

2 Theoretischer Hintergrund

Im folgenden Kapitel sollen drei Variablen, auf denen Therapeutenunterschiede zu erwarten sind, als potentielle Indikatoren für effektive Therapeuten vorgestellt werden. Es handelt sich dabei um das Outcome der Patienten, die Sitzungsanzahl der Therapien und die Dropout-Rate der Therapeuten. Darüber hinaus soll ein Überblick der bisherigen Forschung zu Therapeutenunterschieden auf diesen Variablen angeboten werden. Soweit der aktuelle Forschungsstand es zulässt, werden danach die Zusammenhänge zwischen diesen drei Variablen dargestellt. Abschließend soll ein kurzer Vergleich von zwei grundlegenden statistischen Ansätzen durchgeführt werden, um in die Bayes'sche Statistik einzuführen, derer sich einige der folgenden Analysen bedienen werden.

2.1 Therapeuteneffekte

Zur Einführung in das Thema sollen kurz der Therapeuteneffekt im Allgemeinen und die Vorteile von sogenannten Zufallsfaktoren dargestellt werden. Die Unterschiede zwischen Therapeuten in Bezug auf ein bestimmtes Merkmal ihrer Patienten, im vorliegenden Fall das Outcome, die Sitzungsanzahl und der Dropout-Status, werden im sogenannten Therapeuteneffekt quantifiziert. Dabei können die Therapeuten bei der Auswertung als feste (*fixed effects*) oder als Zufallsfaktoren (*random effects*) modelliert werden. Ältere Studien haben Therapeuten immer als feste Faktoren erhoben, wohingegen aktuelleren Untersuchungen die Möglichkeiten der Zufallsfaktoren zur Verfügung stehen. Ein fester Faktor umfasst eine

bekannte Anzahl an Ausprägungen, die alle realisiert werden. Im Gegensatz dazu werden die Ausprägungen eines Zufallsfaktors als Zufallsstichprobe aus einer zugrundeliegenden Population von Ausprägungen betrachtet. Man geht also davon aus, dass die beobachteten Therapeuten eine repräsentative Stichprobe aller vergleichbaren Therapeuten sind (Baldwin & Imel, 2013). Statistisch gesehen erhält man anstelle eines festen Koeffizienten, der für einen Zusammenhang oder eine Differenz steht, für jeden Therapeuten dessen Abweichung vom Durchschnitt. Die Varianz dieser Abweichungen ist der interessierende Parameter, der die Variabilität der durchschnittlichen Merkmalsausprägung einzelner Therapeuten um den Gesamtmittelwert darstellt. Für das Outcome bedeutet dies zum Beispiel, dass die Variabilität des mittleren Outcomes jedes Therapeuten um das durchschnittliche Outcome aller Patienten beobachtet wird; dies gilt ebenso für die Sitzungsanzahl und die Dropout-Rate. Der Therapeuteneffekt gibt den Anteil der Varianz zwischen den Therapeuten an der Gesamtvarianz des Kriteriums an. Ein großer Effekt bedeutet demnach, dass sich verschiedene Therapeuten stark in der mittleren Ausprägung ihrer Patienten unterscheiden. Der große Vorteil von Zufallsfaktoren besteht darin, dass die Ergebnisse nicht auf die realisierten Ausprägungen beschränkt sind, sondern auf die Population generalisiert werden können. Somit können aus Analysen mit Zufallsfaktoren Aussagen über alle Therapeuten – und nicht nur über die empirisch beobachteten – abgeleitet werden. Aufgrund dieses Vorteils verwenden fast alle aktuellen Studien diese Auswertungsmethode. Die folgende Übersicht über den derzeitigen Forschungsstand zu Therapeuteneffekten auf Outcome, Sitzungsanzahl und Dropout soll sich ebenfalls darauf beschränken.

2.2 Outcome

Die erste Variable, die als Indikator für effektive Therapeuten in Frage kommt, ist das Outcome der Patienten. Bei einem erfolgreichen Therapeuten ist ein positives Outcome zu erwarten, wohingegen Patienten, die von einem weniger erfolgreichen Therapeuten behandelt werden, ein weniger positives oder sogar ein negatives Outcome erreichen können.

2.2.1 Allgemeines zum Outcome

Das Outcome eines Patienten beschreibt das Ergebnis einer Therapie, das nach deren Beendigung erreicht wurde. Es kann auf verschiedene Weisen ausgedrückt werden, wobei man es nach der Messmethode und dem Messinstrument einteilen kann. Methodisch wird entweder der Messwert eines Patienten direkt nach der Therapie, also zum Post-Messzeitpunkt, oder die Veränderung über die Therapie hinweg als Outcome verwendet. Die Veränderungsmessung kann direkt erfolgen, sodass diese retrospektiv eingeschätzt wird, oder sie wird indirekt erhoben, indem ein Differenzwert aus der Post- und der Prä-Messung gebildet wird. Außerdem kann das Outcome vom Patienten selbst, vom behandelnden Therapeuten oder von einem objektiven Dritten eingeschätzt werden. Als Messinstrumente werden störungsspezifische und störungsunspezifische Instrumente eingesetzt. Erhoben werden neben den Symptomen vor allem die generelle Belastung, das Wohlbefinden und das allgemeine Funktionsniveau. Ebenso wird teilweise anstelle der absoluten Veränderung vom Beginn zum Ende der Therapie die klinisch signifikante Veränderung nach Jacobson & Truax (1991) berichtet. Lambert (1983, zitiert nach Lambert, 2013) nimmt an, dass zu den wichtigsten Aspekten des Outcomes die subjektive

Erfahrung der Patienten, das Funktionieren der Patienten in sozialen Rollen und ihr Funktionieren in intimen Beziehungen gehören. Gegen die Nutzung von Outcome-Maßen im psychotherapeutischen Prozess wurden von den Anwendern praktische Gründe, wie Kosten und Zeitaufwand, sowie philosophische, wie deren Relevanz, vorgebracht (Hatfield & Ogles, 2004). Aktuelle Untersuchungen zeigen jedoch die positiven Effekte der regelmäßigen Erhebung von Outcome-Maßen (Lambert et al., 2003). Viele Instrumente sind zudem mittlerweile kostenfrei zugänglich und verkürzte Outcome-Fragebögen erleichtern die Verwendbarkeit im klinischen Alltag (Lambert, 2013; Lambert & Hawkins, 2004). Das Outcome ist das meistgenutzte Merkmal von Patienten zur Untersuchung von Therapeutenunterschieden und bietet sich darüber hinaus als ein Maß für die Effektivität von Therapeuten an.

2.2.2 Therapeuteneffekte und -unterschiede im Outcome

Viele der verfügbaren Studien zu Therapeuteneffekten auf Outcome wurden in zwei großen Meta-Analysen zusammengefasst (Crits-Christoph et al., 1991; Baldwin & Imel, 2013). Bereits vor über 20 Jahren haben Crits-Christoph et al. (1991) aus den Daten von 15 Studien Therapeuteneffekte berechnet, indem sie die Therapeuten a posteriori als Zufallsfaktor modelliert haben. Die Studien stammten aus den Jahren 1965 – 1990 und umfassten insgesamt 27 Behandlungsgruppen, die unterschiedliche Formen von Psychotherapie bekamen, und 141 Therapeuten. Die Autoren fanden sehr heterogene Therapeuteneffekte, die von 0% bis 48.7% reichten. Wurde ein Manual befolgt und waren die Therapeuten erfahren, verringerte sich die Höhe des Therapeuteneffekts signifikant. Diese beiden

Prädiktoren konnten gemeinsam 33% der Varianz in den durchschnittlichen Therapeuteneffekten erklären. Wenn nur die Studien mit den größten Therapeuteneffekten betrachtet wurden, waren es sogar 34%. Im Mittel berichten die Autoren einen moderaten Effekt von 8.6% der Varianz des Outcomes für die Differenz zwischen den Therapeuten. In einer aktuelleren Meta-Analyse wurden die Therapeuteneffekte von 46 Studien mit Therapeuten als Zufallsfaktor zusammengefasst (Baldwin & Imel, 2013). Dabei ergaben sich große Überschneidungen mit den Studien aus der Analyse von Crits-Christoph et al. (1991). Der mittlere Therapeuteneffekt lag für die 46 Studien mit 1,281 Therapeuten und 14,519 Patienten bei 5%, mit einem 95%-Konfidenzintervall (KI) von 3 – 7%. Es zeigte sich, dass der Therapeuteneffekt für naturalistische Studien mit 7%, 95% KI[5; 10], größer ausfällt als für experimentelle Untersuchungen mit 3%, 95% KI[1; 5]. Da bei sogenannten *randomized controlled trials* (RCTs) zur Erhöhung der internen Validität alle unsystematische Variation des Outcomes minimiert werden soll, reduzieren sich auch die Unterschiede zwischen den Therapeuten. Dies wird u.a. durch die Manualisierung des Therapieprozesses erreicht. Dieser Befund deckt sich mit den zuvor präsentierten Ergebnissen von Crits-Christoph et al. (1991) zum Einfluss von Manualen auf die Größe des Therapeuteneffektes.

Die meisten Studien aus den beiden Meta-Analysen fanden im ambulanten Setting statt. Therapeuteneffekte konnten aber auch im stationären Bereich nachgewiesen werden (Dinger, Strack, Leichsenring, Wilmers & Schauenburg, 2008). Wurde das Outcome von den Patienten eingeschätzt, zeigte sich ein Therapeuteneffekt von 3%. Bei einer Therapeutenbeurteilung stieg der Anteil der Therapeutenvarianz an der Gesamtvarianz auf 16% an. Damit liegt die Schätzung bei einer Patientenbeurteilung des Outcomes unter den Ergebnissen aus vergleichbaren Untersuchungen im

ambulanten Sektor, die Effekte von 8.3% respektive 6.2% gefunden haben (Lutz, Leon, Martinovich, Lyons & Stiles, 2007; Lutz et al., im Druck). Um die Größe des Effektes global einordnen zu können, soll zum Vergleich erwähnt werden, dass der Therapeuteneffekt auf die von den Patienten eingeschätzte therapeutische Allianz in derselben Studie 33% betrug (Dinger et al., 2008). Die therapeutische Allianz – auch therapeutische Beziehung genannt – umfasst die Bindung zwischen Patient und Therapeut sowie die Übereinstimmung in den Zielen und den Aufgaben, die zur Zielerreichung notwendig sind (Bordin, 1979). Die Therapeuten unterschieden sich demnach zehnmal stärker in der durchschnittlichen Qualität ihrer therapeutischen Beziehungen zu den Patienten als in dem mittleren Outcome ihrer Patienten.

Neben dem Outcome-Maß zum Post-Messzeitpunkt variiert auch der Zusammenhang der Prä- und Postwerte der Patienten zwischen den Therapeuten. Im stationären Setting konnte die Interaktion zwischen den Therapeuten und der Ausgangsbelastung der Patienten 6.5% der Varianz im vom Patienten beurteilten Outcome aufklären (Dinger et al., 2008). Das heißt, dass eine starke Ausgangsbelastung zwar im Durchschnitt mit einem schlechteren Outcome assoziiert war, dieser Zusammenhang aber nicht für alle Therapeuten gleich stark ausfiel. Im ambulanten Setting wurde mithilfe von Wachstumskurven eine Varianzaufklärung durch Therapeutenunterschiede in der Veränderungsrate der Patienten von 17.3% gefunden (Lutz et al., 2007). Die prozentuale Aufklärung der Varianz des Outcome-Maßes betrug, vergleichbar mit den zuvor dargestellten Befunden, in derselben Studie 8.3%. Unter Berücksichtigung von variablen Zusammenhängen zwischen den Prä- und Post-Werten des Outcomes berichteten Kim, Wampold und Bolt (2006) Therapeuteneffekte von 8 – 12% für eine mittlere Ausgangsbelastung. Wenn die Autoren die Zusammenhänge auf einen

für alle Patienten einheitlichen Wert fixierten, zeigten sich ähnliche Ergebnisse: Für die Gesamtstichprobe (sogenannte *Intent-to-Treat-Analyse*) betrug der Therapeuteneffekt 1 – 9%, für die Substichprobe aus Patienten, die die Therapie abgeschlossen haben, (sogenannte *Completer-Analyse*) betrug er 5 – 10%. In Abhängigkeit von dem Outcome-Maß und dem statistischen Modell lagen die gefundenen Effekte bei 1 – 12% und somit bei durchschnittlich 8% durch Therapeuten aufgeklärte Varianz (Kim et al., 2006).

In die bisher berichteten Therapeuteneffekte, die sich grob zwischen einem und zehn Prozent bewegen, reihen sich die Befunde von Saxon und Barkham (2012) ein. Im Unterschied zu den anderen Studien, wurden die Parameter in dieser Untersuchung mit Markov Chain Monte Carlo (MCMC) Methoden geschätzt, die bei komplexen Modellen zu reliableren Ergebnissen führen (Browne, 2009). Sie fanden Therapeuteneffekte von 7.8%, wenn nur die Ausgangsbelastung der Patienten als Prädiktor aufgenommen wurde, und 6.6%, wenn zusätzlich der *Case Load* der Therapeuten Berücksichtigung fand (Saxon & Barkham, 2012). Unter dem *Case Load* versteht man im Allgemeinen die Anzahl der Patienten und den Schweregrad ihrer Belastung. Saxon und Barkham (2012) haben dazu die Ausgangsbelastung und die Ausprägung der Selbst- und Fremdgefährdung der Patienten herangezogen. Die Varianzaufklärung durch die Therapeuten variierte in Abhängigkeit von dem Schweregrad der Ausgangsbelastung: Bei geringer Belastung lag der Therapeuteneffekt bei unter 3%, während er bei stärkerer Belastung der Patienten auf bis zu 10% anstieg. Gerade bei stark belasteten Patienten sollte der Therapeut demnach entscheidenden Einfluss auf das Outcome haben.

Untersuchungen zur Erklärung der Therapeutenunterschiede in den Outcomes ihrer Patienten haben bisher heterogene und widersprüchliche Ergebnisse geliefert (Baldwin & Imel, 2013). Die konsistentesten Zusammenhänge zeigen sich mit der therapeutischen Allianz. Therapeutenunterschiede von durchschnittlich 9%, 95% KI[2; 18], in der Allianz konnten in einer Meta-Analyse gefunden werden (Baldwin & Imel, 2013). Die Korrelation zwischen der Allianz und dem Outcome wurde in einem Multilevel-Modell innerhalb der Therapeuten nicht signifikant, während derselbe Zusammenhang zwischen den Therapeuten negativ und statistisch signifikant war (Baldwin, Wampold & Imel, 2007). Patienten, die von Therapeuten mit höheren Allianz-Werten behandelt wurden, hatten im Durchschnitt bessere Outcomes. Varianz in den Allianz-Werten innerhalb eines Therapeuten hatte hingegen keinen Einfluss auf das Outcome der Patienten. Das bedeutet, dass sich Therapeuten in ihrer Fähigkeit unterscheiden, starke Allianzen mit ihren Patienten zu bilden, und dies einen Einfluss auf das Outcome der Patienten hat. Für weitere Therapeutenvariablen, wie Ausbildungslevel, Art der Ausbildung, Geschlecht und theoretische Orientierung, konnten keine Effekte auf das Outcome der Patienten gefunden werden (Okiishi, Lambert, Nielsen & Ogles, 2003; Okiishi et al., 2006). Die Unterschiede zwischen den Therapeuten sind demnach das Ergebnis nicht erhobener Therapeutenvariablen.

Bei ihrer Suche nach dem sogenannten *Supershrink* fanden Okiishi et al. (2003) signifikante Unterschiede in der Veränderungsrate der 1,799 Patienten zwischen den 56 beobachteten Therapeuten. Ähnlich wie Lutz et al. (2007) modellierten sie Wachstumskurven, die sie auf Therapeuten-Ebene aggregierten, um die durchschnittliche Verbesserung der Patienten eines Therapeuten zu erhalten. Allerdings berichteten sie keine Werte für

die Therapeuteneffekte, sondern verglichen die Wachstumskurven miteinander. Zwischen den erfolgreichsten und den am wenigsten erfolgreichen Therapeuten, bezogen auf die Symptomverbesserung ihrer Patienten, fanden sich deutliche Differenzen in der Veränderungsrate; über eine Standardabweichung in der Differenz des Outcome-Maßes. So konnten die Autoren die Therapeuten anhand des Ausmaßes der Verbesserung ihrer Patienten in eine Rangreihe bringen. Auch mit einer größeren Anzahl an Patienten pro Therapeut (30 vs. 15) konnten die Ergebnisse repliziert werden (Okiishi et al., 2006). Allerdings veränderte sich die Rangreihe der Therapeuten deutlich, wenn anstelle der Veränderungsrate der Patienten die Prä-Post-Differenz gemessen wurde. Die Korrelation zwischen den beiden Kriterien betrug dabei signifikante, aber moderate $r = .46$ ($p < .001$). Die besten und schlechtesten Therapeuten, nach beiden Kriterien, unterschieden sich statistisch und klinisch signifikant im mittleren Outcome ihrer Patienten.

2.3 Sitzungsanzahl

Neben dem Outcome spielt auch die durchschnittliche Sitzungsanzahl ihrer Patienten eine entscheidende Rolle in der Beurteilung von Therapeuten. Darum ist die Sitzungsanzahl die zweite Variable, die als Indikator für die Effektivität von Therapeuten genutzt werden kann. Bei erfolgreichen Therapeuten kann eine kürzere Sitzungsanzahl und somit ein effizienteres Arbeiten erwartet werden, während weniger erfolgreiche Therapeuten mehr Sitzungen für ihre Therapien benötigen.

2.3.1 Allgemeines zur Sitzungsanzahl

Die Sitzungsanzahl bezeichnet die Anzahl der Sitzungen, die ein Patient zwischen Therapiebeginn und Therapieabschluss oder -abbruch in Behandlung verbracht hat. In Deutschland werden von den Krankenkassen die Kosten für fünf probatorische Sitzungen übernommen, nach denen über das Zustandekommen einer Therapie entschieden werden muss. Wird ein Antrag auf Therapie vom Patienten gestellt und von dem Kostenträger genehmigt, folgen die therapeutischen Sitzungen. Probatorische Sitzungen sind ein deutsches Phänomen, das in den meisten anderen Ländern nicht existiert. Die Anzahl der Sitzungen wird formal von der Art der Therapie (kognitiv-verhaltenstherapeutisch vs. psychoanalytisch oder psychodynamisch) und der beantragten Dauer (Kurzzeit- vs. Langzeittherapie) bestimmt. Unter bestimmten Voraussetzungen können Verlängerungsanträge gestellt werden. Es existieren jedoch große Differenzen bezüglich der Sitzungsanzahl zwischen den Staaten, in denen die Therapie beantragt und durchgeführt wird. Unabhängig von der Anzahl genehmigter Sitzungen kann die Therapie jederzeit beendet werden, indem der Patient sie unilateral abbricht oder der Therapeut, in Absprache mit dem Patienten, die Therapie als regulär beendet einstuft.

Therapeuten belegen deutlich unterschiedliche Rangplätze in Bezug auf ihre Resultate, wenn sie nicht nach der Veränderungsrate ihrer Patienten, sondern nach dem Produkt aus der Veränderung und der Sitzungsanzahl beurteilt werden (Okiishi et al., 2003). Es gilt also, zwischen der Effektivität eines Therapeuten und seiner Effizienz zu unterscheiden (Okiishi et al., 2006). Die bisher dargestellten Befunde bezogen sich auf die reine Veränderung über die Zeit, was einen Indikator für die Effektivität darstellt. Therapeuten, die für eine ähnliche absolute Veränderung (vom

Prä- zum Postzeitpunkt) aber weniger Sitzungen benötigen, sollten als effizienter betrachtet werden. Sie führen bei ihren Patienten schneller zu Verbesserungen, die ein Therapieende ermöglichen, behandeln in der gleichen Zeit mehr Patienten und arbeiten aus gesundheitspolitischer Sicht kostenschonender.

2.3.2 Therapeuteneffekte und -unterschiede in der Sitzungsanzahl

Die Sitzungsanzahl wurde erstmals von Okiishi et al. (2003) zur Bewertung der Outcomes, die Therapeuten produzieren, herangezogen. Die Autoren berichten, dass es signifikante Differenzen in der Behandlungslänge zwischen den Therapeuten mit den längsten und den kürzesten Therapien gab ($t = 9.42, p < .01$). Da die Therapeuten mit der größten Veränderungsrate zeitgleich die kürzesten Therapien durchführten, mussten beide Variablen betrachtet werden, um die tatsächliche Verbesserung der Patienten darstellen zu können. Dazu haben die Autoren die Veränderungsrate mit der durchschnittlichen Sitzungsanzahl multipliziert. Es zeigte sich, dass der Therapeut mit der stärksten Veränderung pro Zeit nur noch auf Platz 38 von 56 lag, wenn die Sitzungsanzahl miteinbezogen wurde (Okiishi et al., 2003). Während Okiishi et al. (2003) also die Zeit mitberücksichtigen mussten, um die absolute Veränderung der Patienten aus der Veränderungsrate ableiten zu können, stehen wir vor dem umgekehrten Fall: Die meisten Analysen der Effektivität von Therapeuten beziehen sich ausschließlich auf das Outcome, obwohl auch die Zeit, die zum Erreichen dieses Outcomes benötigt wird, eine hohe Relevanz besitzt. Um die Sitzungsanzahl in die Analyse miteinzubeziehen, kann anstelle der Veränderungsrate auch die Prä-Post-Differenz der Outcome-Variable betrachtet werden

(Okiishi et al., 2006). Hier finden sich ebenfalls signifikant unterschiedliche Rangplätze für die Therapeuten. Die Autoren empfehlen, sowohl die Veränderung vom Prä- zum Postzeitpunkt als auch die Veränderungsrate zu erheben. Alternativ sollten dieselben Informationen über das gleichzeitige Betrachten der Veränderung und der Sitzungsanzahl erhalten werden.

Eine erste Studie hat diese beiden Variablen in einer Stichprobe aus 44 Therapeuten erhoben, die 349 Patienten behandelt haben (Lutz et al., im Druck). Es zeigte sich ein Therapeuteneffekt von 9% für die Therapielänge – operationalisiert über die Sitzungsanzahl –, wenn für die Anzahl der bewilligten Sitzungen kontrolliert wurde. Es konnten also 9% der Differenzen in der Sitzungsanzahl der Patienten durch Unterschiede zwischen den Therapeuten erklärt werden. Zusätzlich wurde der Zusammenhang zwischen der mittleren Sitzungsanzahl und dem mittleren Outcome der Therapeuten untersucht. Es konnte keine signifikante Beziehung zwischen den beiden Merkmalen gefunden werden; die Korrelation betrug $r = .09$ ($p = .56$). Das bedeutet, dass die Therapeuten mit dem besten Outcome nicht auch die Therapeuten sind, die im Durchschnitt längere oder kürzere Interventionen durchführen (Lutz et al., im Druck).

2.4 Dropout

Selbst hochwirksame therapeutische Methoden, die von kompetenten Therapeuten angewendet werden, können ihre Effektivität nicht entfalten, wenn die Therapie vorzeitig abgebrochen wird. Solche Therapieabbrüche werden auch Dropouts genannt. Deshalb wird die Dropout-Rate eines Therapeuten als dritte Variable aufgenommen, die einen Indikator für

die Effektivität von Therapeuten darstellen kann. Erfolgreiche Therapeuten sollten weniger Therapieabbrüche zu verzeichnen haben als weniger erfolgreiche Kollegen.

2.4.1 Allgemeines zu Dropout

Dropout ist nicht so eindeutig zu definieren, wie es die Sitzungsanzahl ist. Aktuell finden sich vier geläufige Definitionen, die als Operationalisierungen von Therapieabbrüchen genutzt werden: die Einschätzung des Therapeuten (Chisholm, Crowther & Ben-Porath, 1997), das Nichtwahrnehmen des letzten vereinbarten Termins (Hatchett, Han & Cooker, 2002), die Sitzungsanzahl (Baekeland & Lundwall, 1975) und das Ausbleiben eines weiteren Kontaktes nach dem Aufnahmetermin (*intake-only*; Longo, Lent & Brown, 1992).

Das Zeitkriterium ist dabei die invalideste Operationalisierung, da es frühe Abbrüche und frühe aber reguläre Therapieabschlüsse vermischt (Pekarik, 1985b). Weiterhin wird es meist als Median-Split durchgeführt, sodass der Grenzwert in Abhängigkeit von den Studiencharakteristika variiert und ein Vergleich zwischen Studien unmöglich wird. Zudem fand Pekarik (1985b) auf 18 Variablen keine signifikanten Differenzen zwischen Dropouts und Completern, wohingegen sie sich bei einer Therapeutenbeurteilung des Therapieendes auf 11 von 18 Variablen unterschieden. Das Zeitkriterium ist also nicht trennscharf zwischen Patienten, die die Therapie abbrechen und denen, die sie regulär abschließen. Ebenso wird das Kriterium des letzten vereinbarten Termins kritisiert, da somit Patienten, die kurz vor dem Ende der Therapie stehen, fälschlicherweise als Dropouts kategorisiert werden. Zeitgleich werden Therapieabbrecher, die keinen weiteren Termin vereinbaren, nicht als Dropouts gezählt (Pekarik,

1985b). Wird das Nichtwahrnehmen des letzten Termins als Definition benutzt, findet sich eine geringere Dropout-Rate als bei der Therapeutenbeurteilung oder dem Zeitkriterium (Wierzbicki & Pekarik, 1993), wodurch die tatsächliche Rate der Therapieabbrüche unterschätzt werden kann. Die Beurteilung des Abbruchs durch den Therapeuten hingegen ist augenscheinlich, basiert auf der Erfahrung des Therapeuten mit dem Patienten und wird deshalb als überlegene Methode, Dropouts zu definieren, betrachtet (Wierzbicki & Pekarik, 1993). Außerdem können Therapeuten, eher als ein Zeitkriterium, konzeptuell zwischen frühen Dropouts und frühen regulären Abschlüssen differenzieren (Pekarik, 1985b). Eine Einschränkung bei dieser Definition stellt die vergleichsweise geringe Reliabilität dar, da unterschiedliche Therapeuten verschiedene Kriterien heranziehen könnten. Auch dadurch sind ihre Urteile nicht immer akkurat (Garb, 2005; Grove, Zald, Lebow, Snitz & Nelson, 2000).

Einen direkten Vergleich zwischen diesen vier Definitionen von Dropout haben Hatchett und Park (2003) durchgeführt. Die höchste Rate erhielten sie über den Median-Split (53.1%), die geringste über *intake-only* (17.6%). Die Therapeuteneinschätzung und die Klassifikation über das Nichtwahrnehmen der letzten Sitzung führten zu identischen Dropout-Raten (40.8%). Die Übereinstimmung zwischen *intake-only* und der Verfehlung des letzten Termins betrug $\kappa = .26$, zwischen letzterem und der Therapeuteneinschätzung hingegen $\kappa = .62$. Die Definitionen sollten demnach nicht austauschbar benutzt werden, da sie nicht dasselbe messen, wobei die Therapeuteneinschätzung mit dem Verpassen eines vereinbarten letzten Termins stark zusammenhängt. In Anlehnung an Pekarik (1985b) und Garfield (1994) sehen die Autoren das Zeitkriterium und *intake-only* nicht als valide Operationalisierungen von Dropout an (Hatchett & Park, 2003).

Da ihnen die Therapeutenbeurteilung jedoch zu unreliabel ist und diese nur schwer über Studien hinweg verglichen werden kann, schlagen sie als neue Methode eine Definition von Dropout über die reliable Veränderung (*reliable change*, RC) oder die klinisch signifikante Veränderung (*clinically significant change*, CSC; Jacobson & Truax, 1991) vor. In einem empirischen Vergleich der Definitionen zeigte sich, dass die CSC ein deutlich konservativeres Kriterium darstellt als die bisherigen vier Methoden (Swift, Callahan & Levine, 2009). Die Dropout-Rate war auch in dieser Untersuchung am geringsten für *intake-only* (8.1%), gefolgt von dem Nichtwahrnehmen der vereinbarten Sitzung (48.1%) und dem Zeitkriterium (50%). Im Unterschied zu den Raten bei Hatchett und Park (2003) beurteilten die Therapeuten ganze 74.1% der Patienten als Dropouts. Dennoch lag die Dropout-Rate nach dem Kriterium der CSC mit 77% noch höher (Swift et al., 2009). Auch hier hingen, neben der RC und der CSC, nur die Therapeutenbeurteilung und die nichtwahrgenommene letzte Sitzung bedeutsam zusammen. Die Autoren resümieren, dass die vier klassischen Definitionen teils klinisch signifikant verbesserte Patienten als Dropout und teils auch nicht verbesserte als Completer klassifizieren. Diese fehlende Übereinstimmung mit der CSC war für die Therapeutenbeurteilung am geringsten ausgeprägt. Obwohl Swift et al. (2009) diese neue Methode als akkurat, logisch, valide und reliabel beschreiben, bleibt die Frage offen, ob die CSC wirklich ein Kriterium für Dropout sein kann oder als Outcome-Maß auf einer anderen Dimension der Therapeuteneffektivität anzusiedeln sein muss. In dieser Definition ist Dropout mit dem Outcome des Patienten gleichzusetzen. Um eine differenziertere Beurteilung eines Therapeuten zu erlauben, sollte jedoch zwischen dem Outcome seiner Patienten und der Dropout-Rate unterschieden werden können. Zudem schlägt dieses Verfahren bei der Klassifikation von Dropouts fehl, wenn

nicht Symptomreduktion das primäre Ziel der Therapie ist oder anstelle von störungsspezifischen Fragebögen globale Outcome-Maße herangezogen werden (Swift et al., 2009). Solche Probleme können durch die Therapeutenbeurteilung des Dropouts umgangen werden.

Die einzige Differenz zwischen Dropouts und Completern in einer Studie von Bados, Balaguer und Saldaña (2007) zeigte sich bezüglich der Sitzungsanzahl. Im Durchschnitt unterschieden sie sich auch nicht in ihrer Zufriedenheit mit dem Treatment, der Einschätzung der therapeutischen Allianz und der wahrgenommenen globalen Verbesserung, allerdings variierten die Verteilungsformen (Simon, Imel, Ludman & Steinfeld, 2012). Patienten, die nach der ersten Sitzung nicht erneut erschienen waren, berichteten mehr sehr hohe und mehr sehr geringe Werte auf den drei Variablen. So gaben ca. 30% die höchste Zufriedenheit, über 60% die höchste Allianz und über 40% starke Verbesserungen der Symptome an. Gleichzeitig waren diese Patienten mit der geringsten Zufriedenheit mit dem Treatment assoziiert und bei über 25% der betroffenen Patienten hatten sich die Symptome nicht verändert oder sogar verschlechtert. Diese Befunde weisen darauf hin, dass bei Dropouts nicht von einer homogenen Population gesprochen werden kann (vgl. auch Mueller & Pekarik, 2000). Darum sollten neben dem Status (Therapie abgebrochen vs. regulär beendet) auch Outcome-Variablen berücksichtigt werden.

Wird die Dropout-Rate in die Effektivität eines Treatments eingerechnet, so reduziert sich die Effektivität deutlich (Bados et al., 2007). In einem Beispiel der Autoren sinkt der Anteil der verbesserten Patienten von ursprünglich 70% auf 56%, was zeigt, dass die Wirksamkeit der Intervention ohne Berücksichtigung der Dropout-Rate überschätzt wird. Neben diesen methodischen Gründen sprechen auch inhaltliche für eine Aufnahme der Dropout-Rate in die Beurteilung von Therapeuten. Dropouts haben

nicht nur negative Folgen für die Patienten selbst, sondern führen auch zu klinischen, finanziellen und moralischen Problemen bei ihren Therapeuten und den verantwortlichen Institutionen (Pekarik, 1985a; Swift, Greenberg, Whipple & Kominiak, 2012). Darüber hinaus bleibt die Belastung für Angehörige von Patienten, die die Therapie abbrechen, bestehen (Klein, Stone, Hicks & Pritchard, 2003). Deshalb stellen sie einen wichtigen Faktor in Bezug auf Effektivität und Ökonomie therapeutischer Interventionen dar.

2.4.2 Therapeuteneffekte und -unterschiede in der Dropout-Rate

Die Dropout-Rate wurde bisher nur in einer Studie erhoben, welche die Therapeuten als Zufallsfaktor in die Analyse aufgenommen hat (Huppert et al., 2014). Allerdings wurde kein Therapeuteneffekt berechnet. Therapeutenunterschiede in Bezug auf die mittlere Dropout-Rate können somit nicht berichtet werden. Allerdings wurden in anderen Studien einige Therapeutenvariablen identifiziert, die die Dropout-Wahrscheinlichkeit beeinflussen. Signifikante Prädiktoren für einen Therapieabbruch, die auf den Therapeuten zurückzuführen sind, können als Hinweise auf Therapeutenunterschiede betrachtet werden, da sie zwischen den Therapeuten variieren. Die Erfahrung von Therapeuten wurde bereits in der Meta-Analyse von Swift und Greenberg (2012) als Moderator für die Dropout-Rate berichtet. In einem aktuellen Überblicksartikel wurden weitere Therapeutenfaktoren identifiziert, die die Dropout-Rate beeinflussen (Roos & Werbart, 2013). Die Autoren fanden, dass neben der Erfahrung auch das Training und die emotionale Intelligenz des Therapeuten negativ mit der Dropout-Rate assoziiert sind. Weiterhin haben sie einen Einfluss von Beziehungs-

variablen, wie u.a. der Übereinkunft über Probleme, der Allianz, der Zufriedenheit des Klienten sowie der Passung des Geschlechts und der Ethnizität von Patient und Therapeut, auf die Dropout-Rate gefunden. Im Einklang mit bisher berichteten Effekten auf das Outcome, hatte eine geringe Anzahl an Kontakten einen negativen Effekt auf den Therapieabschluss, sodass bei wenigen Kontakten die Dropout-Rate anstieg. Die Befunde der 44 Studien aus den Jahren 2000 bis 2011 sind zwar insgesamt heterogen und die Dropout-Raten und Effektstärken variieren stark, dennoch geben sie einen klaren Hinweis darauf, dass sich die Dropout-Rate der Patienten zwischen Therapeuten unterscheidet und somit ein Therapeuteneffekt zu erwarten ist.

2.5 Zusammenhänge

Der intraindividuelle Zusammenhang zwischen der Anzahl an Therapiestunden und der Verbesserung der Patienten auf diversen Outcome-Maßen wurde in dem bis heute akzeptierten *dose-effect model of change* (Howard, Kopta, Krause & Orlinsky, 1986) beschrieben. Demnach handelt es sich um einen negativ beschleunigten Dosis-Wirkungs-Zusammenhang. Mit zunehmender Sitzungsanzahl werden immer mehr Sitzungen benötigt, um dieselbe Verbesserung zu erreichen. In den frühen Sitzungen hingegen erzielt die Therapie höhere Effekte (Kopta, 2003). Daher proklamieren Barkham et al. (2006) ein *good enough level* (GEL), ab welchem die Patienten aufgrund der ausreichenden Verbesserung die Therapie beenden. Für klinisch bedeutsam belastete Patienten konnten die Autoren zeigen, dass die Raten in reliabler und klinisch signifikanter Verbesserung mit der Anzahl der Sitzungen abnahmen. In den frühen Sitzungen verbesserten sich demnach mehr Patienten als in den späteren. Weiterhin nehmen

die Autoren an, dass die nötige Verbesserung, um das GEL zu erreichen, umso geringer wird, je länger die Therapie andauert. Patienten, die sich bereits lange in Therapie befinden, sind demnach mit geringeren Verbesserungen zufrieden. Barkham et al. (2006) argumentieren, dass von Messung zu Messung immer kleiner werdende Gruppen von sich immer langsamer verbessernden Patienten analysiert wurden, da die stärker belasteten Patienten länger in der Behandlung verweilten. Für jede dieser Gruppen fanden die Autoren lineare Zusammenhänge zwischen der Behandlungsdauer und der Verbesserung, welche aggregiert die negativ beschleunigte Dosis-Wirkungs-Kurve formten. Diese Beobachtungen sind vereinbar mit den Annahmen des GEL-Modells (Barkham et al., 2006). Stulz, Lutz, Kopta, Minami und Saunders (2013) konnten hingegen zeigen, dass der Fit eines statistischen Modells besser wird, wenn es anstelle eines linearen einen loglinearen Zusammenhang zwischen Dosis und Wirkung annimmt. Die Befunde sind konsistent mit dem negativ beschleunigten Zusammenhang im *dose-effect model* von Howard et al. (1986): Die durchschnittliche Veränderungsrate nahm mit steigender Sitzungsanzahl und unabhängig von der Behandlungsdauer ab. Ebenso sind sie mit dem GEL-Modell von Barkham et al. (2006) in Einklang: Schnelle Verbesserung war mit kürzerer Therapiedauer assoziiert, da die Klienten die Therapie beenden, sobald sie ihr GEL erreicht haben. Die langsame Verbesserung in den späteren Sitzungen ist jedoch nicht mit fehlender Verbesserung gleichzusetzen, da mit der Anzahl der Sitzungen über die Patienten hinweg auch die Anzahl der reliabel verbesserten Patienten (negativ beschleunigt) anstieg.

Wird der Zusammenhang des Outcomes mit der Sitzungsanzahl über die Zeit betrachtet, sollte auch die Dropout-Rate miteinbezogen werden. Bei längsschnittlichen Untersuchungsdesigns können sich positive Assoziationen zwischen der Sitzungsanzahl und dem Outcome zeigen, die

jedoch als reine Methodenartefakte auf einen Selektionsfehler zurückzuführen sein können. Dies liegt daran, dass sich von Messzeitpunkt zu Messzeitpunkt die betrachtete Stichprobe verändert, indem selektiv Patienten durch Dropout ausscheiden. Der positive Zusammenhang könnte sich nur deshalb zeigen, weil die eingangs stark belasteten Patienten während des Studienverlaufs ausscheiden (Specka, Finkbeiner & Hagen, 2000). So reduziert sich nur die Anzahl der stark belasteten, nicht aber die der wenig oder gar nicht belasteten Patienten. Während die absolute Anzahl an nicht belasteten Patienten also gleich bleibt, steigt ihr prozentualer Anteil an der Gesamtstichprobe. Dies kann irrtümlicherweise als Verbesserung der Patienten interpretiert werden, wenn die Dropout-Rate nicht berücksichtigt wird.

Neben der Art des Zusammenhangs zwischen Sitzungsanzahl und Outcome liegen zu dessen Stärke bisher eher heterogene Befunde vor (Goldenberg, 2002; Howard et al., 1986; Miller & Berman, 1983). So berichtet Goldenberg (2002), dass in einer kombinierten Stichprobe das Outcome das stärkste Korrelat der Sitzungsanzahl war. Es verlor jedoch seine Position als *single best predictor*, wenn die Stichprobe in Patienten, die abgebrochen haben, (Dropouts) und solche, die die Therapie regulär abgeschlossen haben, (Completer) unterteilt wurde. Unter den Completern hatte der Therapeut (als kategoriale Variable mit vier Ausprägungen) die stärkste Korrelation mit der Sitzungsanzahl. Unter den Dropouts trat kein Prädiktor eindeutig hervor. Es scheinen viele Einflussfaktoren zu existieren, die diesen Zusammenhang moderieren, darunter Klienten-, Therapeuten- und Kontextfaktoren (Steenbarger, 1994). In seinem Review-Artikel fasst Steenbarger (1994) zusammen, dass der Dosis-Wirkungs-Zusammenhang davon abhängt, was gemessen wird (Outcome-Maß), wann gemessen wird und auf wie viele Sitzungen die Therapie ausgelegt ist sowie davon,

wer das Outcome einschätzt (Patient, Therapeut oder ein Dritter). Als weitere Einflussfaktoren nennt er u.a. die Diagnose und die Erwartung des Patienten bezüglich der Therapiedauer, den Einfluss, die Dominanz und die Fähigkeit des Therapeuten, eine Allianz aufzubauen, sowie die Patienten-Therapeuten-Passung und die Ereignisse nach der Therapie. Daraus schlussfolgert der Autor, dass es nicht die eine Funktion gibt, um Therapiedauer und Outcome in Beziehung zu setzen, sondern diese von der aktuellen Phase der Veränderung abhängt. Methodisch müssen zudem Zusammenhänge zwischen und innerhalb von Patienten differenziert werden. Intraindividuell zeigen die zuvor berichteten Befunde einen positiven Zusammenhang zwischen Outcome und Sitzungsanzahl, sodass sich das Outcome eines Patienten verbessert, wenn er mehr Sitzungen wahrnimmt. Auf interindividuelle Zusammenhänge können diese Befunde jedoch nicht generalisiert werden, da zum Beispiel stark belastete Patienten zum einen ein schlechteres Outcome haben als weniger stark belastete und zum anderen mehr Sitzungen benötigen. In diesem Fall würde sich ein negativer Zusammenhang ergeben.

Um eine Erklärung für die unterschiedlichen Zusammenhänge zwischen Sitzungsanzahl und Outcome zu finden, haben Mueller und Pekarik (2000) die erwartete Sitzungsanzahl zu Beginn der Therapie miterhoben. Sie fanden, dass die erwartete und die tatsächliche Sitzungsanzahl deutlich differierten, was sich vor allem bei wenigen Sitzungen zeigte: Von den 230 Patienten erwarteten nur 3.0%, dass die Therapie nur eine oder zwei Sitzungen dauert, wohingegen 13.5% tatsächlich maximal zwei Sitzungen wahrgenommen haben. Eine bis fünf Sitzungen erwarteten immerhin 18.8% der Patienten, wobei die Therapie bei 36.6% nicht mehr als fünf Sitzungen dauerte. Auch die Therapeuten erwarteten längere Therapien.

Die Anzahl der Sitzungen konnte in einer schrittweisen multiplen Regression signifikant u.a. durch die von den Patienten erwartete Anzahl an Sitzungen, den Abschluss des Therapeuten und die Bildung des Patienten vorhergesagt werden (Mueller & Pekarik, 2000). Die Patienten, die früher die Therapie beendeten als erwartet, zeigten auf zwei von drei Outcome-Maßen größere Verbesserungen. Allerdings berichteten die Patienten, deren Sitzungsanzahl in dem erwarteten Bereich lag, eine höhere Zufriedenheit. Die Vorhersage der Dauer durch den Patienten korrelierte in dieser Studie mit der tatsächlichen Anzahl an Sitzungen und die Präzision dieser Vorhersage hing u.a. mit dem Outcome und der Zufriedenheit der Patienten zusammen.

In den bisher dargestellten Untersuchungen wurde die Sitzungsanzahl als Operationalisierung der Therapiedauer herangezogen, jedoch kann diese keine Informationen über die vergangene Zeit während der Therapie liefern. Deshalb sollte der Effekt von Therapiedauer und Sitzungsanzahl analysiert werden, um eine differenziertere Aussage über den Zusammenhang von Dosis und Wirkung treffen zu können. Bei 74 erwachsenen Ambulanzpatienten war das Outcome nicht über die Dauer oder die Sitzungsanzahl vorhersagbar, sehr wohl aber über deren Interaktion (Reardon, Cukrowicz, Reeves & Joiner, 2002). Die Interaktion klärte inkrementelle 6.2% der Varianz im Outcome auf. Bei Patienten mit wenigen Sitzungen (≤ 11) war eine längere Dauer mit schlechterem Outcome assoziiert, während für Patienten mit mehr Sitzungen kein Zusammenhang zwischen der zeitlichen Dauer und dem Outcome bestand. Lange Therapiezeiträume sind demnach destruktiv, wenn wenige Sitzungen stattfinden. Um die Dichte der Sitzungen zu messen, haben die Autoren die Sitzungen pro Woche berechnet. Es zeigte sich, dass die Dichte bei Patienten mit kürzeren Therapien höher war

und lange Therapien oft mit einer geringen Dichte einhergingen. Bei Patienten mit kurzen Therapien (≤ 5.5 Monate) war eine höhere Dichte zudem mit einem besseren Outcome assoziiert, während bei längeren Therapien eine höhere Dichte mit einem schlechteren Outcome in Beziehung stand. Dies erklären die Autoren darüber, dass Patienten mit einer starken Ausgangsbelastung länger in Therapie bleiben als weniger belastete Patienten (Reardon et al., 2002). Die Befunde liefern trotz einiger Einschränkungen – dadurch, dass sich mehr Patienten in der Untersuchung verschlechtert als verbessert haben – interessante Einsichten in die differentiellen Zusammenhänge des Outcomes mit der Sitzungsanzahl und der Therapiedauer.

2.6 Frequentistische und Bayes'sche Statistik

Viele Studien zu Therapeuteneffekten berichten nur die Effektstärken, da aufgrund der deutlich zu geringen Stichprobenumfänge für Multilevel-Modelle (vgl. Abschn. 3.4.2) und der daraus resultierenden geringen Teststärke auch bedeutsame Effekte nicht signifikant werden. Dennoch ist die Signifikanz eines Ergebnisses nach wie vor das zentrale Thema in der angewandten Statistik. Ein bisher wenig beachtetes Problem bei der Interpretation von Signifikanztests sind die weder reliablen noch objektiven p -Werte, die die Wahrscheinlichkeit für ein Ereignis unter Annahme der Nullhypothese ausdrücken sollen. Die Einschränkungen dieser Methode wurden von Nuzzo (2014) darauf zurückgeführt, dass es sich um ein Hybridsystem aus zwei unterschiedlichen statistischen Entwicklungen handelt, welches nicht im Sinne der Autoren angewendet wird. Dabei wurde der von Fisher (1922) eingeführte p -Wert, der keinen definitiven Test darstellt, mit dem Regelsystem von Neyman und Pearson (1967) ver-

bunden, obwohl der p -Wert auf diese Weise nie benutzt werden sollte. Fishers (1922) p -Wert sollte lediglich die Entscheidung vereinfachen, ob eine Hypothese einen zweiten Blick wert ist. So bedeutet ein empirischer Wert von $p = .01$ nicht, dass das gefundene Ergebnis nur mit einer einprozentigen Wahrscheinlichkeit falsch positiv ist; auch wenn die gängige Methodenlehre uns das glauben lässt. Der p -Wert kann nicht rückwärts arbeiten und Aussagen über die zugrundeliegende Realität machen (Nuzzo, 2014), denn dazu fehlt die Information über die Basiswahrscheinlichkeit, dass ein realer Effekt vorliegt. Da diese nicht in dem p -Wert enthalten ist, bleibt er invariant gegenüber Veränderungen der Wahrscheinlichkeit: Bei gleichbleibendem p -Wert wird ein Ergebnis immer unwahrscheinlicher, je weniger plausibel die Hypothese ist. Ein weiteres Problem ist die Überzeugungskraft eines p -Wertes, der von dem relevanteren Effekt ablenkt. Die praktische Relevanz eines Ergebnisses enthält aber viel wichtigere Informationen als ein p -Wert, weshalb Cumming (2013) zum Berichten von Effektstärken und Konfidenzintervallen rät.

Um Informationen über die Wahrscheinlichkeit eines wahren Effektes berücksichtigen zu können, kann der p -Wert durch Methoden ersetzt werden, die Erweiterungen der Bayes' Regel darstellen. Es handelt sich um ein Theorem, das beschreibt, wie man Wahrscheinlichkeit als die Plausibilität eines Ergebnisses sehen kann, anstatt diese als Häufigkeit eines Ergebnisses zu betrachten. Diese Methoden berechnen, wie sich Wahrscheinlichkeiten verändern, wenn neue Daten hinzukommen (Nuzzo, 2014). Man unterscheidet also zwischen statistischen Ansätzen, die Wahrscheinlichkeit als relative Häufigkeit definieren, (frequentistische Ansätze) und Ansätzen, die auf dem Bayes' Theorem aufbauen (Bayes'sche Ansätze). Die beiden Ansätze werden bei Li, Lingsma, Steyerberg und Lesaffre (2011) zusammengefasst vorgestellt. In der frequentistischen Statistik ist die

Wahrscheinlichkeit eines Ereignisses demnach das Limit der relativen Häufigkeit dieses Ereignisses in einer großen Anzahl an Studien. Die Ergebnisse werden so interpretiert, als hätte man den Versuch unendlich oft durchgeführt. Man schätzt den unbekannten aber festen Parameter θ (Theta) durch die maximale Erhöhung der Wahrscheinlichkeit (Maximum Likelihood). Auf der Grundlage des geschätzten θ wird die Vorhersage geleistet, deren Unsicherheit nur auf den Sampling-Eigenschaften des geschätzten Wertes von θ basiert (Feller, 1957).

Im Vergleich dazu kombiniert der Bayes'sche Ansatz eine a priori Wahrscheinlichkeit mit den tatsächlich beobachteten Daten, um dadurch eine aktualisierte a posteriori Wahrscheinlichkeit zu erhalten (Bernardo & Smith, 1994). Der Parameter θ bekommt eine Wahrscheinlichkeitsverteilung, die das a priori Wissen über diesen Parameter widerspiegelt. Zur Berechnung müssen hochdimensionale Integrale gebildet werden, was heute über Markov Chain Monte Carlo (MCMC) Methoden geleistet werden kann (Gelfand & Smith, 1990). MCMC Methoden approximieren das gesuchte Integral über ein Monte Carlo Sampling (Gibbs Sampling und Metropolis-Hastings Updating). Eine genauere Darstellung dieser Sampling-Methoden ist hier nicht vorgesehen, der interessierte Leser sei aber auf Hadfield (2014a) verwiesen. Die a posteriori Schätzungen basieren hierbei auf der a priori Verteilung der Parameter und der Wahrscheinlichkeit. Dadurch variieren die Schätzungen bei gleichen Daten und gleicher Wahrscheinlichkeit, wenn unterschiedlich informative a priori Werte benutzt werden. Sind die a priori Werte nicht informativ, gleichen die a posteriori Schätzungen den Ergebnissen der Maximum Likelihood (ML) Methoden (Li et al., 2011). Über diese a priori Parameter können die in der frequentistischen Statistik fehlenden Informationen über die Wahrscheinlichkeit, dass ein realer Effekt vorhanden ist, in die statistische Testung einfließen.

Um die Unterschiede der beiden Ansätze zu verdeutlichen, sollen die geschätzten Wahrscheinlichkeiten einander gegenübergestellt werden. So ist die Likelihood der frequentistischen Ansätze die Wahrscheinlichkeit der Daten, gegeben die Modellparameter μ (Mittelwert in der Population) und σ^2 (Varianz in der Population):

$$\Pr(y|\mu, \sigma^2) \quad (2.1)$$

Die Modellparameter werden als fest und bekannt angenommen. Das Problem dabei ist, dass die Daten beobachtet wurden und die Parameterwerte nicht bekannt sind (Hadfield, 2014a). Bei Bayes'schen Ansätzen hingegen wird die Wahrscheinlichkeit der Modellparameter, gegeben die beobachteten Daten, geschätzt:

$$\Pr(\mu, \sigma^2|y) \quad (2.2)$$

Wenn man diese Wahrscheinlichkeit umformt, erhält man folgende proportionale Darstellung:

$$\Pr(y|\mu, \sigma^2) \Pr(\mu, \sigma^2) \quad (2.3)$$

Der erste Teil stellt nun die Likelihood dar, der zweite Teil die a priori Annahmen über die Werte, die die Modellparameter annehmen können (Hadfield, 2014a). Das größte Problem in der Anwendung dieser sicherlich vorteilhaften Methode sind die a priori Annahmen an sich. In Abhängigkeit davon, was als a priori Verteilung in die Analyse eingeht, verändern sich die a posteriori Schätzungen deutlich. Dies erfüllt zwar genau

die gewünschten Voraussetzungen, erfordert aber auch valide Werte für die a priori Verteilungen, welche häufig nicht bekannt sind. Um aber einen konzeptuellen Mehrwert über die frequentistischen Ansätze hinaus zu erhalten, müssen informative a priori Verteilungen vorgegeben werden.

2.7 Zusammenfassung und Herleitung der Hypothesen

Das Ziel der vorliegenden Untersuchung ist es, die Therapeuteneffekte auf das Outcome, die Sitzungsanzahl und die Dropout-Rate zu quantifizieren. Betrachtet werden primär die Effektstärken (Cumming, 2013). Bedeutsame Therapeuteneffekte auf die drei Kriterien würden eine Differenzierbarkeit von effektiven und weniger effektiven Therapeuten belegen. Die drei Merkmale können dann als Maße für den Erfolg der therapeutischen Arbeit nutzbar gemacht werden. Weiterhin soll der Zusammenhang zwischen den Therapeuteneffekten untersucht werden. Bei einem nur geringen Zusammenhang sind die Indikatoren nicht konfundiert und somit zusätzlich zu den jeweils anderen Variablen nützlich, um effektive Therapeuten zu identifizieren.

Die Schätzung des Therapeuteneffekts auf Outcome stellt eine Replikation der in Abschnitt 2.2.2 zusammengefassten Ergebnisse dar und unterstützt aufgrund der leicht abweichenden Rahmenbedingungen die Generalisierbarkeit der bisherigen Befunde. Um die Schätzung des Therapeuteneffekts auf die Sitzungsanzahl, der erst einmal berichtet wurde (Lutz et al., im Druck), und des Therapeuteneffekts auf die Dropout-Rate, der zuvor noch nie publiziert wurde, erweitert die vorliegende Studie den Forschungsstand. Auch die Zusammenhänge zwischen den Therapeuteneffekten wurden zuvor nur von Lutz et al. (im Druck) berichtet. Nach der Studie von Saxon und Barkham (2012) ist dies zudem die erste Untersuchung, die

zur Schätzung der Therapeuteneffekte – und in der Psychotherapieforschung überhaupt – MCMC Methoden einsetzt. Der vierte Punkt, der zum Neuigkeitswert dieser Untersuchung beiträgt, ist die Anwendung eines multivariaten Auswertungsdesigns, das drei Kriterien vereint, die unterschiedlichen Verteilungsformen folgen.

Bezüglich des Outcomes der Patienten finden sich bei sehr heterogenen Befunden im Mittel geringe bis moderate Therapeuteneffekte um 7% bei naturalistischen Studien (Baldwin & Imel, 2013). Gerade bei weniger stark belasteten Patienten ist von kleinen Effekten auszugehen (Saxon & Barkham, 2012). Es existieren bisher kaum Befunde zu Therapeuteneffekten auf die Sitzungsanzahl (Okiishi et al., 2006), allerdings gibt eine erste Studie Hinweise darauf, dass die Effekte mit 9% größer sein könnten als die auf Outcome (Lutz et al., im Druck). Für die Dropout-Rate wurden bisher nur Therapeutenvariablen, nicht aber Therapeuteneffekte untersucht (Roos & Werbart, 2013). Die gefundenen signifikanten Zusammenhänge zwischen Therapeutenvariablen und der Dropout-Rate geben aber Grund zu der Annahme, dass auch diese Variable systematisch zwischen den Therapeuten variiert. Es wird ein ähnlicher Effekt wie auf das Outcome erwartet. Die erste Hypothese (H1) lautet daher:

H1: Es existieren geringe bis moderate Therapeuteneffekte auf den Variablen Outcome, Sitzungsanzahl und Dropout. Die Variabilität der Therapeuten in der mittleren Sitzungsanzahl ist dabei größer als die im mittleren Outcome.

H1a: Der Therapeuteneffekt auf Outcome beträgt 7%.

H1b: Der Therapeuteneffekt auf die Sitzungsanzahl beträgt 9%.

H1c: Der Therapeuteneffekt auf Dropout beträgt 7%.

Darüber hinaus sollen neben den klassischen Maximum Likelihood Schätzungen auch MCMC Methoden eingesetzt werden, bei denen präzisere Schätzungen erwartet werden (Nuzzo, 2014). Diese bieten darüber hinaus die Möglichkeit, die drei interessierenden Variablen in einem multivariaten Modell simultan auszuwerten. Diese Modellierung bringt u.a. den Vorteil mit sich, dass die drei potentiell zusammenhängenden Variablen in ihren Residuen kovariieren dürfen, wodurch die Schätzungen der Varianzen weiter optimiert werden. Die Hypothesen werden daher sowohl mit ML als auch mit MCMC Schätzmethoden überprüft. Für die Modelle mit Parameterschätzungen über die MCMC Methode werden von den frequentistischen Modellen abweichende Schätzungen der wahren Werte erwartet; mit der größten Differenz zwischen den Schätzungen der frequentistischen univariaten Modelle einerseits und denen des Bayes'schen multivariaten Modells andererseits.

Das multivariate Modell erleichtert zudem die Betrachtung der Zusammenhänge zwischen den Therapeuteneffekten. Die zweite Forschungsfrage dieser Arbeit fokussiert die Abhängigkeit der drei Therapeuteneffekte. Der Zusammenhang zwischen der Sitzungsanzahl eines Patienten und seinem Outcome ist intraindividuell positiv ausgeprägt (Stulz et al., 2013), kann allerdings interindividuell verschiedene Richtungen aufweisen. Wenn für die Ausgangsbelastung eines Patienten kontrolliert wird, sollten die Variablen positiv zusammenhängen (Howard et al., 1986), vor

allem bei einer kombinierten Stichprobe aus Abbrüchen und regulären Abschlüssen (Goldenberg, 2002). Die vielen Einflussfaktoren (Steenbarger, 1994) machen eine Vorhersage jedoch schwierig. Geht man davon aus, dass Patienten mit größerer Sitzungsanzahl ein besseres Outcome erreichen, wenn für ihre Prä-Werte kontrolliert wird, kann ein positiver Zusammenhang zwischen den Therapeuteneffekten auf Sitzungsanzahl und Outcome erwartet werden. Der Zusammenhang der Therapeuteneffekte auf Sitzungsanzahl und Dropout sollte negativ ausfallen, da Therapeuten kürzere Therapien mit einer höheren Wahrscheinlichkeit als Abbrüche bewerten. Diesen Zusammenhang konnten Bados et al. (2007) bereits bei Patienten finden. Der Therapeuteneffekt auf die Dropout-Rate sollte zudem negativ mit dem auf das Outcome assoziiert sein, da sich Patienten, die die Therapie abbrechen, im Durchschnitt weniger verbessern sollten. Somit erreichen Therapeuten, die mehr ihrer Patienten als Dropouts klassifizieren, ein schlechteres mittleres Outcome. Die Zusammenhänge sollten jedoch relativ gering ausfallen, weshalb alle drei Merkmale zur Identifikation überdurchschnittlich effektiver Therapeuten herangezogen werden müssen. Die zweite Hypothese (H2) lautet zusammenfassend:

H2: Der Therapeuteneffekt auf Outcome kovariiert positiv mit dem auf die Sitzungsanzahl und negativ mit dem auf die Dropout-Rate. Die Effekte auf Sitzungsanzahl und Dropout kovariieren ebenfalls negativ.

Therapeuteneffekte auf Outcome, Sitzungsanzahl und
Dropout

Multivariate Multilevel-Analyse mit
Markov-Chain-Monte-Carlo-Schätzung
Schwartz, B.

2017, XVII, 124 S. 10 Abb., Softcover

ISBN: 978-3-658-16471-3