den weiteren Therapieverlauf. Patientinnen deren Symptomwerte schon in den ersten Wochen innerhalb der Bandbreite erfolgreicher Verläufe liegen, haben eine höhere Chance auf Therapieerfolg, als Patientinnen, deren Scores sich ausserhalb dieses Bereichs bewegen. Dies gilt insbesondere für Patientinnen, die mit einer mittleren bis ausgeprägten Symptomschwere starten. Patientinnen mit extrem starken Symptomen haben insgesamt eine eher schlechtere Prognose, daher ist bei ihnen der Vorteil des Prognose-Indikators weniger stark ausgeprägt. Konstruiert man den

Prognose-Indikator nach 8 Wochen Behandlung so hält sich der Vorteil der innerhalb der Bandbreite liegenden Patientinnen und ist überdies nicht mehr vom anfänglichen Schweregrad abhängig. Die Ergebnisse dieses Abschnittes sprechen für die Brauchbarkeit der Bandbreite von Referenzverläufen bei der Prognose des weiteren Verlaufes und damit auch für ihren Einsatz in einem verlaufsorientierten QM-System. Allerdings sind aufgrund der Datenerhebung für andere Studienzwecke (vgl. Kächele 1999) Einschränkungen zu beachten.

14 Diskussion und Ausblick

14.1 Diskussion

14.1.1 Der Zusammenhang zwischen Therapie-Response und Therapieerfolg

Ein Hauptziel der Arbeit bestand darin zu untersuchen, ob die Vorhersage des Therapieerfolgs unter Hinzunahme bedeutsamer Ereignisse aus dem Therapieverlauf verbessert werden kann. Das hier vorgeschlagene Konzept des Therapie-Response fordert als wichtigstes Kriterium die prognostische Relevanz dieses Ereignisses für den weiteren Therapieverlauf. Je nach dem, welche Operationalisierung nun für Response gewählt wurde, fiel auch der Zusammenhang mit dem Therapieergebnis unterschiedlich aus.

Die Operationalisierung als ‘verbessertes Wohlbefinden’ zeigte dabei den niedrigsten Zusammenhang. Der Grund dafür mag zum einen in der Erfassung dieses Konstruktes mit nur einer Frage bezüglich der Veränderungseinschätzung der Patientin liegen. Das Wohlbefinden wird nicht darin direkt erwähnt. Daher besteht die Gefahr, daß das Konstrukt mit dieser Frage nicht adäquat erfaßt wird. Ein anderer Grund könnte in einer hohen Fluktuation des Wohlbefindens bestehen: die Frage erfaßt die Einschätzung des Patienten zum jetzigen Augenblick und kann durch aktuelle Stimmungsschwankungen beeinflußt sein. Diese müssen aber nicht unbedingt dem ‘durchschnittlichen’ Wohlbefinden der vorangegangenen Woche entsprechen; aufgrund der hohen Schwankungsbreite mag der Zusammenhang nur gering ausfallen. Eine sorgfältigere Erfassung des Konstruktes könnte durchaus höhere Zusammenhänge bringen.

Bei den symptomorientierten Operationalisierungen war der Zusammenhang zum Therapieergebnis deutlich höher. Das mag auf den ersten Blick nicht verwunderlich erscheinen, da mit der Reduktion der Symptome der Weg zum Therapieerfolg ja schon zum Teil beschritten ist; das Interesse liegt allerdings darauf, ob die erreichten Veränderungen darüberhinaus stabil bleiben. Beide Operationalisierungen, Verbesserung der Symptome und moderate Symptomatik, hängen mit der anfänglichen Symptomschwere zusammen, welche wiederum mit dem Therapieergebnis zusammenhängt. Daher muß

überprüft werden, ob Therapie-Response und anfänglicher Schweregrad nicht die gleiche Information beinhalten. Wenn dies – wie bei den Anorexien und Therapie-Response als moderate Symptomatik (siehe S. 66) – der Fall ist, kann mit Response kein zusätzlicher Informationsgewinn erzielt werden.

Schließlich ist darauf zu achten, daß das Kriterium für Therapieerfolg auch tatsächlich erreicht wird. In der Stichprobe der Anorexien wurde das Erfolgskriterium nur von gut 10% der Patientinnen erreicht, was die überwiegend starke Chronifizierung und den bekannten langfristigen Verlauf der Störung widerspiegelt. Hier ist zu überlegen, ob für Therapieerfolg eine niedrigere Schwelle als hier (nur noch Residualsymptomatik) angesetzt werden sollte. Im Prinzip stellt dies jedoch die gleiche Fragestellung wie die Definition von Therapie-Response dar: welches Kriterium ist bedeutsam, um als partieller Therapieerfolg gewertet zu werden? Auch hier ist neben der aktuellen Verbesserung des Zustandes der Patientin (z.B. in sehr schweren Fällen keine künstliche Ernährung mehr, oder keine Notwendigkeit eines stationären Aufenthaltes) die weitere Prognose von wichtiger Bedeutung.

Der nicht mehr vorhandene Zusammenhang zwischen Response und Therapieerfolg nach 2.5 Jahren bei Bulimie kann mit den sich rasch verändernden verhaltensnahen Symptomen (Eß- und Brechanfälle) erklärt werden: sie können – im Gegensatz zu Symptomen der Anorexie – sehr schnell ‘aufgegeben’ werden oder wieder auftreten, so daß es hier wahrscheinlich ist, zum Katamnesezeitpunkt die Patientinnen in den unterschiedlichsten Phasen des Verlaufes anzutreffen. Daher ist es gerade bei Bulimie sinnvoll, vermehrt Erfolgskriterien einzusetzen, die die Dauer und Anzahl vorhergehender Rückfälle mit einbezieht (vgl. Frank et al. 1991).

14.1.2 Zeitpunkt des Therapie-Response:

Aufgrund des Designs der Studie konnte der Symptomstatus der Patientinnen nur in einem Abstand von jeweils 4 Wochen erhoben werden. Da die meisten Patientinnen nur ca. 11 bis 12 Wochen in Behandlung waren, standen nur die Erhebungen nach 4 und (bei einer reduzierteren Stichprobe) nach 8 Wochen zur Verfügung, um Response zu definieren. Der Vergleich dieser beiden Zeitpunkte brachte keine deutlichen Unterschiede zu Tage. Ein stärkerer Zusammenhang zwischen Response nach 8 Wochen und Therapieerfolg liegt intuitiv nahe, dies ist jedoch nur dann gegeben, wenn durch den Response nach 8 Wochen der Anteil an erfolgreichen Patientinnen bei den Respondern erhöht wird – der Anteil an erfolgreichen Patientinnen bei Respondern nach 4 bzw. 8 Wochen ist hier jedoch sehr ähnlich. Lediglich bei der Operationalisierung als ‘moderate Symptome’ ist der Zusammenhang zum Therapieerfolg etwas höher.

Ein dichteres Zeitraster bei der Untersuchung der Verläufe wäre jedoch notwendig, um den ‘günstigsten’ Zeitpunkt zur Erfassung von Therapie-Response genauer zu ermitteln.

Eine wöchentliche Erhebung hätte neben einer feineren Bestimmung des Zeitpunktes von Therapie-Response den Vorteil, auch kurzfristigere Veränderungen zu erfassen und diese ggf. von Therapie-Response zu unterscheiden: die einmalige Erhebung des Zustandes nach 4 Wochen sagt nichts über die Stabilität der eingetretenen Veränderung aus; sie kann eine kurzfristige Verbesserung mit gleich anschließender Verschlechterung oder auch Teil einer schon seit 3 Wochen anhaltenden Verbesserung sein. Künftige Untersuchungen über Therapie-Response sollten daher Erhebungen in einem deutlich dichteren Zeitraster als 4 Wochen vornehmen.

14.1.3 Response und weiterer Verlauf der Symptomatik:

Die 4 beschriebenen Verlaufstypen belegen den unterschiedlichen Verlauf von Respondern und Non-Respondern. Da bei beiden Gruppen, sowohl erfolgreiche und nicht erfolgreiche Verläufe anzutreffen sind, lassen sich für die 4 möglichen Kombinationen Referenzverläufe konstruieren. In der Anwendung dieser Referenzen für den Vergleich mit aktuell laufenden Therapien, ist der Vergleich mit den erfolgreichen Verläufen am sinnvollsten, um zu entscheiden, inwiefern die Symptomatik im Bereich dessen liegt, was bei erfolgreichen Therapien üblicherweise zu beobachten ist.

Diese 4 unterschiedlichen Verlaufsformen stellen eine relativ grobe Vereinfachung dar, die sich nicht zuletzt durch das grobe Zeitraster der Erhebungen ergibt. Es sind eine ganze Reihe weiterer Verlaufstypen denkbar, die sich jedoch in dieser Untersuchung nicht untersuchen ließen. So könnte Therapie-Response bei einigen Patienten erst nach einer bestimmten Periode gleichbleibender Symptomatik eintreten. Dies hätte einen durch einen deutlichen 'Bruch' gekennzeichneten Verlauf zur Folge. Eine andere Variante besteht in der Reduktion der Anzahl von Rückfällen. Der Symptomverlauf wäre durch ein ständiges Auf und Ab charakterisiert, dessen Frequenz sich im Laufe der Zeit reduziert. Die Modellierung solcher Verlaufsgestalten erfordert jedoch eine größere Anzahl von Messungen in engeren Intervallen, um bestimmte Abschnitte identifizieren zu können.

14.1.4 Zur Methodik der Konstruktion von Referenzverläufen

Die Abbildungen in Abschnitt 12.3 verdeutlichen, daß die Aussagekraft der Referenzverläufe nicht zuletzt von ihrer Bandbreite abhängt. Je weniger sich erfolgreiche und nicht erfolgreiche Verläufe überlappen, desto deutlicher lassen sie sich unterscheiden. Nach Gleichung 2 (S. 90) hängt die Bandbreite sowohl von den Schwankungen der einzelnen Messungen um den geschätzten Verlauf als auch von der Variabilität der einzelnen Verläufe ab. Ein Instrument mit großem Meßfehler (d.h. stark schwankenden Messungen) wird daher zu einer großen Bandbreite der Referenzverläufe führen, genauso wie eine ungeeignete Auswahl des mathematischen Verlaufsmodells (vgl. Tabelle

21, S. 92).

Leider sind von dem hier benutzten Instrument keine Testgütekriterien vorhanden. Die geringe Anzahl der Fragen und ihre relativ grobe Skalierung lassen jedoch eine vergleichsweise niedrige Reliabilität vermuten. Das 4-wöchige Erhebungszeitraster trägt, wie bereits erwähnt, zusätzlich zu einer nur groben Erhebung des Verlaufs bei. Daß unter diesen suboptimalen Bedingungen dennoch ein Ergebnis erzielt werden konnte, das zu einer Verbesserung der Vorhersage des Therapieerfolges führt, spricht zum einen für das Konzept des Therapie-Response und zum anderen auch dafür, mit reliablen Instrumenten und einem feineren Erhebungszeitraster noch bessere Ergebnisse zu erzielen.

Um die gefundenen Ergebnisse noch genauer auf ihre prädiktive Validität hin zu überprüfen, ist eine Kreuzvalidierung mit einer *neuen* Stichprobe nötig. So ist es denkbar, daß die Ergebnisse zwar auf die vorliegende Stichprobe zutreffen, aber sich bei einer anderen Stichprobe nicht bewähren. Dieser wichtige Punkt konnte in dieser Arbeit trotz großer Stichprobe nicht mehr angegangen werden: dazu müßte aus allen 4 Verlaufsgruppen jeweils ein Drittel als 'Test-Stichprobe' zurückbehalten werden, damit erfolgreiche und nicht-erfolgreiche Responder und Non-Responder in gleichem Verhältnis wie in der Ausgangsstichprobe vertreten sind. Dadurch wird die Stichprobengröße zur Konstruktion der Referenzverläufe auf mindestens 50 (bei den erfolgreichen Non-Respondern) reduziert und bewegt sich damit nahe der Mindestgrenze von 30 bis 50 Patienten, die zur reliablen Schätzung der Verlaufparameter und insbesondere deren Variabilität benötigt werden (Snijders & Bosker 1999). Da das Prinzip eines verlaufsorientierten QM auch so hinreichend deutlich wird, wurde auf die Kreuzvalidierung verzichtet. Bevor dieses System (ggf. mit anderen Instrumenten und bei anderen Störungen) zur praktischen Anwendung kommt, ist eine Kreuzvalidierung allerdings unerlässlich. In einer Anwendung eines solchen Systems bei anderen Störungsbildern ist zusätzlich auf möglicherweise von den hier beobachteten Verlaufsgestalten abweichenden Mustern zu achten, die dann Referenzmodelle mit anderen mathematischen Modellen erforderlich machen.

14.1.5 Prognose-Indikatoren

Der untersuchte Prognose-Indikator (die Einteilung, ob sich der beobachtete Symptomwert innerhalb oder außerhalb der Bandbreite erfolgreicher Verläufe befindet) ist ein bedeutsamer Prädiktor des Therapieerfolgs, der stärker als die zu Beginn der Behandlung zur Verfügung stehende Information ist. Er bringt zusätzlich zu Therapie-Response mehr Varianzaufklärung und eine leicht verbesserte Klassifikation. In Verbindung mit der anfänglichen Symptomschwere und Therapie-Response lassen sich nach 4 Wochen Patientinnen mit einem besonders hohen Risiko für einen ungünstigen Therapieaus-

gang zuverlässig identifizieren. Dennoch stellt sich die Frage, ob sich die aufwendige Konstruktion der Grenzen auszahlt. Dies ist schwierig zu beurteilen. Angesichts der Größe der gefunden Effekte, unterscheidet der Indikator zwischen Gruppen mit deutlich unterschiedlicher Erfolgs-Wahrscheinlichkeit. Die Vorhersage der Klassifikation in erfolgreiche und nicht erfolgreiche Behandlungen wird aber nur geringfügig besser als die Base-Rate von 70%. Wie im vorigen Abschnitt bereits erwähnt, sind unter optimaleren Bedingungen durchaus größere Verbesserungen der Vorhersage zu erwarten. Eine hundertprozentig korrekte Vorhersage ist sicherlich ein zu hoch gestecktes Ziel, da im Verlauf einer Therapie plötzlich Ereignisse eintreten können, die eine bereits stattgefundenen Verbesserung wieder zunichte machen können oder umgekehrt, die Lösung eines Problems doch noch ermöglichen. Daher kann das Modell nur als Stütze dienen, das empirisch abgesicherte Fakten liefert – aber keine fehlerfreie Vorhersage. Der Vorteil liegt darin, die Größe des Fehler abschätzen zu können: die hier vorgestellten Analysen erlauben im Vergleich zu den erfolgreichen Patientinnen eine treffsicherere Vorhersage der nicht erfolgreichen Patientinnen .

Lutz et al. (1999) untersuchten in ihrem Modell in ähnlicher Weise die Bedeutung von Prognose-Indikatoren, die sich aus der Bandbreite der Verläufe ableiten. Sie fanden, daß Patienten, deren Symptomwert während der Behandlung zweimal die errechnete Grenze überschritt, ein dreimal höheres Risiko für Mißerfolg hatten als andere Patienten. Dies ist im Vergleich zu den hier berichteten zweistelligen Odds-Ratios vergleichsweise wenig. Die Gründe für den größeren Unterschied in dieser Arbeit sind vor allen Dingen in der Berücksichtigung des Therapie-Response und den unterschiedlichen Verlaufsformen zu suchen, im ausschließlichen Vergleich mit den positiv verlaufenden Therapien, in der Berücksichtigung von anderen Variablen bei der Modellierung der Erfolgchancen und schließlich in der homogeneren Stichprobe von ausschließlich bulimischen Patientinnen. All diese Punkte bewirken eine höhere Informationsausschöpfung in dem Sinne, daß bei der Konstruktion der Referenzverläufe und der Vorhersage des Therapieerfolgs mehr Faktoren berücksichtigt werden und damit eine höhere Genauigkeit erreicht wird.

Der Vergleich des Symptomstatus zu verschiedenen Zeitpunkten mit den Referenzverläufen eröffnet einen neuen Raum zur Konstruktion von weiteren Prognose-Indikatoren. Der Vergleich der Symptomschwere zu mehreren Zeipunkten oder im Verhältnis zu *beiden* Referenzen (erfolgreiche und nicht erfolgreiche Verläufe) erlaubt es, komplexere und potentiell vorhersagekräftigere Indikatoren zu entwickeln. So könnten beispielsweise die Erfolgchancen derjenigen Patientinnen besonders niedrig sein, wenn sie 2 mal (oder 2 mal direkt hintereinander) außerhalb der Bandbreite erfolgreicher Verläufe lagen. Eine andere Möglichkeit bestünde darin, analog zur ‘clinical significance’, aus der Bandbreite der erfolgreichen und nicht erfolgreichen Verläufen zu jedem Zeitpunkt einen Cutoff zu berechnen, der angibt, ob die aktuelle Symptomschwere eher zu den erfolgreichen oder nicht erfolgreichen Verläufen gerechnet werden kann. Solche

Indikatoren müssen, wie in dieser Arbeit demonstriert, auf ihren prognostischen Wert hin untersucht werden.

14.1.6 Ist ein verlaufsorientiertes QM durchführbar?

Ein verlaufsorientiertes QM-System erfordert sicherlich einen höheren organisatorischen Aufwand, um während der Therapie die benötigten Informationen zu sammeln, aufzubereiten und zurückzumelden. Dies mag zunächst als Nachteil erscheinen. Die Entwicklung kürzerer, aber dennoch in den wesentlichen Aspekten aussagekräftiger Instrumente (Lambert et al. 1996, Nisslmüller 1999, Kordy et al. 2000), der inzwischen alltäglich gewordene Einsatz moderner EDV, die verbesserte Prognose nach einer gewissen Zeit der Behandlung und die damit verbundene praktische Relevanz zur Unterstützung anstehender Behandlungsentscheidungen machen ein solches System sowohl realisierbar als auch attraktiv.

Die in Zukunft zu beantwortende 'Gretchenfrage' bleibt darüberhinaus, ob es gelingt, mit solchen oder ähnlichen Systemen die Qualität der Behandlung, d.h. die Behandlungsergebnisse bzw. die Zuteilung der Patienten zu den geeignetsten Therapieformen zu verbessern.

14.2 Zusammenfassung der wichtigsten Ergebnisse

An dieser Stelle werden nochmals die wichtigsten Ergebnisse der in Abschnitt 4 formulierten Ziele der Arbeit abschließend dargestellt. Die Ergebnisse aus Teil I. und II. lassen sich wie folgt zusammenfassen:

14.2.1 Konzeptualisierung und Operationalisierung von Therapie-Response

In Anlehnung an Frank et al. (1991) ist Therapie-Response eine sichtbare Verbesserung des (symptomatischen) Zustandes, der anzeigt, daß sich ein Patient auf dem Wege der Besserung befindet. Er ist gekennzeichnet durch eine mittlere Symptomschwere (keine volle Symptomatik, aber nicht notwendigerweise vollständige Besserung), ist der Behandlung zuzuschreiben und besitzt – als wichtigstes Kriterium – Aussagekraft bezüglich des Therapieergebnisses bzw. des zu erwartenden Verlaufs der Symptomatik. Im Allgemeinen Modell der Psychotherapie (Orlinsky & Howard 1987) kann Therapie-Response als Kumulation von 'Mikroergebnissen' verstanden werden, die positiv auf den weiteren Therapieverlauf zurückwirken.

Therapie-Response wurde in dieser Arbeit anhand der Kernsymptome von Anorexie und Bulimie als Verbesserung des Wohlbefindens, als das Erreichen eines bestimmten Symptomschweregrades und als Verringerung der Symptomatik um einen bestimmten

Betrag operationalisiert; er wurde nach 4 Wochen und – sofern noch weit genug vom Entlassungszeitpunkt entfernt – nach 8 Wochen registriert.

14.2.2 Prädiktive Validität von Therapie-Response

Zusammenhang zwischen Response und Therapieerfolg: Patientinnen mit Therapie-Response haben eine höhere Wahrscheinlichkeit für Therapieerfolg (Odds-Ratios zwischen 2 und 6). Dies gilt für alle Operationalisierungen mit Ausnahme der Verbesserung des Wohlbefindens bei Anorexie, wo dieser Effekt nicht in Erscheinung trat. Dieser Zusammenhang bleibt auch nach der Kontrolle relevanter Eingangsinformationen und Behandlungsparameter in vergleichbarer Größenordnung bestehen. Die beiden symptomorientierten Operationalisierungen von Therapie-Response zeigen im Vergleich zur Verbesserung des Wohlbefindens einen höheren Zusammenhang zum Therapieergebnis. Es macht keinen Unterschied, ob Therapieerfolg aus der Patienten- oder Therapeutenperspektive erhoben wurde.

Der Zusammenhang zwischen Response und Therapieerfolg nach 2.5 Jahren fiel sowohl bei Anorexie als auch bei Bulimie schwächer aus: bei Bulimie ließ sich bei keiner Operationalisierung ein Zusammenhang mehr feststellen. Bei Anorexie haben die Responder noch eine 2 bis 3 mal so hohe Wahrscheinlichkeit auf Therapieerfolg nach 2.5 Jahren, allerdings war der Unterschied statistisch nicht klar absicherbar. Ein verbessertes Wohlbefinden bei Anorexie dagegen war deutlich mit einer besseren Erfolgchance nach 2.5 Jahren verbunden.

Zeitpunkt der Erhebung von Therapie-Response: Der Zusammenhang zwischen Response und Ergebnis ist bei dieser Stichprobe weitgehend unabhängig vom Zeitpunkt der Erfassung: Response nach 8 Wochen führt zu keinem nennenswerten Zugewinn an Genauigkeit der Vorhersage des Therapieerfolges im Vergleich zu Response nach 4 Wochen. Ein leichter Vorteil des Response nach 8 Wochen ist nur in Relation zum Therapieergebnis aus der Therapeutenperspektive zu beobachten.

Verbesserung der Vorhersage des Therapieergebnisses durch Response: Bei Anorexie konnte die Vorhersage des Therapieerfolgs durch die Hinzunahme des Therapie-Response *nicht* verbessert werden. Dies lag an der unausgeglichene Erfolgsrate bei Anorexie (10% Erfolg, 90% Mißerfolg) und am Zusammenhang der Eingangsvariablen mit Therapie-Response (s.u.). Bei Bulimie konnte eine Verbesserung der Vorhersage des Therapieergebnisses durch Response erzielt werden, was sich in einer höheren Rate an richtigen Klassifikationen und einem erhöhten Ausmaß an aufgeklärter Varianz in den Modellen niederschlug.

14.2.3 Vorhersage von Therapie-Response

Therapie-Response hing bei den meisten Operationalisierungen nur sehr schwach mit den Eingangscharakteristika der Patientinnen zusammen. Dennoch ist Therapie-Response nicht immer unabhängig von den Eingangsvariablen der Patientinnen: je nach gewählter Operationalisierung von Therapie-Response besteht ein stärkerer oder schwächerer Zusammenhang. Response als ‘Verbesserung des Wohlbefindens’ war aus der Anfangsinformation der Patientinnen nicht vorherzusagen; anders bei den symptomorientierten Operationalisierungen: bei Anorexie hatte ein Großteil der Responder (definiert als ‘Erreichen einer moderaten Symptomatik’) schon bei Behandlungsbeginn nur relativ schwache Symptomatik. Response bringt hier keine neue Information und verbessert daher auch die Vorhersage des Therapieerfolgs nicht (s.o.).

Bei Bulimie besteht ein logischer Zusammenhang zwischen dem anfänglichen Schweregrad und Response: nur Patientinnen mit einer Mindest-Symptomschwere von 1.5 (definiert als Mindestmaß an ‘Verringerung der Symptomatik’) können überhaupt Response zeigen – in der vorliegenden Stichprobe war dies bei über 80% der Fall, so daß der Anteil an Non-Respondern nicht automatisch der Anteil der Patientinnen mit niedriger Anfangssymptomatik war und daher die Analysen nicht verzerrte.

14.2.4 Therapie-Response und Therapieverlauf

Am Beispiel der Bulimien konnte gezeigt werden, daß Responder und Non-Responder über den Behandlungszeitraum hinweg unterschiedliche Verlaufsformen haben. Gemäß der Definition von Response als Mindestmaß an Symptomverringering zeigen Responder eine starke Symptomreduktion innerhalb der ersten 4 Wochen, die dann in Abhängigkeit des Therapieergebnisses erhalten bleibt oder bei negativem Ausgang durch einen Anstieg der Symptome wieder verlorenght. Non-Responder zeigen in den ersten 4 Wochen eine deutlich geringere Symptomreduktion, die bei nicht erfolgreichen Patientinnen nach 1 bis 2 Monaten zum Stillstand kommt, während sie bei erfolgreichen Patientinnen länger anhält und so nach längerer Zeit als bei den Respondern zum Verlust der Symptome führt.

Teil III. der Arbeit demonstrierte die Anwendung von Therapie-Response in einem verlaufsorientierten QM. Die nächsten 3 Abschnitte fassen deren Ergebnisse zusammen:

14.2.5 Eigenschaften eines verlaufsorientierten QM-Systems

Vier Eigenschaften eines verlaufsorientierten QM wurden herausgearbeitet:

1. die Verwendung kurzer, veränderungssensitiver und doch valider Instrumente, die mehrfache Erhebungen ohne all zu großen Mehraufwand erlauben.

2. die Definition verlaufsorientierter Normen, die einen Vergleich des aktuellen Verlaufs mit der Bandbreite üblicherweise zu beobachtender, *positiv verlaufender* Therapien ermöglichen.
3. Die komplexere Verlaufsinformation erfordert eine computergestützte Erfassung, Verwaltung und grafische Aufbereitung der erhobenen Daten, die ein rasches Vorgehen und schnelle Rückmeldung gewährleisten, damit Patient und Therapeut noch im Verlauf der Therapie davon profitieren können.
4. Ein verlaufsorientiertes QM-System sollte Prognose-Indikatoren beinhalten, die eine genauere und abgesichertere Prognose des Therapieergebnisses unterstützt, um entsprechend darauf abgestimmte Behandlungsschritte einleiten zu können.

14.2.6 Konstruktion von Verlaufsstandards

Mit Hilfe des in Abschnitt 12.1.2 beschriebenen Verfahren der Hierarchischen Linearen Modelle lassen sich die durchschnittlichen Verläufe von Respondern und Non-Respondern konstruieren. Besonderer Vorteil besteht darin, daß sie

- zu jedem Zeitpunkt den Mittelwert *und* die Streuung der Symptomsschwere angeben
- die unterschiedlichen Verlaufsformen von Respondern und Non-Respondern berücksichtigen
- die individuellen Informationen vom Behandlungsbeginn integrieren

Da für den Verlauf einer Behandlung der Vergleich vom durchschnittlichen Verlauf der Normalpopulation und der einer Patientenpopulation (analog zum Vorgehen zur Berechnung der ‘clinical significance’) nicht möglich ist, bietet sich die Unterscheidung in ‘erfolgreiche’ und ‘nicht erfolgreiche’ Verläufe an. Die daraus resultierende Hypothese für ein verlaufsorientiertes QM ist, daß sich die Behandlung günstig entwickelt, so lange die Symptome sich im Spektrum erfolgreicher Verläufe befinden. Dieser Hypothese wurde mit Hilfe der Prognose-Indikatoren nachgegangen.

14.2.7 Verbesserte Abschätzung der Erfolgswahrscheinlichkeit durch Prognose-Indikatoren

Die hier vorgestellte Herangehensweise bietet 2 Prognose-Indikatoren an, die eine Verbesserung der Abschätzung des weiteren Verlaufes ermöglicht. Zum einen ist die Erfolgswahrscheinlichkeit bei Respondern höher als bei Non-Respondern; zum anderen

liefert der Vergleich von aktueller Symptomschwere und der Bandbreite weitere Information: liegt die Symptomschwere einer Patientin nach 4 oder 8 Wochen außerhalb des Bereiches der erfolgreichen Verläufe, so verschlechtern sich die Aussichten auf Therapieerfolg. Ein Prognosemodell, welches sowohl Anfangsinformationen als auch die Prognose-Indikatoren beinhaltet, verbessert die Vorhersage des Therapieerfolges im Vergleich zu einem Modell, das nur die Anfangsinformationen einbezieht.

14.3 Ausblick

In der Konsequenz des hier vorgestellten Ansatz lassen sich drei noch nicht ausgearbeitete Bereiche erkennen, die weiterer konzeptioneller und empirischer Arbeit bedürfen:

Zunächst verdeutlichen die gezeigten Beispielverläufe (vgl. Abschnitt 12.3, S. 96f.), daß es sich beim Vergleich des aktuellen Verlaufes mit den entsprechenden Referenzen um ein *Klassifikationsproblem* handelt. Gehört der gerade beobachtete Verlauf zur Klasse der erfolgreichen oder der nicht erfolgreichen? Die Grafiken zeigen, daß sich die Bandbreiten erfolgreicher und nicht erfolgreicher Verläufe je nach Ausgangslage und je nach Entwicklung in den ersten Wochen zum Teil deutlich überlappen. Der hier vorgestellte Prognose-Indikator ignoriert die nicht erfolgreichen Verläufe; eine Klassifikationsregel, die *beide* Referenzen berücksichtigt, könnte evtl. eine noch exaktere Prognose ermöglichen. So könnten etwa die Entfernungen eines beobachteten Symptomwertes vom Mittelwert beider Referenzen (ausgedrückt in Perzentilen der Verteilungen) in Verhältnis gesetzt werden, um diesen Index als weiteren Prognose-Indikator zu überprüfen.

Ein anderer Punkt besteht im Abschätzen des Zeitpunktes, zu dem eine Behandlungsentscheidung ansteht. Der Vergleich von aktuellen Therapieverläufen mit entsprechenden Referenzverläufen bietet Anregungen zur therapeutischen Reflektion der bisherigen Arbeit und eine empirisch abgesichertere Prognose über das Therapieergebnis. Zusammengenommen kann dies zwar eine Erleichterung für Entscheidungen zur adaptiven Indikation bedeuten, wünschenswert wäre aber an dieser Stelle nicht nur eine sicherere Prognose des Therapieergebnisses, sondern insbesondere eine Prognose über den zu erwartenden weiteren Weg der symptomatischen Entwicklung. Ließe sich der *weitere* symptomatische Verlauf mit hinreichender Genauigkeit vorhersagen, wären zeitlich bezogene Aussagen möglich, die eine Abschätzung des *Zeitpunktes* erlauben, zu dem eine Entscheidung ansteht oder bis zu dem man noch ohne Intervention warten kann. Die Extrapolation der symptomatischen Entwicklung könnte die Abschätzung des Zeitintervalles erleichtern, bis zu dem ein bestimmtes Ergebnis erreicht sein sollte. Ist dies dann nicht der Fall und die Prognose nach einiger Zeit negativ, so zeigte dies Handlungsbedarf an – wird das Ereignis schneller erreicht, so könnten schon früher im Therapieverlauf Interventionen zur Erhaltung des erreichten Fortschritts eingeleitet

werden.

Der hier vorgestellte Vergleich mit verschiedenen Verlaufstypen in Abhängigkeit vom Therapie-Response tut dies auf der Grundlage von Modellen, die den *gesamten* Verlauf modellieren. Eine interessante Variante der Verlaufsuntersuchung könnte darin liegen, nicht den gesamten, sondern nur den *anschließenden* Verlauf *nach* Therapie-Response zu modellieren. Die Verlaufsparameter dieses anschließenden Teilverlaufes könnten mit Hilfe der HLM-Methode zu den verschiedensten Faktoren in Beziehung gesetzt werden, um eine tatsächliche Vorhersage des weiteren Verlaufes auf der Grundlage der Ausgangsinformation und des bisherigen Verlaufes zu ermöglichen. Der Vorteil bestünde in einer verbesserten Vorhersage des *Verlaufes der Symptomatik*, während in dieser Arbeit die Vorhersage der *Erfolgswahrscheinlichkeit* mit Hilfe der Verlaufsmodelle verbessert wurde. Dazu ist jedoch eine Stichprobe von Verläufen nötig, die ein dichteres Erhebungsraster mit mehr Messungen als in der vorliegenden Stichprobe beinhaltet, da die ersten (beiden) Messungen ja nicht zum weiteren Verlauf nach Therapie-Response gehören und damit für die Modellierung wegfallen.

Mit dem vorigen Gedanken eng verknüpft ist der abschließende Punkt: Die genannten Modelle sind zwar von prognostischer Relevanz, aber sagen nichts darüber aus, welche *Art* von Maßnahmen die Prognosen implizieren. Was bedeutet es, wenn z.B. nach 12 Sitzungen noch keine sichtbare Verbesserung eingetreten sind? Ein Rückmeldung innerhalb des QM zeigte lediglich an, daß die Behandlung nicht so verläuft, wie erfolgreiche Behandlungen dies üblicherweise tun. Ein solches Signal erfordert eine Bewertung der zur Verfügung stehenden *Alternativen*. Erhöhen sich die Chancen auf Besserung wenn der Therapeut oder die Therapeutin gewechselt wird, oder wenn die Behandlung abgebrochen wird? Was soll bei Therapieabbruch die Alternative sein? Diese Frage erfordert Untersuchungen darüber, welche Faktoren zu einem unerwünschten Therapieverlauf führen können (z.B. Mohr 1995) und ob und wie nicht erfolgreich behandelte Patienten zu geeigneteren Behandlungen verwiesen werden können. In diesem Sinn bezieht sich der Ausdruck 'adaptive Indikation' nicht nur auf einen spezifischen Therapieabschnitt, sondern auch auf den gesamten Zeitraum, in dem eine Person behandlungsbedürftig ist. Dies entspricht dem sog 'stepped care approach', bei dem in Abhängigkeit der vorigen Behandlungsabschnitte und dessen Ergebnissen geeignete Nachfolgebehandlungen bestimmt werden.

Diese drei Punkte verdeutlichen die Bereiche einer verlaufsorientierten Qualitätsmanagement-Strategie, die noch weiterer Präzisierung und anschließender Validierung bedürfen: zum einen die Modellierung der Einzelverläufe selbst in adäquateren mathematischen Modellen, weiterhin die Entwicklung von Strategien und Indikatoren, welche die Prognose des weiteren Verlaufes und damit auch zeitlich bezogene Aussagen ermöglichen und schließlich die Evaluation der aufgrund des Qualitätsmanagements getroffenen Entscheidungen.

Literatur

- Angst, J. (1988). Risikofaktoren für den Verlauf affektiver Störungen. In Zerssen, D. v., Müller, D. v. (Hrsg.), *Affektive Störungen*, S. 99–110. Berlin: Springer-Verlag.
- Bastine, R. (1992). *Klinische Psychologie*. Kohlhammer: Stuttgart.
- Beutler, L. & Clarkin, J. (1990). *Systematic treatment selection: toward targeted therapeutic interventions*. New York: Brunner/Mazel, Inc.
- Breslin, F., Sobell, M., Sobell, L., Buchan, G. & Cunningham, J. (1997). Toward a stepped care approach to treating problem drinkers: the predictive validity of within-treatment variables and therapist prognostic ratings. *Addiction* 92(11), 1479–1489.
- Bryk, A. & Raudenbush, S. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin* 101(1), 147–158.
- Bryk, A. & Raudenbush, S. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage.
- Cleveland, W. (1993). *Visualizing Data*. Summit, New Jersey: Hobart Press.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist* 45, 1304–1312.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist* 49, 997–1003.
- Dawson, R., Lavori, P., Coryell, W., Endicott, J. & Keller, M. (1998). Maintenance strategies for unipolar depression: an observational study of levels of treatment and recurrence. *Journal of Affective Disorders* 49, 31–44.
- de Jong-Meyer, R., Hautzinger, M., Rudolf, G., Strauß, W., & Frick, U. (1996). Prädiktions- und Verlaufsanalysen bei kombinierter psychologischer und medikamentöser Therapie endogen depressiver Patienten. *Zeitschrift für klinische Psychologie* 25(2), 110–129.
- Dew, M., Reynolds, C., Houck, P., Hall, M., Buysse, D., Frank, E. & Kupfer, D. (1997). Temporal profiles of the course of depression during treatment. predictors of pathways toward recovery in the elderly. *Archives of General Psychiatry* 54(11), 1016–1024.
- Dilling, H., Mombour, W. & Schmidt, M. (1993). *Internationale Klassifikation psychischer Störungen. Kapitel V. (2. Aufl.)*. Bern: Verlag Hans Huber.
- Donabedian, A. (1982). *The criteria and standards of quality*. Ann Arbor: Health Administration Press.

- Elkin, I. (1994). The NIMH treatment of depression collaborative research program: where we began and where we are. In In Bergin, A.E. & Garfield, S. B. (Hrsg.), *Handbook of Psychotherapy and Behavior Change* (4. Aufl.), S. 114–142. New York: John Wiley & Sons, Inc.
- Fairburn, C. & Cooper, Z. (1993). The eating disorder examination (12th ed.). In Wilson, C. F. G. (Hrsg.), *Binge eating: Nature, assessment and treatment*, S. 317–360. New York: Guilford.
- Fiedler, P. (1997a). *Persönlichkeitsstörungen* (3. Aufl.). Weinheim: Psychologie Verlags Union.
- Fiedler, P. (1997b). 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* 26(2), 138–149.
- Frank, E., Prien, R., Jarrett, R., Keller, M., Kupfer, D., Lavori, P., Rush, A. & Weissman, M. (1991). Conceptualization and rationale for consensus definitions of terms in major depression disorder. *Archives of General Psychiatry* 48(Sept.), 851–855.
- Frank, J. (1973). *Persuasion and healing* (2. Aufl.). Baltimore: Johns Hopkins University Press.
- Franke, G. (1995). *Die Symptom Checkliste von Derogatis - Deutsche Version (SCL-90-R)*. Göttingen: Beltz Test.
- Gibbons, R., Hedeker, D., Elkin, I., Waterneaux, C., Kraemer, H., Greenhouse, J., Shea, T., Imber, S., Sotsky, S. & Watkins, J. (1993). Some conceptual and statistical issues in analysis of longitudinal psychiatric data. *Archives of General Psychiatry* 50, 739–750.
- Goldstein, H. (1995). *Multilevel Statistical Models* (2. Aufl.). New York: John Wiley & Sons Inc.
- Grawe, K. (1998). *Psychologische Therapie*. Göttingen: Hogrefe.
- Grawe, K. & Braun, U. (1994). Qualitätskontrolle in der Psychotherapiepraxis. *Zeitschrift für klinische Psychologie* 23(4), 242–267.
- Hebebrand, J., Hesecker, H., Himmelmann, W., Schäfer, H. & Remschmidt, H. (1994). Altersperzentilen für den Body-Mass-Index aus Daten der nationalen Verzehrstudie einschließlich einer Übersicht zu relevanten Einflußfaktoren. *Aktuelle Ernährungsmedizin* 19, 259–265.

- Heidenreich, T. & Hoyer, J. (1998). Stadien der Veränderung in der Psychotherapie. *Verhaltenstherapie und psychosoziale Praxis* 30(4), 381–402.
- Hosmer, D. & Lemeshow, S. (1989). *Applied logistic regression*. New York: Wiley & Sons.
- Howard, K., Brill, P., Lueger, R., O'Mahoney, M. & Grant, R. (1993). *Integra outpatient tracking assessment*. Philadelphia: Compass Information Services, Inc.
- Howard, K., Kopta, S., Krause, M. & Orlinsky, D. (1986). The dose-effect relationship in psychotherapy. *American Psychologist* 41(2), 159–164.
- Howard, K., Lueger, R., Maling, M. & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology* 61, 678–685.
- Howard, K., Moras, C., Brill, P., Martinovich, Z. & Lutz, W. (1996). The evaluation of psychotherapy: Efficacy, effectiveness, patient progress. *American Psychologist* 51.
- Illardi, S. & Craighead, W. (1994). The role of nonspecific factors in cognitive-behavior therapy for depression. *Clinical Psychology: Science and Practice* 1, 138–156.
- Jacobson, N. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Clinical and Consulting Psychology* 59(1), 12–19.
- Kächele, H. & Kordy, H. (1992). Qualitätssicherung: Erläuterungen zu einem Reiz- und Modewort. *Zeitschrift für psychosomatische Medizin* 38, 310–324.
- Kächele, H. & Kordy, H. (1996a). Ergebnisforschung in der Psychotherapie und Psychosomatik. In T. v. Uexküll (Hrsg.), *Psychosomatische Medizin* (5. Aufl.), S. 490–501. München Wien Baltimore: Urban & Schwarzenberg.
- Kächele, H. & Kordy, H. (1996b). Indikation als Entscheidungsprozess. In T. v. Uexküll (Hrsg.), *Psychosomatische Medizin* (5. Aufl.), S. 352–362. München Wien Baltimore: Urban & Schwarzenberg.
- Kächele, H. (1999). Eine multizentrische Studie zu Aufwand und Erfolg bei psychodynamischer therapie von Eßstörungen. Studiendesign und erste Ergebnisse. *Psychotherapie, Psychosomatik, Medizinische Psychologie* 49, 100–108.
- Kächele, H. (eingereicht). Aufwand und Ergebnis stationärer psychodynamischer Psychotherapie von Eßstörungen in Deutschland: Ergebnisse einer multizentrischen Studie. *Psychotherapy Research* .
- Keller, M., Lavori, P., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P. & Andreasen, N. (1987). The longitudinal interval follow-up evaluation. A com-

- prehensive method for assessing outcome in prospective longitudinal studies. *Archives of general psychiatry* 44, 540–548.
- Kordy, H. & W. Hannover (1998). Beobachten, Dokumentieren, Bewerten, Steuern: Qualitätsmanagement in der stationären Psychotherapie. In Vogel, A. L. H. (Hrsg.), *Qualitätssicherung in der Psychotherapie und der psychosozialen Versorgung*, S. 355–373. München: DGVT-Verlag.
- Kordy, H., Hannover, W., Richard, M. & Lambert, M. (angenommen). Der Ergebnisfragebogen EB-45: Normierung und Validierung der deutschen Version des Outcome Questionnaire von M. Lambert. *Zeitschrift für klinische Psychologie*.
- Kordy, H. & Lutz, W. (1995). Das Heidelberger Modell: Von der Qualitätskontrolle zum Qualitätsmanagement stationärer Psychotherapie durch EDV - Unterstützung. *Psychotherapie Forum* 3, 197–206.
- Kordy, H., Richard, M., Herrmann, A., Murphy, F., Treasure, J. & Charpentier, P. (1999). A computer assisted eating disorders specific quality management system: EQUAL-TREAT. *European Eating Disorders Review* 7, 239–258.
- Kordy, H. & Senf, W. (1985). Überlegungen zur Evaluation psychotherapeutischer Behandlungen. *Psychotherapie und medizinische Psychologie* 35, 207–212.
- Laireiter, A. & Vogel, H. (1998). *Qualitätssicherung in der Psychotherapie und psychosozialen Versorgung*. München: DGVT Verlag.
- Lambert, M. & Bergin, A. (1994). The effectiveness of psychotherapy. In Bergin, A.E. & Garfield, S. B. (Hrsg.), *Handbook of psychotherapy and behavior change* (4. Aufl.), S. 143–189. New York: John Wiley & Sons, Inc.
- Lambert, M., Burlingame, G., Umphress, V., Hansen, N., Vermeersch, D., Clouse, G. & Yanchar, S. (1996). The reliability and validity of the outcome questionnaire. *Clinical Psychology and Psychotherapy* 3, 106–116.
- Lambert, M., Vermeersch, D. & Hansen, N. (1999). The implementation of a quality management system into clinical practice: a controlled experiment. 30th annual international meeting, Society for Psychotherapy Research, Braga, Portugal, 16-20 June.
- Lavori, P., Dawson, R. & Mueller, T. (1994). Causal estimation of time-varying treatment effects in observational studies: application to depressive disorder. *Statistics in medicine* 13, 1089–1100.
- Littell, R., Milliken, G., Stroup, W. & Wolfinger, R. (1996). *SAS System for Mixed Models*. Cary, NC: SAS Institute Inc.
- Lueger, R. (1993). Assessing quality in outpatient psychotherapy: Implications for

- designing and selecting cost-effective mental health care benefits. Technical report, International foundation of employee benefit plans. Inc, Wisconsin.
- Lueger, R. (1995). Ein Phasenmodell der Veränderung in der Psychotherapie. *Der Psychotherapeut* 40, 267–278.
- Lueger, R. (1998). Using feedback on patient progress to predict the outcome of psychotherapy. *Journal of Clinical Psychology* 54(3), 383–393.
- Lutz, W. (1997). *Evaluation eines Qualitätssicherungsprogramms in der Psychotherapie*. Regensburg: Roderer.
- Lutz, W., Martinovich, Z. & Howard, K. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology* 67(4), 571–577.
- Lutz, W., Stammer, H., Leeb, B., Dötsch, M., Bölle, M. & Kordy, H. (1996). Das Heidelberger Modell der aktiven internen Qualitätssicherung stationärer Psychotherapie. *Psychotherapeut* 41, 25–35.
- Lyons, J., Howard, K., O'Mahoney, M. & Lish, J. (1997). *The measurement and management of clinical outcomes in mental health*. New York: Jon Wiley & Sons, Inc.
- Margraf, J. (1996). *Lehrbuch der Verhaltenstherapie (Bd. 1 und 2)*. Berlin: Springer.
- Meyer, A.-E., Richter, R., Grawe, K., Graf v.d. Schulenberg, J.-M. & Schulte, B. (1991). *Forschungsgutachten zu Fragen eines Psychotherapeutengesetzes*. Bonn: Bundesministerium für Jugend, Familie, Frauen und Gesundheit.
- Mohr, D. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology Science and Practice* 2(1), 1–27.
- Morgan, H. & Hayward, A. (1988). Clinical assessment of anorexia nervosa. *British Journal of Psychiatry* 152, 367–371.
- Nisslmüller, K. (1999). Ein Fragebogen zum Ergebnismonitoring im Rahmen des Qualitätsmanagements in der Psychotherapie. Erste Validierung und Normierung des Outcome Questionnaires von Michael Lambert an einer deutschen Stichprobe. Unveröffentlichte Diplomarbeit, Universität Trier.
- Olmsted, M., Kaplan, A., Rockert, W. & Jacobsen, M. (1996). Rapid responders to intensive treatment of bulimia nervosa. *International Journal of Eating Disorders* 19(3), 279–285.
- Orlinsky, D. & Howard, K. (1987). A generic model of psychotherapy. *Journal of Integrative and Eclectic Psychotherapy* 6(1), 6–27.

- Paul, G. (1969). Behavior modification research: Design and tactics. In C. Franks (Hrsg.), *Behavior Therapy: Appraisal and Status*, S. 29–62. New York: Mc Graw-Hill.
- Prochaska, J. & Norcross, J. (1994). *Systems of Psychotherapy. A Transtheoretical Analysis* (3. Aufl.). Pacific Grove, California: Brooks/Cole.
- Prochaska, J., Velicer, W., Guadagnoli, E., Rossi, J. & DiClemente, C. (1991). Patterns of change: Dynamic topology applied to smoking cessation. *Multivariate Behavioral Research* 26(1), 83–107.
- Quitkin, F., McGrath, P., Stewart, J., Ojepek-Welikson, K., Taylor, B., Nunes, E., Deliyannides, D., Agosti, V., Donovan, S., Petkova, E. & Klein, D. (1996). Chronological milestones to guide drug change. *Archives of General Psychiatry* 53, 785–792.
- Reinecker, H. & Fiedler, P. (1997). *Therapieplanung in der modernen Verhaltenstherapie*. Lengerich: Papst.
- Richard, M. (1996). Die Auswertung von Therapieverläufen mit Hilfe Hierarchischer Linearer Modelle. Unveröffentlichte Diplomarbeit, Universität Heidelberg.
- Schulte, D. (1996). *Therapieplanung*. Hogrefe: Göttingen.
- Schulte-Bahrenberg, T. & Schulte, D. (1993). Change of psychotherapy goals as a process of resignation. *Psychotherapy Research* 3, 153–165.
- Searle, S. (1988). Mixed models and unbalanced data: Wherefrom, whereat and whereto? *Communications in Statistics - Theory and Methods* 4(17), 935–968.
- Searle, S., Casella, G. & McCulloch, C. (1992). *Variance Components*. New York: John Wiley & Sons, Inc.
- Snijders, T. & Bosker, R. (1999). *Multilevel Analysis*. London: Sage Publications Ltd.
- Sperry, L. (1997). Treatment outcomes: An overview. *Psychiatric Annals* 27(2), 95–99.
- Sperry, L., Brill, P., Howard, K. & Grissom, G. (1996). *Treatment outcomes in psychotherapy and psychiatric interventions*. New York: Brunner/Mazel.
- Strober, M., Freeman, R. & Morrell, W. (1997). The long-term course of severe anorexia nervosa in adolescents: Survival analysis of recovery, relapse, and outcome predictors over 10–15 years in a prospective study. *International Journal of Eating Disorders* 22, 339–359.
- Strupp, H. & Hadley, S. (1977). A tripartite model of mental health and therapeutic outcomes. *American Psychologist* (März), 187–196.

-
- Tang, T. & DeRubeis, R. (1999). Reconsidering rapid early response in cognitive behavioral therapy for depression. *Clinical Psychology: Science and Practice* 6(3), 1999.
- Thiel, A. & Paul, T. (1988). Entwicklung einer deutschsprachigen Version des Eating-Disorder-Inventory (EDI). *Zeitschrift für Differentielle und Diagnostische Psychologie* 9(4), 267–278.
- Umphress, V., Lambert, M., Smart, D., Barlow, S. & Clouse, G. (1997). Concurrent and construct validity of the outcome questionnaire. *Journal of Psychoeducational Assessment* 15, 40–55.
- Velicer, W., Rossi, J., Prochaska, J. & DiClemente, C. (1996). A criterion measurement model for health behavior change. *Addictive Behaviors* 21(5), 555–584.
- Vogel, G. & Schulte, D. (1997). Methoden- und verlaufsorientierte Strategien von Psychotherapeuten. *Zeitschrift für Klinische Psychologie* 26(1), 38–49.
- Wilson, G. (1998). Manual-based treatment and clinical practice. *Clinical Psychology: Science and Practice* 5(3), 363–375.
- Wilson, G. (1999). Rapid response to cognitive behavior therapy. *Clinical Psychology: Science and Practice* 6(3), 289–292.
- Wittchen, H.-U. (1989). *Diagnostisches und Statistisches Manual psychischer Störungen*. Weinheim; Basel: Beltz.

15 Anhang

Der Kurzbeurteilungsbogen aus der multizentrischen Studie, MZ-ESS.

Kurzbeurteilung PatientIn

Klinik-Code: Datum:
 Code-Nr.Pat:

1. Jetziges Körpergewicht in kg (unbekleidet): kg

2. Fürchten Sie sich davor, dick zu werden bzw. zuzunehmen?

- ich habe keine Angst davor, dick zu werden bzw. zuzunehmen
- ich fürchte mich gelegentlich davor, dick zu werden bzw. zuzunehmen
- ich bin in ständiger Sorge, dick zu werden bzw. zuzunehmen
- der Gedanke daran, dick zu werden bzw. zuzunehmen, löst bei mir Panik aus
- ich möchte eher sterben als dick sein bzw. zuzunehmen

3. Wie finden Sie Ihren Körper (Bitte kreuzen Sie jeweils an, wo zwischen den beiden Polen Sie sich selbst sehen)

viel zu dünn	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	viel zu dick
anziehend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	abstoßend
wohlproportioniert	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unproportioniert
muskulös	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	schlaff
weiblich (bei Frauen)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unweiblich (bei Frauen)
männlich (bei Männern)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	unmännlich (bei Männern)

4. Wie oft durchschnittlich haben Sie die betreffenden Maßnahmen in den letzten 4 Wochen angewandt? (bitte machen Sie ein Kreuz in jede Zeile)

gar nicht	weniger als 1 x pro Woche	1 x pro Woche	2-3 x pro Woche	täglich	mehrmals täglich	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Erbrechen
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Einnahme von Abführmitteln
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Diät bzw. niederkalorische Nahrung
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sonstiges:

5. Wie oft kommen Eßanfälle in den letzten 4 Wochen durchschnittlich vor?

.... mal/Tag oder mal/Woche oder seltener mal/Monat oder gar nicht

6. Hatten Sie in den letzten 4 Wochen Ihre Regelblutung?

- ja nein bei Männern: trifft nicht zu

7. Nehmen Sie derzeit die Antibaby-Pille ein?

- ja nein bei Männern: trifft nicht zu