

BEDEUTUNG, METHODIK UND ORGANISATION VON FOLLOW-UP-STUDIEN

Follow-up-studies: importance, methodology, organization

Stefan Klingberg, Gerhard Buchkremer

Zusammenfassung

Psychische Störungen haben oft eine mehrjährige Vorgeschichte, bilden sich nicht immer bis Therapieende vollständig zurück und neigen zu Rezidiven. Daher sollten Prädiktionsanalysen zum Verlauf psychischer Störungen sowie Studien zur Wirksamkeit von Interventionen mittel- bis langfristige Zeiträume für die Verlaufsbeurteilung berücksichtigen. Dieser Forderung wird nur eine begrenzte Zahl an Studien gerecht. Bei der Durchführung solcher Follow-up-Studien sind verschiedene methodologische Schwierigkeiten zu beachten, v.a. die Wahl des angemessenen Zeitraums, die Art der Hypothesenprüfung sowie der Umgang mit fehlenden Daten. Mögliche Problemlösungen zu diesen Themen werden diskutiert. Darüber hinaus werden organisatorische Fragen erörtert, die für den Erfolg von Follow-up-Studien ebenfalls eine wichtige Rolle spielen. Follow-up-Studien sind relativ aufwändig und bedürfen einer substanziellen finanziellen Unterstützung. Hierzu wird eine Beispielrechnung vorgelegt. Abschließend werden ethische Fragen kurz diskutiert.

Schlüsselwörter Follow-up-Studien – Methodologie – Therapieforschung - Drop-out-Problem
Follow-up studies: importance, methodology, organization

1. Einleitung

Die Wirksamkeit von Interventionen bei Menschen mit psychischen Störungen muss hinsichtlich mittel- bis langfristiger Zeiträume beurteilt werden. Psychische Störungen, wie z.B. Substanzabhängigkeit, schizophrene Psychosen, affektive Störungen, Angststörungen, beruhen auf langfristigen Entwicklungen, sind potenziell durch längerfristig bestehende Symptome und funktionelle Einschränkungen gekennzeichnet (Keller, 1999; Mancuso et al., 1993; Huber et al., 1980). Sie entwickeln sich teilweise über mehrere Jahre (z.B. Schizophrenien: vgl. Häfner et al., 1999). Auch nach Beginn der Therapie können Einschränkungen fortbestehen. Daher werden in Bezug auf viele Störungen Studien über mittel- bis langfristige Verläufe gefordert (Keller und Boland, 1998; Stapleton et al., 1998; Versiani et al., 1997; Weiden et al., 1996; Schweizer und Rickels, 1996).

Es ist das Ziel von Follow-up-Studien, den langfristigen Verlauf psychischer Störungen zu berücksichtigen. Ihre prospek-

Summary

Mental disorders are characterized by a long-term course. At the end of a therapy there are remaining symptoms in many cases. In addition, relapse is not uncommon. Predictor analyses as well as clinical trials should therefore observe the course of the illness on the medium or long term. However, the number of studies conducting such long-term Follow-up is limited. There are several methodological problems, such as the choice of an adequate Follow-up time, the kind of hypothesis testing or the options in dealing with missing values. We discuss some solutions according to these problems. Furthermore we discuss organizational questions that are of same importance for the success of a study. Follow-up studies are expensive and time consuming. They need substantial financial support. We give an example for a financial calculation. Finally we discuss some ethical questions.

Keywords Follow-up studies - methodology - therapy research - Drop-out-problem

tive Anlage erlaubt die Bewertung von Therapien im Rahmen von Interventionsstudien und die Suche nach Verlaufsprädiktoren in Prognose- oder Prädiktorstudien.

Obwohl die Notwendigkeit von Follow-up-Studien unbestritten ist, bestehen gerade auf diesem Gebiet große Defizite. Dies gilt nicht nur für psychotherapeutische, sondern auch für pharmakologische Interventionen (Thase, 1999). Bei der Einführung neuer Medikamente wird vor allem ein kurzfristiger Wirksamkeitsnachweis gefordert, Langzeitdaten werden hingegen sehr viel seltener erhoben und publiziert. Besonders deutlich kommt dies im folgenden Statement von Kazdin (1994) zum Ausdruck: „In current therapy research, evidence indicates that the inclusion of Follow-up assessment in the study is the exception rather than the rule.“ Der wesentliche Grund für diese Situation ist in erster Linie der erhebliche Aufwand, der mit Follow-up-Studien verbunden ist. Ohne eine

substanzielle Förderung solcher Studien durch Drittmittelgeber ist der erforderliche Aufwand nur schwer zu leisten.

Im Folgenden werden Ziele und Fragestellungen von Follow-up-Studien beschrieben. Dabei wird insbesondere auf Fragen der Therapieevaluation und auf einige Aspekte von Prädiktionsanalysen fokussiert. Es folgen methodische Überlegungen zu Designfragen, zur Behandlung von Drop-outs und zur Studienorganisation. Nach einer Beispielrechnung zu Kosten von Follow-up-Studien wird abschließend auf ethische Fragen eingegangen.

2. Ziele von Follow-up-Studien

Follow-up-Studien prüfen, ob ein mit der Krankheit in Verbindung stehendes Merkmal, das zu Studienbeginn erfasst wurde, bei einem späteren Zeitpunkt (nach der Therapie oder bei weiteren Follow-up-Messungen) qualitativ oder quantitativ verändert ist. Diese längsschnittliche Betrachtungsweise ist vor allem für zwei Gruppen von Fragestellungen relevant: für Therapieevaluation und für Prädiktionsanalysen.

2.1 Therapieevaluation

Kazdin und Kendall (1998) haben 7 grundlegende Schritte der Therapieforschung formuliert:

1. Konzeptualisierung der Dysfunktion
2. Forschung zum Verlauf der Dysfunktion
3. Konzeptualisierung der Behandlung
4. Spezifikation der Behandlung
5. Prüfungen des Therapieergebnisses
6. Therapieprozessforschung
7. Tests von Randbedingungen und Moderatorvariablen

Demnach vollzieht sich die Entwicklung von Therapieverfahren in verschiedenen Schritten. Relativ am Beginn der Therapieentwicklung steht Forschung zum Verlauf der Dysfunktion. Erforderlich ist zunächst ein sehr detailliertes Wissen über die Dynamik der Krankheitsentstehung, die hierbei relevanten Variablen und ihre zeitliche Entwicklung. Dies erfordert Follow-up-Studien. Konzeption und Ausarbeitung von Interventionen sollten sich hieran anschließen. Follow-up-Studien sind ebenfalls unter dem Punkt „Prüfung des Therapieergebnisses“ durchzuführen.

Aus der Methodologie zur Pharmaka-Prüfung ist eine Unterteilung in Forschungsphasen bekannt (vgl. z.B. Spiegel, 1988). Als Phase III wird die erweiterte klinische Prüfung bezeichnet, die sich methodisch vor allem durch die Durchführung von randomisierten klinischen Studien kennzeichnet. Unter der Phase III werden ausdrücklich auch Langzeitstudien genannt. Die Wirksamkeit von Therapieverfahren muss auch über längere Zeiträume nachgewiesen werden. Vergleichbares gilt erst recht für Phase-IV-Studien zur Prüfung des Therapieerfolgs unter Praxisbedingungen. Der Langzeitaspekt ist hier nicht wegzudenken.

Therapieerfolg oder Wirksamkeit sind vielschichtige Begriffe. Baumann (1997) hat darauf differenziert hingewiesen und

die Aspekte Multimodalität, Abstraktionsgrad, Vergleichsmaßstab und Zeitabhängigkeit herausgearbeitet, auf die hier in Kürze eingegangen werden soll.

Bei einer Therapiestudie kann es nicht *die* Wirksamkeit einer Intervention geben, sondern nur eine Wirksamkeit in bestimmter Hinsicht. Die Wirksamkeit insbesondere von Psychotherapie muss multimodal unter Einbeziehung aller relevanten diagnostischen Ebenen (subjektiv, objektiv, apparativ) und Datenquellen (Patienten, Therapeuten, Angehörige) erfasst werden. Es ist nicht ausreichend, z.B. nur den Wegfall eines Symptoms zu erheben. Es ist möglich, dass durch den Wegfall eines Symptoms auf anderen Ebenen, wie z.B. Lebensqualität, eine Verschlechterung eintritt.

Der Abstraktionsgrad von Wirksamkeitsaussagen kann stark variieren. Prädiktoren eines Krankheitsverlaufes oder eines Therapieerfolgs können operational mit biologischen oder verhaltensnahen Parametern über die Symptom- und Syndromebene bis hin zu hypothetischen Konstrukten (z.B. „Narzissmus“, „Bewältigungsstil“ etc.) erfasst werden.

Wenn die Wirksamkeit einer Therapiemaßnahme bestimmt werden soll, ist ein Vergleichsmaßstab erforderlich. Hierfür können je nach Fragestellung unterschiedlich geartete Kontrollgruppen herangezogen werden: Unbehandelte Kontrollgruppen, Warte-Kontrollgruppen, Placebo-Kontrollgruppen, unspezifisch behandelte Kontrollgruppen, Kontrollgruppen mit konkurrierender Therapie.

Die Zeitabhängigkeit der Wirksamkeit ist in Bezug auf das Thema dieser Arbeit der wichtigste Aspekt. Die Wirksamkeit kann zu unterschiedlichen Zeiten und im Verlaufe des Krankheitsgeschehens differieren. Aus der Schizophreniebehandlung gibt es z.B. Hinweise, dass psychotherapeutische Verfahren zur Rückfallverhütung im ersten Jahr nach der Klinikentlassung noch nicht wesentlich zur Senkung der Rückfallrate beitragen, jedoch im späteren Zeitverlauf ihre Wirkung entfalten (Hornung et al., 1999). Ebenso ist denkbar, dass psychopathologische Symptome im Therapieverlauf erst allmählich zurückgehen, dann über einen bestimmten Zeitraum stabil bleiben und schließlich in der Intensität wieder zunehmen. Verschiedene Therapien können dabei ganz unterschiedliche zeitliche Verläufe in Bezug auf unterschiedliche Symptombdimensionen hervorbringen. Weitere Beispiele finden sich bei Kazdin (1994), die zeigen, dass nicht ungeprüft davon ausgegangen werden kann, dass Therapieeffekte, gleich welcher Art, über die Zeit stabil sind.

Zentrale Forderungen an Therapieevaluationen unter dem Blickwinkel des Themas dieser Arbeit sind daher:

1. Erfassung von störungsangemessenen, langen Follow-up-Zeiträumen: Aus den einleitend dargestellten Überlegungen ergibt sich, dass die Zeitperspektive für Wirksamkeit oft angemessener in Jahren als in Monaten oder gar Wochen anzugeben ist.

2. Mehrdimensionale Diagnostik: Daten verschiedener Quellen (Patient, Angehörige, Beurteiler) und Ebenen (objektiv, subjektiv, Rating) müssen gesammelt werden. Diese Forderung gilt auch und gerade für Langzeitverläufe.
3. Komplexe Wirksamkeitsaussagen: Legt man einen längeren Follow-up-Zeitraum zugrunde und verfolgt die beschriebene Strategie der mehrdimensionalen Diagnostik, folgt daraus, dass Therapiestudien immer komplexe Ergebnisse aufweisen. Aussagen zur Wirksamkeit können nur in dieser Komplexität angemessen sein.

Der dritte Punkt steht allerdings in Widerspruch zu einer zentralen statistischen Forderung für konfirmative Wirksamkeitsuntersuchungen, nämlich der Begrenzung auf sehr wenige Haupthypothesen bzw. der Begrenzung der Zahl durchzuführender statistischer Tests. Der wesentliche Grund hierfür ist die Inflationierung des Alpha-Fehlers bei einer hohen Zahl von statistischen Tests an der gleichen Stichprobe (Bortz und Döring, 1995). Dem kann u.a. dadurch begegnet werden, dass aus einer umfangreichen Untersuchungsbatterie a priori festgelegt wird, welche Variable und welcher Test zur Prüfung der Haupthypothese vorgesehen sind. Die weiteren Tests sind entsprechend als explorativ anzusehen und so zu interpretieren. Der beschriebenen Komplexität von Wirksamkeitsaussagen kann als Konsequenz daraus nur aufgrund von mehreren Studien adäquat Rechnung getragen werden. Dies bedeutet jedoch in der Praxis eine enorme Erhöhung des notwendigen Aufwandes.

Ein möglicher Nebeneffekt von Follow-up-Studien ist, dass durch die regelmäßigen Kontakte mit den Untersuchern die Studienteilnehmer oft bessere Therapieergebnisse aufweisen, als sie unter „naturalistischen“ Bedingungen zu erwarten wären. Die Patienten erfahren eine zumeist intensivere Betreuung als allgemein üblich. U. a. hat dies die Konsequenz, dass auf Krisen schneller reagiert werden kann. Zudem repräsentieren die Studienstichproben häufig eine systematisch selektierte Teilpopulation. Die Behauptung, so gewonnene Stichproben seien weniger schwer erkrankt, ist aber schwierig zu belegen. Entsprechende Gegenbeispiele finden sich bei Carroll et al. (1999) sowie Amori und Lenox (1989). Als Konsequenz ist zu fordern, dass der Selektionsprozess in klinischen Studien in allen Phasen (Screening, Rekrutierung, Randomisierung, Follow-up in unterschiedlichen Zeiträumen) zu dokumentieren und hinsichtlich der Selektionsfaktoren kritisch zu prüfen ist. Dies ist für die Beurteilung der Generalisierbarkeit der resultierenden Wirksamkeitsaussagen entscheidend.

2.3 Prädiktionsanalysen

In Prädiktionsanalysen wird nach Variablen gesucht, die entweder den Krankheitsverlauf allgemein unter naturalistischen Bedingungen oder den Krankheitsverlauf unter spezifischen Behandlungsbedingungen oder den Therapieerfolg einer Therapie vorhersagen.

Grundsätzlich sind solche Studien prospektiv anzulegen, da nur in diesem Design hypothesenprüfende Prädiktions-

analysen möglich sind. Prospektiv im engeren Sinne bedeutet, dass die Fragestellung und der vollständige Untersuchungsplan vor Beginn der Untersuchung festliegen. In einem weiteren, weniger eng gefassten Sinne werden Studien generell dann als prospektiv angesehen, wenn von bestimmten Eingangsbedingungen (z.B. Therapieintervention) ausgehend, im Verlauf einer Zeitachse auf die Outcome-Kriterien (z.B. Therapieeffekt, Rückfallhäufigkeit etc.) geschaut wird, auch wenn dies nicht a priori geplant und vorgesehen war. Die Prüfung von expliziten Zusammenhangshypothesen ist so allerdings nicht möglich.

Bei der retrospektiven Sicht wird von den Wirkkriterien auf der Zeitachse rückwärts auf Merkmale der auslösenden Bedingungen gesehen. In entsprechenden Analysen können deshalb oft die Daten nicht in der erforderlichen Weise (z.B. aus Krankenakten o.Ä.) rekonstruiert werden. Noch wichtiger ist der Sachverhalt, dass retrospektiv eine intentionsgeleitete Analyse der Daten nicht mehr sicher vermeidbar ist.

Wie auch bei der Therapieevaluation ist bei Prädiktionsanalysen der Zeitraum der Prädiktion ein wesentlicher Faktor. Oft sind Variablen für einen Zeitraum prädiktiv, für einen anderen dagegen nicht. Daraus folgt, dass der Prädiktionszeitraum aus der Fragestellung abzuleiten und zu begründen ist.

Wie bei der Therapieevaluation ergibt sich auch hier durch die Vielzahl der Variablen und resultierenden Tests das Problem des multiplen Testens. Gerade Prädiktionsanalysen verleiten durch die in der Regel hohe Zahl verfügbarer Variablen zu datengeleiteten Post-hoc-Analysen mit unkontrollierter Erhöhung der Irrtumswahrscheinlichkeit. Dies kann in nahezu beliebigen, nicht aussagekräftigen Ergebnissen resultieren. Eine der überzeugendsten Lösungen ist die Durchführung von hypothesenprüfenden Prädiktionsanalysen, in denen am Beginn einer Follow-up-Untersuchung festgelegt wird, welche sehr begrenzte Zahl an Prädiktoren, die für die Fragestellung relevant sind, erfasst und im weiteren Verlauf überprüft werden. Wird dieser Untersuchungsplan seriös verfolgt, so schützt dies vor der retrospektiven und letztlich willkürlichen Suche nach signifikanten Korrelationen in großen Datensätzen.

Zusammenfassend können folgende wesentliche Forderungen für Prädiktionsanalysen genannt werden:

1. Die Prädiktionszeiträume und die verschiedenen diagnostischen Verfahren sind in Bezug auf die Fragestellung zu begründen.
2. Wann immer möglich, sollten hypothesenprüfende Prädiktionsanalysen durchgeführt werden mit a priori formulierten Zusammenhangshypothesen, die potenziell auch nicht bestätigt werden können.
3. Nach Möglichkeit sollten bei Follow-up-Studien Messverfahren eingesetzt werden, die eingeführt sind, hinsichtlich ihrer Gütekriterien geprüft sind und einen Vergleich mit analogen nationalen und internationalen Studien ermöglichen.

Methodik von Follow-up-Studien

Design-Fragen

Designs von Follow-up-Studien kennzeichnen sich durch folgende Bestimmungsstücke:

1. Untersuchungszeitraum: Die Untersuchung wird für einen festgelegten Zeitraum, z.B. 1, 2, 5 oder 10 Jahre durchgeführt.
2. Untersuchungsfrequenz: Ebenso wird die Untersuchungsfrequenz von vornherein festgelegt, z.B. monatlich, vierteljährlich oder in größeren Intervallen.
3. Untersuchungsmerkmale: Die zu untersuchenden Merkmale werden zu Beginn der Untersuchung festgelegt.
4. Fragestellungen: Die mit den aufgeführten Merkmalen zu beantwortenden Fragestellungen werden a priori formuliert. Die Fragestellungen können jedoch unterschiedlich sein und entstammen dem Bereich der Therapieevaluation, der Prädiktion sowie der Epidemiologie.

Der prospektive Charakter der Studie ergibt sich daraus, dass alle genannten Merkmale vor Beginn der Untersuchung festgelegt werden. Zur Sicherung des prospektiven Charakters sollte die Fragestellung publik gemacht werden.

2.3 Drop-out-Problem

Der Begriff „Drop-out“ umfasst mehrere mögliche Ereignisse im Laufe einer Studie. Folgende Sachverhalte werden unter dieser Überschrift angesprochen:

1. Ablehnung des Beginns einer Studienintervention trotz vorheriger Zustimmung, z.B. wegen Unzufriedenheit mit dem Randomisationsergebnis
2. Abbruch der Studienintervention nach ihrem Beginn aus verschiedenen Gründen:
 - Ablehnung durch den Patienten wegen Unzufriedenheiten mit der Therapiegestaltung
 - Abbruch aus organisatorischen Gründen (z.B. Umzug des Patienten)
 - Abbruch seitens des Therapeuten wegen notwendiger, nicht protokollkonformer Interventionen (z.B. bei Kriseninterventionen wegen Suizidalität o.ä.)
3. Nachträgliche Feststellung eines Ausschlusskriteriums einer Studie und somit nachträglicher Ausschluss trotz initialer Studienaufnahme und –rekrutierung
4. Ablehnung der weiteren Studienteilnahme durch den Patienten. Mit Studienteilnahme ist hier die Teilnahme an Untersuchungen und Interviews sowie die Bearbeitung von Fragebögen etc. gemeint

Grundsätzlich bietet sich demgemäß die Unterscheidung von Therapieabbruch und Studienabbruch an, da hierfür jeweils unterschiedliche Konsequenzen erforderlich sind.

In Bezug auf den Therapieabbruch gilt für randomisierte klinische Studien seit längerem als eine der zentralen methodischen Forderungen die „intent-to-treat“-Analyse (zu Details

siehe Hollis und Campbell, 1999). Damit ist gemeint, dass die Hauptauswertung einer Studie alle Patienten einbeziehen sollte, die randomisiert wurden, nicht nur diejenigen, bei denen die Therapie wie geplant durchgeführt wurde (dies würde u.a. als „completer“-Analyse, according-to-protocol-Analyse oder „as-treated“-Analyse bezeichnet). In „completer“-Analysen kann die Strukturgleichheit der randomisierten Gruppen durch Therapieabbrüche vollständig aufgehoben sein, insbesondere bei höheren Therapieabbruch-Raten. Dies ist gerade in Studien über längerfristige und aufwändigere Therapien ein Problem (z.B. bei stationären Behandlungen).

Die Forderung nach intent-to-treat-Analysen gilt dabei vorrangig für die so genannten „large pragmatic trials“ (Hotopf et al., 1999), die nicht in erster Linie auf die Kausalität der Therapieeffekte (Hypothese: Therapie bewirkt Veränderung) abheben, sondern auf die Stärkung der empirischen Evidenz und Plausibilität einer Intervention in einem multikonditionalen Umfeld, in dem viele Einflüsse nur schwer beeinflussbar und konstant zu halten sind. „Pragmatic trials“ sind demzufolge keine klassischen Experimente. Jedoch kann durch die Randomisierung bei ausreichend großen Stichproben auch eine Vielzahl unbekannter Einflussfaktoren kontrolliert bzw. zwischen den Gruppen gleichgehalten werden (Hsu, 1989).

Ein wesentlicher Einwand von Therapeuten gegen „intent-to-treat“-Analysen ist, dass Patienten dazugezählt werden, die kaum oder unzureichend behandelt wurden. Das entscheidende Gegenargument ist (zumindest in Bezug auf „pragmatic trials“), dass auch Therapieabbrüche der jeweiligen Therapiestrategie zur Last zu legen sind, da sie, auf die Grundgesamtheit bezogen, eine unzureichende Therapiewirkung implizieren. Darüber hinaus gilt die Forderung nach „intent-to-treat“-Analysen umso mehr, je höher die Therapieabbruch-Rate in einer Studie ist et al. (Petrak et al., 1999).

Um bei Therapiestudien sowohl eine „intent-to-treat“-Auswertung als auch eine „as-treated“-Auswertung zu ermöglichen, sollen tunlichst alle Patienten, die am Beginn der Studie rekrutiert werden konnten, in alle Nachuntersuchungen einbezogen werden. Das bedeutet, dass auch Patienten, die die Studienintervention abgebrochen haben, in den Nachuntersuchungen berücksichtigt werden sollten. Dies erfordert eine zusätzliche Aufklärung des Patienten und ein geschicktes motivierendes Vorgehen, was vor allem auf Patienten zutrifft, die einer Kontrollgruppe zugeteilt worden sind.

Als Drop-out sollten nur die Patienten bezeichnet werden, von denen in einer Studie keine Daten mehr erfasst werden können, z.B. weil sie die weitere Mitarbeit ablehnen, sich suizidiert haben oder unbekannt verzogen sind. Diese Patienten sind als Studienabbrecher (im Unterschied zu Therapieabbrechern) zu bezeichnen. Die Daten aller übrigen Patienten sind im Sinne der oben beschriebenen „intent-to-treat“-Analyse zu verwenden. Studienabbrecher in dieser Definition stellen insofern ein Problem dar, als sie in Therapiestudien die Strukturgleichheit der Gruppen und damit die interne Validität gefährden. In

Prädiktionsanalysen kann durch eine hohe Drop-out-Rate eine erhebliche Verzerrung der Prädiktion resultieren. Die Minimierung der Studienabbruchrate ist daher ein wichtiges Ziel in der Durchführung von Follow-up-Studien. Hierzu sind nicht zuletzt auch organisatorische Maßnahmen zu treffen.

Studienabbrüche führen zu fehlenden Daten. Zur Behandlung dieser „missing data“ bestehen verschiedene Vorschläge, die vorwiegend auf verschiedenen Imputationsverfahren beruhen. Imputation bedeutet das Ersetzen fehlender Daten durch erschlossene oder berechnete, immer aber hypothetische Angaben. Häufig verwendete Alternativen sind „last observation carried forward“, Imputation aufgrund plausibler Therapie-Szenarien oder Imputation aufgrund der Daten der in der Studie verbliebenen Patienten et al. (Mazumdar et al., 1999). Jede dieser Alternativen ist kritikwürdig, daher besteht jeweils die Frage nach der am besten vertretbaren unter prinzipiell ungünstigen Alternativen. Die Konsequenz sollte sein, „missing data“ so gut wie möglich, nicht zuletzt auch durch die geeignete Wahl des Hauptzielkriteriums der Studie zu vermeiden (Lavori, 1992).

3.3 Organisatorische Fragen

Aus den bisherigen Ausführungen ergibt sich als die zentrale organisatorische Anforderung an Follow-up-Studien, Patienten und Studienmitarbeiter für die Studie zu gewinnen und in der Studie zu halten. Dabei ist die Langfristigkeit des Vorhabens für alle Seiten die wesentliche Herausforderung. Im Einzelnen sollten u.a. folgende Aspekte berücksichtigt werden: Drittmittelgeber, tragende Institution, Studienleitung und Studienmitarbeiter müssen sich über den langfristig zu erbringenden Aufwand und die erforderlichen Ressourcen im Klaren sein und entsprechende Unterstützung und Engagement aufbringen. Gute Follow-up-Studien sind kosten- und arbeitsaufwändig. Sie sind ohne ausreichende Drittmittelförderung nicht durchführbar.

In Bezug auf die Patienten sind mehrere Punkte zu beachten, die alle die Motivation zur Mitarbeit berühren.

Bereits zu Beginn der Studie muss dem Patienten in verbindlicher Form zugesichert werden, dass alle Informationen, die über ihn gesammelt werden, anonymisiert und nur so ausgewertet werden. Die Schweigepflicht ist in jedem Fall zu wahren und darf nur durch ausdrückliche und besondere Schweigepflichtsentbindung aus einem wichtigen Grund gebrochen werden. Der Proband muss absolute Sicherheit darüber haben, dass von ihm gegebene Informationen unter keinen Umständen missbraucht werden können.

Dies schließt strukturelle Maßnahmen ein. Neben selbstverständlichen Dingen wie der anonymen Kodierung der Daten ist auch z.B. an die anonyme Absenderangabe zu denken. Es sollte ein neutraler Absender genannt werden, der einen Rückschluss auf eine psychiatrische Klinik oder eine psychotherapeutische Institution nicht ermöglicht. Es sollte z.B. von einem Universitätsklinikum im Absender die Rede sein und nicht von der Universitätsklinik für Psychiatrie und Psychotherapie.

Es hat sich bewährt, dass die Probanden immer von derselben Person im Verlauf der Studie betreut werden. D.h. jedoch nicht, dass diese den Patienten betreuende Person auch alle Nachuntersuchungen selbst durchführen muss. Im Rahmen von Therapiestudien kann z.B. ein „blindes Rating“ notwendig werden. Hierbei weiß der Nachuntersucher nicht, welche therapeutische Intervention der Proband erfahren hat. Es hat sich bewährt, dass auch in solchen Fällen die Nachuntersuchung von der dem Probanden bereits bekannten Person moderiert und geleitet wird, ohne jedoch die „Blindheit“ des Ratings zu gefährden.

Der aussichtsreichste Weg, eine möglichst hohe Haltequote in der Studie zu erreichen, ist die persönliche Beziehung zwischen dem Untersucher und dem Probanden. Es ist deshalb außerordentlich wichtig, dass eine konstante Beziehungsperson für den Probanden zur Verfügung steht. Nach Möglichkeit sollte diese Beziehungsperson den Probanden in die Studie einführen und ihn bis zum Ende begleiten. Während des ganzen Studienzeitraumes sollte diese Beziehungsperson den Kontakt, notfalls auch außerhalb der Follow-up-Termine, aufrechterhalten. Dies kann geschehen durch Einladungen zu Symposien oder Beratungsgesprächen. Die Verfügbarkeit einer konstanten Beziehungsperson für den Probanden ist noch nicht gleichbedeutend mit der Notwendigkeit, dass diese Beziehungsperson auch alle Untersuchungen persönlich durchführt. Die Beziehungsperson kann den Probanden sehr viel besser als neue, fremde Personen zur Teilnahme an einzelnen Untersuchungen motivieren.

Falls der Proband zwar bereit ist, sich nachuntersuchen zu lassen, jedoch nicht in die Klinik oder in die die Untersuchung durchführende Institution kommen will, müssen zur Vermeidung von Studienabbrüchen auch Hausbesuche und Telefonkontakte erwogen werden. Die Durchführung von Hausbesuchen erhöht natürlich den Aufwand, was letztlich wieder ein Problem der finanziellen Ausstattung der Studie ist. Hausbesuche können jedoch zur Senkung der Studienabbruch-Rate beitragen. Bei Telefonkontakten ist zu erwägen, inwieweit hierdurch die interessierenden Variablen tatsächlich erhoben werden können. Objektive Daten wie Wiederaufnahmen oder Arbeitstätigkeit sind hier noch eher zu erfassen als z.B. psychopathologische Beurteilungen.

In der Regel wird man den Probanden zumindest die anfallenden Unkosten, die mit der Nachuntersuchung verbunden sind, erstatten müssen.

Um die Motivation zu erhöhen, an Nachuntersuchungen teilzunehmen, werden häufig auch Probandengelder in Aussicht gestellt und bezahlt. Manche Follow-up-Studien lassen sich ohne Probandengelder nicht mit ausreichendem Erfolg durchführen. Andererseits können Probanden, die aus altruistischen Gründen „sich der Forschung zur Verfügung stellen“, demotiviert werden. Hier ist ein sehr sorgfältiges Vorgehen, das ethische und motivationale Gesichtspunkte berücksichtigt, geboten.

Kosten

Mehrfach war schon angesprochen, dass Follow-up-Untersuchungen erhebliche Kosten verursachen. Für die folgende Beispielrechnung legen wir eine Studie mit folgenden Merkmalen zugrunde:

- Kontrollierte Therapiestudie mit zwei Therapiearmen, die zur Erlangung ausreichender statistischer Power 2x50 Patienten benötigt.
- Monatliche Untersuchung der Studienpatienten über einen Zeitraum von 2 Jahren. Die monatlichen Untersuchungen sind für die zeitnahe Erfassung der psychischen Symptomatik erforderlich. Somit sind 2400 Untersuchungskontakte zur realisieren.
- Untersuchungsdauer jeweils 60 Minuten. Der Zeitaufwand für die Mitarbeiter wird unter Berücksichtigung der Organisation und der Dokumentation mit mindestens 2,5 Stunden angesetzt (30 Minuten Organisation/Terminabsprachen, 60 Minuten Untersuchung, 60 Minuten Dokumentation). Die Zeitersparnis durch ausfallende Termine wegen Nicht-Erscheinen des Patienten wird durch den organisatorischen Mehraufwand aufgewogen.
- Ein BAT-IIa/2-Mitarbeiter kann demnach 32 Patienten betreuen (80 Monatsstunden/2,5 Stunden Untersuchungsdauer). Für die Studie werden 3 BAT-IIa/2-Mitarbeiter benötigt.
- Für die Anlage und Pflege der Datenbank, für das Datenmonitoring und statistische Auswertung ist eine weitere Stelle BAT IIa/2 vorzusehen.
- Für die Dateneingabe wird eine studentische Hilfskraft benötigt.

Kostenberechnung:

Mitarbeiter:	DM 384.000,-
4 Mitarbeiter x DM 48000,- ¹ x 2 Jahre	
Studentische Hilfskraft	DM 40.800,-
DM 20400,- ¹ x 2 Jahre	
Fahrtkosten Probanden	DM 60.000,-
2400 Termine x DM 25,- ²	
Sachkosten und Verbrauchsmaterial	<u>DM 25.000,-²</u>
inkl. PC	<u>DM 509.800,-</u>
Gesamtkosten	

¹ Nach den aktuellen Richtsätzen der DFG

² Pauschaler, überschlagsmäßiger Ansatz

Die ermittelte Summe berücksichtigt ausschließlich die Kosten für die Follow-up-Untersuchung. Kosten für die Therapie sind nicht berücksichtigt.

Die zugrunde gelegten Beispieldaten repräsentieren eine keineswegs „große“ Studie. Längere Untersuchungszeiten und höhere Fallzahlen erhöhen den Aufwand entsprechend.

3.5 Ethische Fragen

Die ethischen Fragen können hier nur angerissen werden. Einen differenzierten Überblick geben Helmchen & Vollmann (1999). Grundsätzlich sollten alle klinischen Studien von einer Ethikkommission auf ihre Unbedenklichkeit hin geprüft worden sein. Die Ethikkommissionen prüfen dabei zunehmend häufig auch, ob Aufwand und Ertrag einer Studie in angemessenem Verhältnis stehen. Von daher ist zu berücksichtigen, dass angesichts des hohen Aufwandes von Follow-up-Studien die entsprechenden Fragestellungen immer in besonderer Weise begründet werden müssen.

Bereits zum Beginn der Studie, spätestens bei der Stuserhebung müssen die Patienten darüber informiert werden, dass Follow-up-Untersuchungen stattfinden werden. Die Patienten müssen hierfür ihr Einverständnis geben. Die Zustimmung der Patienten nach ausführlicher Aufklärung ist der entscheidende Punkt. Die Zustimmung sollte in der Regel schriftlich erfolgen, nur in begründeten Ausnahmefällen mündlich, dann aber vor Zeugen und unter Dokumentation der Gründe, die eine schriftliche Zustimmung verhindern.

Von ethischer Seite sind zwei Kategorien von Untersuchungen zu unterscheiden. Heilversuche und Experimente. Therapiestudien gehören der Kategorie Heilversuch an. Sie sind nur dann gerechtfertigt, wenn die Therapie nach dem aktuellen Wissensstand zwar aussichtsreich ist (auf der Basis vorläufiger wissenschaftlicher Ergebnisse), jedoch noch nicht im strengen Sinne als wissenschaftlich in ihrer Wirksamkeit geprüfte und erfolgreiche Therapie belegt ist. Insbesondere in Randomisierungen muss der Patient explizit einwilligen. Experimente hingegen erbringen keinen potenziellen Nutzen für den individuellen an der Studie teilnehmenden Patienten. Prädiktionsanalysen, die nicht im Kontext von Therapien durchgeführt werden, müssen hier eingeordnet werden. Auf den fehlenden Nutzen für den Einzelnen ist dabei hinzuweisen. Gerade in solchen Studien wird ohne Aufwandsentschädigung die Motivation zur langfristigen Mitarbeit gering sein.

Patienten, die nicht ihr Einverständnis zur Durchführung von Follow-up-Studien gegeben haben, dürfen nicht in die Nachuntersuchungen einbezogen werden. Hat ein Patient jedoch zu Beginn der Studie sein Einverständnis erklärt, sollte er zu jeder Follow-up-Untersuchung eingeladen werden, es sei denn, er zieht im Verlauf der Studie sein Einverständnis ausdrücklich zurück.

Patienten haben jederzeit das Recht, aus einer Studie auszuscheiden. Dies muss respektiert werden. Die daraus entstehenden Drop-out-Probleme müssen hinter dem ethischen Aspekt zurückstehen.

Insgesamt bestehen für Follow-up-Untersuchungen keine spezifischen ethischen Probleme. Alle an anderer Stelle geführten Diskussionen um Einwilligungsfähigkeit, Aufklärung, informierte Zustimmung etc. gelten in analoger Weise auch hier.

Literatur

- Amori G, Lenox RH: Do volunteer subjects bias clinical trials? *J Clin Psychopharmacol* 1989; 9:321-327.
- Baumann U: Wie objektiv ist die Wirksamkeit der Psychotherapie? In: C Mundt, M Linden, W Barnett (Hrsg.): *Psychotherapie in der Psychiatrie*. Wien: Springer, 1997 27-32.
- Bortz J, Döring, N: *Forschungsmethoden und Evaluation*. Berlin: Springer, 1995.
- Carroll KM, Nich C, McLellan AT, McKay JR, Rounsaville BJ: 'Research' versus 'real-world' patients: representativeness of participants in clinical trials of treatments for cocaine dependence. *Drug Alcohol Depend* 1999; 54:171-177.
- Häfner H, Löffler W, Maurer K, Hambrecht M, an-der-Heiden W: Depression, negative symptoms, social stagnation and social decline in the early course of schizophrenia. *Acta Psychiatr Scand* 1999; 100:105-118.
- Helmchen H, Vollmann J: Ethische Fragen in der Psychiatrie. In: H Helmchen, F Henn, H Lauter, N Sartorius (Hrsg.): *Psychiatrie der Gegenwart*, Bd. 2: *Allgemeine Psychiatrie*. Berlin: Springer, 1999 521-577.
- Hollis, S, Campbell, F: What is meant by intention to treat analysis? Survey of published randomized controlled trials. *BMJ* 1999; 319:670-674.
- Hornung WP, Feldmann R, Klingberg S, Buchkremer G, Reker T: Long-term effects of a psychoeducational psychotherapeutic intervention for schizophrenic outpatients and their key-persons - results of a five-year follow-up. *Eur Arch Psychiatry Clin Neurosci* 1999; 249:162-167.
- Hotopf M, Churchill R, Lewis G: Pragmatic randomized controlled trials in psychiatry. *Br J Psychiatry* 1999; 175:217-223.
- Hsu LM: Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. *J Consult Clin Psychol* 1989; 57:131-137.
- Huber G, Gross G, Schüttler R, Linz M: Longitudinal studies of schizophrenic patients. *Schizophr Bull* 1980; 6:592-605.
- Kazdin AE: Methodology, design, and evaluation in psychotherapy research. In: AE Bergin, SL Garfield (eds.): *Handbook of psychotherapy and behavior change*. New York: Wiley, 1994 19-70.
- Kazdin AE, Kendall PC: Current progress and future plans for developing effective treatments: Comments and Perspectives. *J Clin Child Psychol* 1998; 27:217-226.
- Keller MB: The long-term treatment of depression. *J Clin Psychiatry* 1999; 60 Suppl 17:41-45.
- Keller MB, Boland, RJ: Implications of failing to achieve successful long-term maintenance treatment of recurrent unipolar major depression. *Biol Psychiatry* 1998; 44:348-360.
- Lavori PW: Clinical trials in psychiatry: should protocol deviation censor patient data? *Neuropsychopharmacology* 1992; 6:39-48.
- Mancuso DM, Townsend MH, Mercante DE: Long-term follow-up of generalized anxiety disorder. *Compr Psychiatry* 1993; 34:441-446.
- Mazumdar S, Liu KS, Houck, PR, Reynolds CF: Intent-to-treat analysis for longitudinal clinical trials: coping with the challenge of missing values. *J Psychiatr Res* 1999; 33:87-95.
- Petrak F, Hardt J, Nickel R, Tiber Egle, U: Checkliste zur Bewertung der wissenschaftlichen Qualität kontrollierter psychotherapeutischer Interventionsstudien (CPI). *Psychotherapeut* 1999; 44:390-393.
- Schweizer E, Rickels K: The long-term management of generalized anxiety disorder: issues and dilemmas. *J Clin Psychiatry* 1996; 57 Suppl 7:9-12.
- Spiegel R: *Einführung in die Psychopharmakologie*. Stuttgart: Huber, 1988.
- Stapleton JA, Sutherland G, Russell MA: How much does relapse after one year erode effectiveness of smoking cessation treatments? Long-term follow up of randomized trial of nicotine nasal spray. *BMJ* 1998; 316:830-831.
- Thase ME: Redefining antidepressant efficacy toward long-term recovery. *J Clin Psychiatry* 1999; 60 Suppl 6:15-19.
- Versiani M, Amrein R, Montgomery SA: Social phobia: long-term treatment outcome and prediction of response - a moclobemide study. *Int Clin Psychopharmacol* 1997; 12:239-254.
- Weiden P, Aquila R, Standard J: Atypical antipsychotic drugs and long-term outcome in schizophrenia. *J Clin Psychiatry* 1996; 57 Suppl 11:53-60.

Dr. phil. Dipl. Psych. Stefan Klingberg
 Osianderstr. 24
 D-72076 Tübingen
 Tel.: 07071 2982330 • Fax: 07071 294141
 Email: stefan.klingberg@med.uni-tuebingen.de