

MISTILTEINN 2004/5

Sonderdruck

Jahrgang	5
Seite	44 – 73
Autor	Renatus Ziegler
Titel	Möglichkeiten und Grenzen klinischer Studien mit Iscador
Copyright	Verein für Krebsforschung, Arlesheim/Schweiz
ISSN	1660-1173
Verlag	Kooperative Dürnau

Klinisch-therapeutische Forschung zur Wirksamkeit von Arzneimitteln

Einführung: Ursachenforschung, Bedingungsforschung, Effektforschung

Klinische Studien mit Mistelpräparaten im allgemeinen und mit Iscador im besonderen stehen in einem Spannungsfeld, dessen gegensätzliche Pole nur mit grossem Forschungsaufwand zu überbrücken sind. Auf der einen Seite müssen die konkreten Fragen und Bedürfnisse von Patienten und den sie behandelnden Ärzten aufgegriffen und Erkenntnishilfen zur Entscheidungsfindung zur Verfügung gestellt werden. Hier steht die individuelle Erkenntnisfähigkeit und praktische Erfahrung von Ärzten und Pflegenden im Vordergrund; deren diagnostische und therapeutische Kompetenz soll genutzt und gefördert werden. Hinzu kommt die Notwendigkeit und Herausforderung – gerade im Bereich der Komplementärmedizin (Horneber et al. 2003) –, die individuellen Entscheidungsmöglichkeiten von Patientinnen und Patienten zu berücksichtigen, zu ermöglichen, ja letztlich aktiv zu fördern. Hier steht das spezifische Verhältnis zwischen Arzt, Pflegenden und Patient im Zentrum. In diesem Rahmen werden die diagnostischen und therapeutischen Massnahmen initiiert und dem weiteren Gang der Entwicklung angepasst. Hier liegt das Feld, wo in der klinischen Anwendung von Mistelpräparaten im allgemeinen und von Iscador im besonderen die unmittelbarsten und fruchtbarsten Erfahrungen vorliegen und weit zurückreichen.

Auf der anderen Seite machen es die spätestens seit den 90er Jahren des 20. Jahrhunderts wichtiger werdenden Bedingungen der Neu- und Nachzulassung von Arzneimitteln und ihrer Erstattung durch die Krankenkassen notwendig, dass gemäss den definierten Standards der gegenwärtigen konventionellen Medizin geforscht, evaluiert und publiziert wird. Dadurch ergibt sich im Idealfall zugleich eine Bestätigung und Vertiefung der individuellen ärztlichen Erfahrung, mit anderen Worten, eine (im Durchschnitt zutreffende) Erweiterung des am einzelnen Patienten erlebten Geschehens auf ein ganzes Kollektiv von Patienten verschiedensten klinischen und seelischen Charakters.

Das Ziel der ersten Art klinisch-therapeutischer Forschung ist die Ergündung des «Heilungsprozesses» beim individuellen Patienten in dessen spezifischer Situation. Was hier «Heilung» bedeutet, ist nicht von vornherein definierbar, sondern Ergebnis der Zusammenarbeit von Arzt, Pflegenden und Patient. Stellt man die Erringung, Förderung und Erhaltung der Freiheitsfähigkeit, der individuellen Autonomie des kranken Menschen und der ihn betreuenden Ärzte und Pflegenden in den Mittelpunkt aller therapeutischen Bemühungen, so wird sich auch der Erfolg einer Therapie daran messen müssen, inwiefern sie dem Patienten und den beteiligten Ärzten und Pflegenden eine Weiterentwicklung ihrer Autonomie ermöglicht. Hierbei kann die Linderung, Stabilisierung und Eliminierung physisch manifester Krankheitssymptome ein entscheidender Faktor sein, auch wenn es kein alleiniges Kriterium für die Wirksamkeit einer Therapie mehr ist. (Siehe dazu und zum folgenden die Untersuchung des Begriffs der therapeutischen Wirksamkeit auf der Grundlage des ethischen Individualismus Rudolf Steiners 1894/1918 in Ziegler 2004b.)

Von fundamentaler Bedeutung für solche therapeutischen Verhältnisse ist die innere Verbindlichkeit, mit der sich alle Beteiligten zu einer Verantwortungsgemeinschaft zusammenfinden und damit das gesamte klinisch-therapeutische Geschehen selbst gestalten. Aufgrund des so angestrebten tieferen Verständnisses der Krankheit, ihrer Ursachen und ihrer Therapiemöglichkeiten beim individuellen kranken Menschen kann hier von *Ursachenforschung* und entsprechender *Ursachentherapie* gesprochen werden. Für letzteres hat Rudolf Steiner den Ausdruck «rationale Therapie» (siehe zum Beispiel Steiner 1920) eingeführt.

Von ganz anderen Voraussetzungen gehen die von Behörden und Krankenkassen in Zusammenarbeit mit der Scientific/Medical Community und den Verbänden der Pharmazeutischen Industrie aufgestellten Regeln aus: In deren Zentrum stehen hohe Qualitätsstandards für quantitative und statistisch gesicherte Vergleiche zweier verschieden therapierter Gruppen (z.B. konventionelle onkologische Therapie vs. konventionelle onkologische Therapie plus Iscador-Therapie) aus demselben Patientenkollektiv, das heisst im hier gegebenen Falle aus einer durch bestimmte Eigenschaften (Ein- und Ausschlusskriterien) festgelegten Auswahl an Krebs erkrankter Menschen. Hierbei ist es für die Verlässlichkeit der Resultate einer klinischen Studie zur Prüfung eines Arzneimittels wie Iscador entscheidend, dass keinerlei individueller Bezug von Patienten, Arzt und Pflegenden zur angewendeten Therapie und deren Beurteilung vorhanden ist oder zustande kommt. Andernfalls gibt es einen systematischen Fehler, das heisst die Studienresultate sind verzerrt, da dann keine strenge Gleichartigkeit (Strukturgleichheit) der beiden Gruppen mehr besteht (abgesehen von der geplanten

verschiedenen Behandlung). Die Strukturgleichheit der Vergleichsgruppen kann einerseits durch eine zufällige Zuordnung (Randomisation) der Patienten in die entsprechenden Therapiegruppen gewährleistet werden und andererseits durch eine Verblindung (Maskierung) von Arzt, Pflegenden und Patienten gegenüber der eingesetzten Therapie. Letzteres bedingt natürlich die Existenz eines dem eingesetzten Arzneimittel in Aussehen und Anwendung möglichst gleichartigen, aber unwirksamen Kontrollstoffes (Plazebo). Dieser Ansatz kann *Effektforschung* und die darauf aufbauende Therapie *Effekttherapie* genannt werden. (Dieser Ansatz ist noch nicht alt; zur historischen Genese der gegenwärtig relevanten Prinzipien der klinischen Forschung, siehe Kiene 2001 und Ziegler 2004a.) Letzteres ist im wesentlichen die Basis für den Ansatz der sogenannten «Evidence-based Medicine» (Sackett et al. 1999).

Ein Zwischenweg, bei welchem die praktikable und qualitativ hochstehende Dokumentation von Einzelfällen im Vordergrund steht, bei der aber auch nach typischen Mustern im Ablauf einer Krankheit und deren therapeutischer Behandlung gesucht werden kann, bei dem folglich auch Fallserien interessant werden, sei hier *Bedingungsforschung* genannt. Hier wird im Einzelfall nach klar erkennbaren notwendigen Bedingungen gesucht, unter denen bestimmte Folgen, insbesondere Krankheits- und Heilungsprozesse, eintreten. Im Vordergrund steht dabei die Erfassung aller therapeutischen Massnahmen sowie die Messung der Körperfunktionen (z.B. Tumorbefund) und der Lebensprozesse (z.B. Blutbild, Immunstatus) sowie die qualitativ und quantitativ auswertbare Erfassung von seelisch-geistigen Prozessen durch Fragebogen und Interviews. Daraus ergibt sich im Ideal eine Abfolge von Ereignissen, die in einem weitgehend notwendigen Verhältnis stehen: Wenn bestimmte Massnahmen ergriffen werden, so treten bestimmte Folgen ein.

Studienqualität

Es ist eine historische Tatsache, dass die Qualität und Verlässlichkeit von klinischen Studien durch die strengen Anforderungen an die Durchführung und Berichterstattung von Wirksamkeitsforschung im Sinne der Effektforschung verbessert wurde: Es bedarf einer grossen planerischen und logistischen Sorgfalt, um eine brauchbare randomisierte Studie durchzuführen.

Es ist nun eine grosse Herausforderung an die Planung und Durchführung von Bedingungsforschung, die aus der Situation der Effektforschung gewohnten hohen Qualitätsstandards an Präzision und Verlässlichkeit der Studiendaten auf die Situation der Bedingungsforschung zu übertragen. Dieser Herausforderung haben sich insbesondere Helmut Kiene und Gunver Kienle gestellt und unter dem Stichwort «Cognition-based Medicine»

entsprechende Arbeitsgrundlagen grundsätzlicher und methodischer Art ausgearbeitet (Kiene 2001; Kienle und Kiene 2003a, 2003b; Kienle et al. 2004). Bei diesen Autoren wird anstatt von Bedingungsforschung, Bedingungen und Folgen, von Ursachen und Wirkungen gesprochen und dementsprechend von Kausalverhältnis oder von der Kausalgestalt, welche vermöge der ideellen Methode der Gestalterkenntnis erfasst werden können. So führen diese Autoren aus, dass, wenn ein unmittelbarer ideeller Zugriff auf die Kausalgestalt nicht möglich ist, eine diese Gestalt abbildende Methode in Frage kommt, die bei klinischen Ereignissen nach verschiedenen Arten von Korrespondenzen zwischen möglichen Ursachen und möglichen Wirkungen (zeitliche Muster, räumliche Bezüge, prozessuale Funktionen, morphologische Beziehungen, Dosis-Wirkung-Beziehungen etc.) sucht. Dies führt zu konkret handhabbaren Regeln, die direkt auf klinische Situationen anwendbar sind.

Mit den im Rahmen der «Cognition-based Medicine» aufgestellten Kriterien zur *Wirksamkeitsbeurteilung im Einzelfall* lassen sich mit hoher Sicherheit diejenigen Bedingungen herausfinden und weiter untersuchen, die für das Eintreten bestimmter Folgen notwendig sind, das heisst ohne welche das in Frage stehende Ereignis (Linderung der Symptome, Stabilisierung oder Heilung der Krankheit) nicht eintreten würde.

Gunver Kienle und Helmut Kiene haben ausserdem zusammen mit Klinikern wiederholt darauf aufmerksam gemacht, dass es irreversible und gefährliche Folgen für die Wertschätzung, Ausbildung und Weiterentwicklung der individuellen ärztlichen Urteilskraft und des individuellen Verhältnisses zwischen Arzt, Pflegenden und Patient haben kann, wenn Forschung und Therapie *nur* noch auf der Grundlage randomisierter Studien, also im Rahmen der Effektforschung, gemacht würde (Kienle et al. 2003b, Kienle und Kiene 2004a, Kienle und Kiene 2003b). In diesem Sinne muss Effektforschung zumindest durch Bedingungsforschung, das heisst durch die Methoden der «Cognition-based Medicine», ergänzt werden.

Die Ursachenforschung muss das eigentliche Ziel jeder Erforschung von Krankheitsursachen und der sich an diese anschliessenden Erforschung sachgemässer therapeutischer Massnahmen sein. Trotz grosser theoretischer und praktischer Schwierigkeiten, welche diese Art der Forschung bietet, muss dieses Ziel aufrecht erhalten werden. Denn hier geht es um tiefgreifende Auseinandersetzungen des Arztes mit dem individuellen Patienten, die bis in die Erkenntnis und Handhabung von Lebenskräften sowie von seelischen und geistigen Prozessen, letztlich bis in Schicksalsbeziehungen vorstossen kann (siehe dazu Steiner 1910). In diesem Sinne können Bedingungs- und Effektforschung nur als Hilfsmittel

aufgefasst werden, welche diesem Ziele dienen. Sie dürfen nicht zum Selbstzweck der klinisch-therapeutischen Forschung werden. Zusammen stellen sie jedoch praktisch und gegenwärtig umsetzbare Forschungsmethoden zur Verfügung und erfüllen in verschiedener Hinsicht zumindest einige der Anforderungen von Behörden und Krankenkassen sowie der Scientific Community und der Patientenverbände.

Aus diesem Blickwinkel ergibt sich für die im folgenden betrachteten vergleichenden klinischen Studien mit Iscador eine ganz andere Perspektive, als diejenige der «Evidence-based Medicine». Deren meist positive Ergebnisse hinsichtlich Wirksamkeit und Verträglichkeit der Iscador-Therapie können – im wesentlichen unabhängig von der Qualität der jeweiligen Studie – als Bestätigung der täglichen klinischen Erfahrung erlebt werden. Solche Studien sind dann deshalb bedeutsam, weil sie im Grossen das zeigen, was im Kleinen am Krankenbett oder in der Ambulanz täglich erlebt wird. (Einen Eindruck von dieser Erlebnisqualität mit Mistelpräparaten im allgemeinen und mit Iscador im besonderen geben mannigfaltige Einzelfallberichte und Fallserien, siehe dazu Kienle und Kiene 2003b, S. 469–484, Kienle und Kiene 2004c). Aus dem Gesichtspunkt der «Evidence-based Medicine» dagegen stehen die quantitativ vergleichenden Studien im Vordergrund. Einzelfälle und Fallserien gelten in der Regel nur als anekdotische Ergänzungen oder Ausschmückungen der harten Endresultate qualitativ hochstehender klinischer Studien.

Vergleichende Studien mit Iscador

Im folgenden stehen die Anforderungen der klinischen Effektforschung im Vordergrund, ihre Möglichkeiten und Grenzen im Bereich der Misteltherapie im allgemeinen und der Iscador-Therapie im besonderen. Die zentralen Stichworte in diesem Zusammenhang sind Studienqualität, Wirksamkeit und Sicherheit von Mistelpräparaten, insbesondere von Iscador. Ein charakteristischer Vorteil quantitativ-vergleichender Studien ist, dass sie, bei hoher Studienqualität, die Wirksamkeit und Sicherheit eines Arzneimittels weitgehend unabhängig von individuellen Charakteristiken von Patienten, Pflegenden, Ärzten und Kliniken oder onkologischen Schwerpunktsprachen in einem statistischen Sinne nachweisen. Bei entsprechendem Studienerfolg ist damit garantiert, dass z.B. das Arzneimittel Iscador im Durchschnitt mit hoher Wahrscheinlichkeit sicher und wirksam ist für einen breiten Anwendungsbereich unter den verschiedensten Umständen. Als Hintergrund für die Beurteilung des Stellenwertes vergleichender klinischer Studien mit Mistelpräparaten, insbesondere mit Iscador, folgen einige grundsätzliche Hinweise zur Struktur sowie zu den Vorteilen und Nachteilen vergleichender Studien (siehe dazu auch Ziegler 2004c, Reeves 2004a).

Strukturen vergleichender Studien

Experimentelle und beobachtende Forschung

Das zentrale Fundament eines Studienplans (Studienprotokoll, Prüfplan) für die therapeutische Effektforschung ist eine klare klinische Fragestellung, zum Beispiel: Hat die Therapie mit Iscador zusätzlich zur Standardtherapie im Vergleich zur Standardtherapie alleine eine verlängernde Wirkung auf die Überlebenszeit?

Es gibt zwei Arten von Fragestellungen, je nachdem ob die klinische Studie als Experiment oder als Beobachtung angelegt ist. Bevor man sich diesen unterschiedlichen Forschungsfragen zuwenden kann, muss genauer zwischen Intervention und Exposition unterschieden werden. Damit können dann die beiden Arten von Forschungsfragen näher charakterisiert werden (siehe Kasten 1).

Bei der experimentellen Forschung kann der forschende Arzt (immer das informierte Einverständnis der Patienten vorausgesetzt) bei der Gruppenzuteilung entweder den Patienten die Intervention durch einen nicht durch medizinische oder psychische Faktoren bedingten Prozess, in der Regel einen Zufallsprozess (Randomisation), zuordnen; oder er kann Sachkenntnisse und/oder Präferenzen

Eine **Intervention** ist eine gut definierte Behandlung, eine Therapie der Patienten, welche für den Zweck der Forschung eingesetzt wird, das heisst, die gemäss Studienplan begonnen, beendet oder variiert wird (oder werden kann) durch den forschenden Arzt (Studienleiter).

Eine **Exposition** ist ein definierter therapeutischer Vorgang oder Einflussfaktor, der im Rahmen des Forschungsprojektes durch den behandelnden Arzt nicht ohne weiteres eingesetzt, gestoppt oder beendet werden kann (z.B. eine vom Patienten gewünschte Iscador-Therapie). Dies bedeutet, dass die Auswahl und die konkrete Durchführung der Therapie unabhängig von der geplanten und implementierten Studie festgelegt und realisiert wird, zum Beispiel aufgrund von Sachkenntnis und Präferenzen des behandelnden Arztes oder des Patienten.

Experimentelle Forschungsfrage: Hat die Intervention (Bedingung, «Ursache», insbesondere die zu untersuchende Therapie, etwa Iscador-Therapie) zusammen mit den Expositionen und den äusseren Umständen (prognostische Faktoren, Kovariablen, zum Beispiel Begleittherapien) einen Einfluss auf die Zielgrösse («outcome», «Wirkung», zum Beispiel Überlebenszeit oder Lebensqualität) im Vergleich mit einer anderen Intervention oder keiner Intervention (z.B. nur konventionelle Basistherapie)?

Beobachtende oder epidemiologische Forschungsfrage: Ist die Exposition (Bedingung, «Ursache», insbesondere die zu untersuchende Therapie, Risikofaktor) zusammen mit den äusseren Umständen (prognostische Faktoren, Kovariablen, zum Beispiel Begleittherapien) assoziiert oder korreliert mit der Zielgrösse («outcome», «Wirkung», zum Beispiel Überlebenszeit) im Vergleich zu anderen Expositionen?

Kasten 1: Experimentelle und beobachtende Forschungsfragen.

des Arztes und/oder des Patienten berücksichtigen. Auch im letzteren Fall, das heisst bei einer nicht randomisierten vergleichenden experimentellen klinischen Studie, wird die Therapie erst im Rahmen der Studie initiiert, durchgeführt und begleitet, steht also nicht unabhängig von der Studie fest (andernfalls handelt es sich nicht um eine experimentelle Studie). Im Falle der beobachtenden oder epidemiologischen Forschung kommt im Rahmen der Studie keine solche Zuordnung in Betracht. Die Entscheide werden ausserhalb des direkten Einflussbereiches des Studienleiters und damit ausserhalb der Studie gefällt und nur in ihren Auswirkungen beobachtet. Die wichtigsten Zeitstrukturen der klinischen Forschung sind in Kasten 2 und Figur 1 dargestellt.

Eine Beobachtungsstudie heisst epidemiologische **Querschnittstudie** oder **transversale Studie** (Figur 1), wenn alle Beobachtungen oder Messungen der Studienteilnehmer im wesentlichen zum selben Zeitpunkt gemacht werden. Bei dieser Art Forschung geht es vor allem um die Beschreibung eines Zustandes einer bestimmten Population, das heisst einer durch klinische Kriterien (Tumorlokalisation, Tumorstadium etc.) definierten Gruppe von Patienten. Die Studie heisst **longitudinal**, wenn die Studienteilnehmer während einer gewissen Zeit periodisch (mindestens zweimal) untersucht werden hinsichtlich Veränderungen in der Zielgrösse nach im Voraus definierten Expositionen oder Interventionen.

Die longitudinale Forschung und ihr Datensammlungsschema heisst **vorwärts gerichtet** (Figur 1), wenn sie mit der Erfassung der Interventionen oder Expositionen beginnt, und zu einem späteren Zeitpunkt die Auswirkungen auf die Zielgrösse (zum Beispiel Lebensqualität) untersucht («von den Bedingungen zu den Folgen»). Bei solchen Studien werden die Patienten aufgrund von medizinischen Basisdaten ausgewählt, unabhängig von der ins Auge gefassten Zielgrösse. Gruppen von Patienten, die über einen gewissen Zeitraum in einer Studie verfolgt werden, heissen Kohorten und die entsprechenden Studien **Kohortenstudien**.

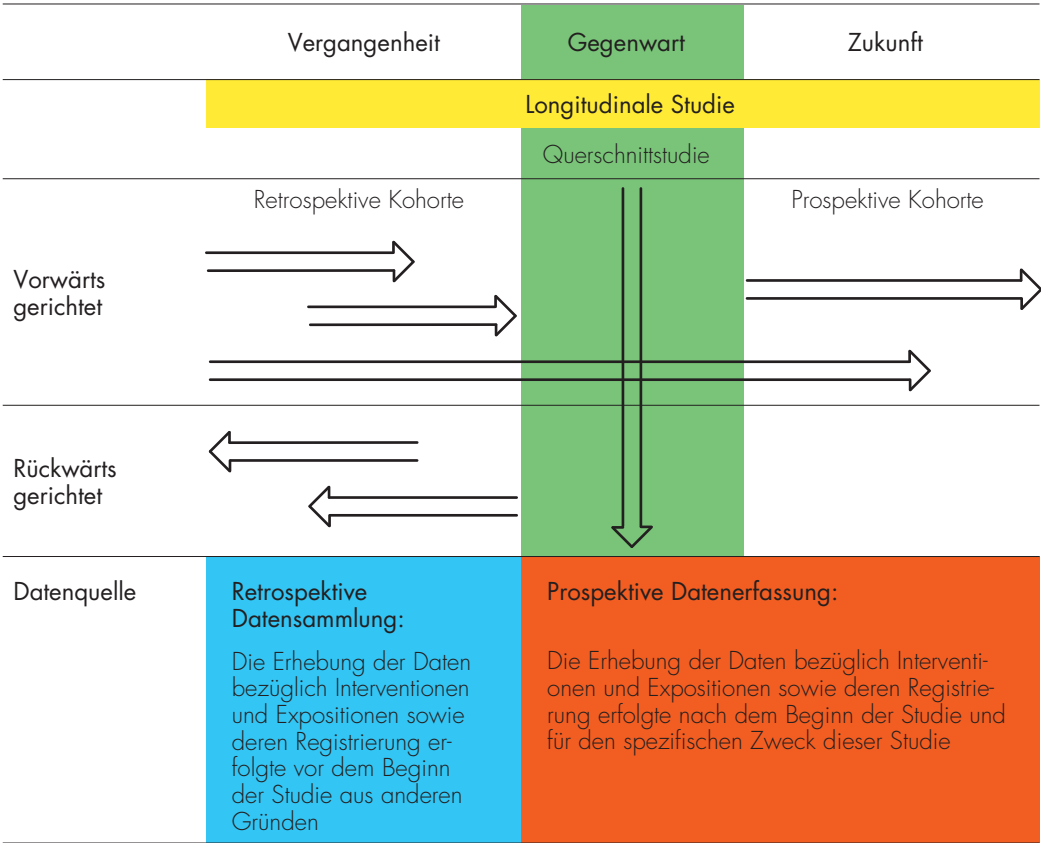
Die longitudinale Forschungsrichtung und ihr Datensammlungsschema heisst **rückwärts gerichtet**, wenn die Studienteilnehmer gemäss den Werten einer Zielgrösse (outcome, zum Beispiel Krankheitszustand oder bestimmter Wert der Lebensqualität) ausgewählt und registriert werden und danach nach Expositionen oder Einflussfaktoren (zum Beispiel Rauchen) gesucht wird («von den Folgen zu den Bedingungen»).

Das Datensammlungsschema heisst **prospektiv** (auch prolektiv oder aktuell), wenn die Erfassung der Interventionen und Expositionen ausschliesslich für den Zweck des Experimentes oder der Beobachtungsstudie gemacht werden, das heisst, wenn die Datenerfassung erst nach der Formulierung der exakten Forschungsfrage und der Ausarbeitung einer detaillierten Forschungsplanung (Protokoll) begonnen wurde (Figur 1). Die wesentlichen Vorteile der prospektiven Datenerfassung liegen in der Kontrolle über die Qualität, den Umfang und die Vollständigkeit der Daten vermöge einer guten Planung und Qualitätskontrolle (Monitoring).

Das Datensammlungsschema heisst **retrospektiv** (auch retrolektiv oder historisch), wenn die Studiendaten aus medizinischen Archiven (Patientendaten von Arztpraxen und Kliniken), Datenregistern (z. B. regionale Krebsregister) etc. gesammelt werden, bei welchen Personen, die nicht in die Studie involviert gewesen sind, die Daten gesammelt haben. Mit anderen Worten, die Datenerfassung fand bereits unter anderen Umständen und mit anderen Schwerpunkten statt, bevor die aktuelle Forschung geplant und durchgeführt wird.

Kasten 2: Zeitstruktur der klinischen Forschung: Forschungsrichtung und Datensammlung (siehe dazu Figur 1).

Aufgrund der vor allem im deutschsprachigen Europa weit verbreiteten und sehr individuell gehandhabten Misteltherapie im allgemeinen und der Iscador-Therapie im besonderen ist experimentelle Forschung in diesen Ländern kaum mehr zu realisieren. Für die hierfür geforderte Standardisierung der Therapie und die Kontrolle der Therapie-Optionen durch den Studienleiter ist wenig Bereitschaft oder Verständnis vorhanden. Der experimentelle Ansatz (auch ohne Randomisierung) ist jedoch durchaus vorstellbar bei vergleichenden Untersuchungen verschiedener Therapie-Verfahren, z.B. auf der Grundlage verschiedener Mistel-Wirtsbäume



Figur 1: Zeitstruktur der klinischen Forschung: Unterscheidung von vorwärts und rückwärts gerichteter longitudinaler Forschung sowie prospektiver und retrospektiver Datenerhebung (siehe dazu Kasten 2).

oder verschiedener Begleittherapien (ergänzende Arzneien, künstlerische Therapien, physiotherapeutische Anwendungen). Dies wird insbesondere dann interessant, wenn es um die Optimierung der konkreten Details einer Therapie geht (sogenannter Therapieoptimierungsvergleich TOPV, siehe Porzsolt et al. 1996). Im Rahmen einer beobachtenden Forschung sind solche Vergleiche in der Regel nicht möglich oder zumindest aufgrund der grossen Therapievelfalt nur unter sehr erschwerten Bedingungen (z.B. sehr grosse Probandenzahlen) erfassbar.

Von der Forschungsidee zur konkreten Umsetzung

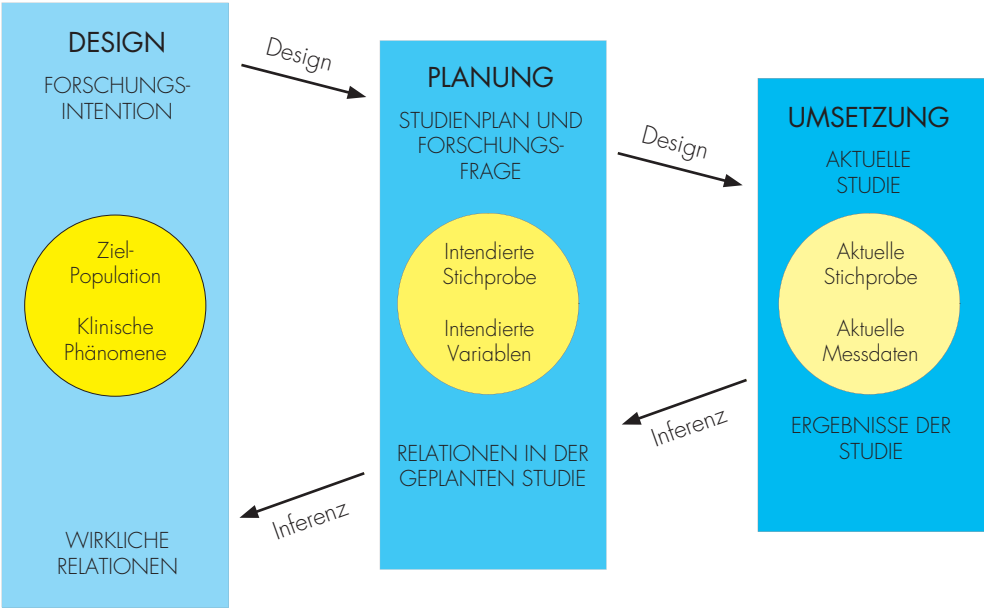
Im *Designstadium* der Wirksamkeitsforschung wird die Forschungsfrage reflektiert und formuliert. Es wird nach der tatsächlichen Relation bestimmter medizinischer Phänomene, das heisst nach der zutreffenden Beziehung bestimmter Krankheiten (z.B. Krebs) und deren Therapie (z.B. konventionelle Therapie und zusätzlicher Iscador-Therapie) zu klinisch manifesten Phänomenen für eine umfassende Zielpopulation (Patientenkollektiv) gesucht. In der Wirksamkeitsforschung wird dabei angenommen, dass die Interventionen und Expositionen mit den medizinisch feststellbaren Tatsachen in einem kausalen Verhältnis (im Sinne eines Verhältnisses von Bedingung und Folge) stehen.

Diese (ideale) Situation wird dann in einen *Studienplan* (Prüfplan, Protokoll) umgesetzt, welcher die konkrete Forschungsfrage spezifiziert, die intendierte Patientenstichprobe (z.B. Mammakarzinom eines bestimmten Stadiums) und die intendierten Messvariablen (z.B. Immunstatus, bestimmte Facetten der Lebensqualität, progressionsfreies Intervall, Überlebenszeit) festlegt. Die «wahre» Relation im medizinischen Forschungsuniversum wird ersetzt durch die im Rahmen der Studie erfassbaren Relationen (Figur 2).

Das nächste Stadium ist die *Umsetzung* oder *Implementierung*, die zur aktuellen Studie führt. Aufgrund verschiedener Einschränkungen (z. B. ausgewähltes Patientenkollektiv vermöge der Bedingungen einer bestimmten Klinik) und unvorhergesehener Umstände (z. B. zu viele Komorbiditäten, zu hohes Alter, Studienverweigerer, Verweigerer der Randomisation) wird die intendierte Stichprobe auf die aktuelle Stichprobe reduziert und die intendierten Messvariablen auf die tatsächlichen gemessenen Variablen (z. B. weniger Immunparameter, nur lückenhafte Erfassung der Lebensqualität) reduziert.

Im Stadium der *Analyse* werden die Studiendaten analysiert. Die Zielvariablen und die prognostischen Variablen, insbesondere die Therapiemassnahmen, werden hinsichtlich ihrer Beziehungen (Assoziationen) analysiert und wenn möglich im Sinne eines kausalen Verhältnisses interpretiert.

Bei jeder Studie gehen vom Designstadium zum konkreten Studienplan viele Aspekte verloren, die man ursprünglich gerne in der Studie gehabt hätte, die sich aber aus praktischen Gründen nicht realisieren liessen. Für Mistelstudien im allgemeinen und Iscador-Studien im besonderen trifft dies auch zu. Die grössten Schwierigkeiten entstehen jedoch beim Übergang von der Planung zur Umsetzung. Aufgrund eindeutiger Präferenzen und oft weit fortgeschrittener Krankheitsstadien ist die Rekrutierung von Teilnehmern für eine prospektive Studie jeweils ein sehr zeitaufwendiger und kostenintensiver Prozess, bei dem immer wieder Kompromisse (das heisst Anpassungen der Ein- und Ausschlusskriterien, Neuberechnung der notwendigen Fallzahl) gemacht werden müssen, die einen gravierenden Einfluss auf das untersuchte Patientenkollektiv haben.



Figur 2: Zeitlicher Ablauf der klinischen Forschung: Reduktion der ursprünglichen Forschungsintention auf die praktische Studienfrage und der ins Auge gefassten Zielpopulation auf die aktuelle Patientenstichprobe der Studie (verändert nach Hulley et al. 2001, Figur 1.6).

Beispiel: Randomisierte Studie zum nichtkleinzelligen Bronchialkarzinom

Hintergrund

An der von der deutschen Bundesarbeitsgemeinschaft für Rehabilitation (BAR) im Auftrag der Bundesversicherungsanstalt für Angestellte (BfA) in den Jahren 1978 bis 1987 durchgeführten multizentrischen Studie über Krebszusatztherapien beim fortgeschrittenen nichtkleinzelligen Bronchialkarzinom (Dold et al. 1991) lassen sich fast alle der hier diskutierten Möglichkeiten und Grenzen von Studien zur Iscador-Therapie demonstrieren. Deshalb wird sie hier mit einiger Ausführlichkeit vorgestellt.

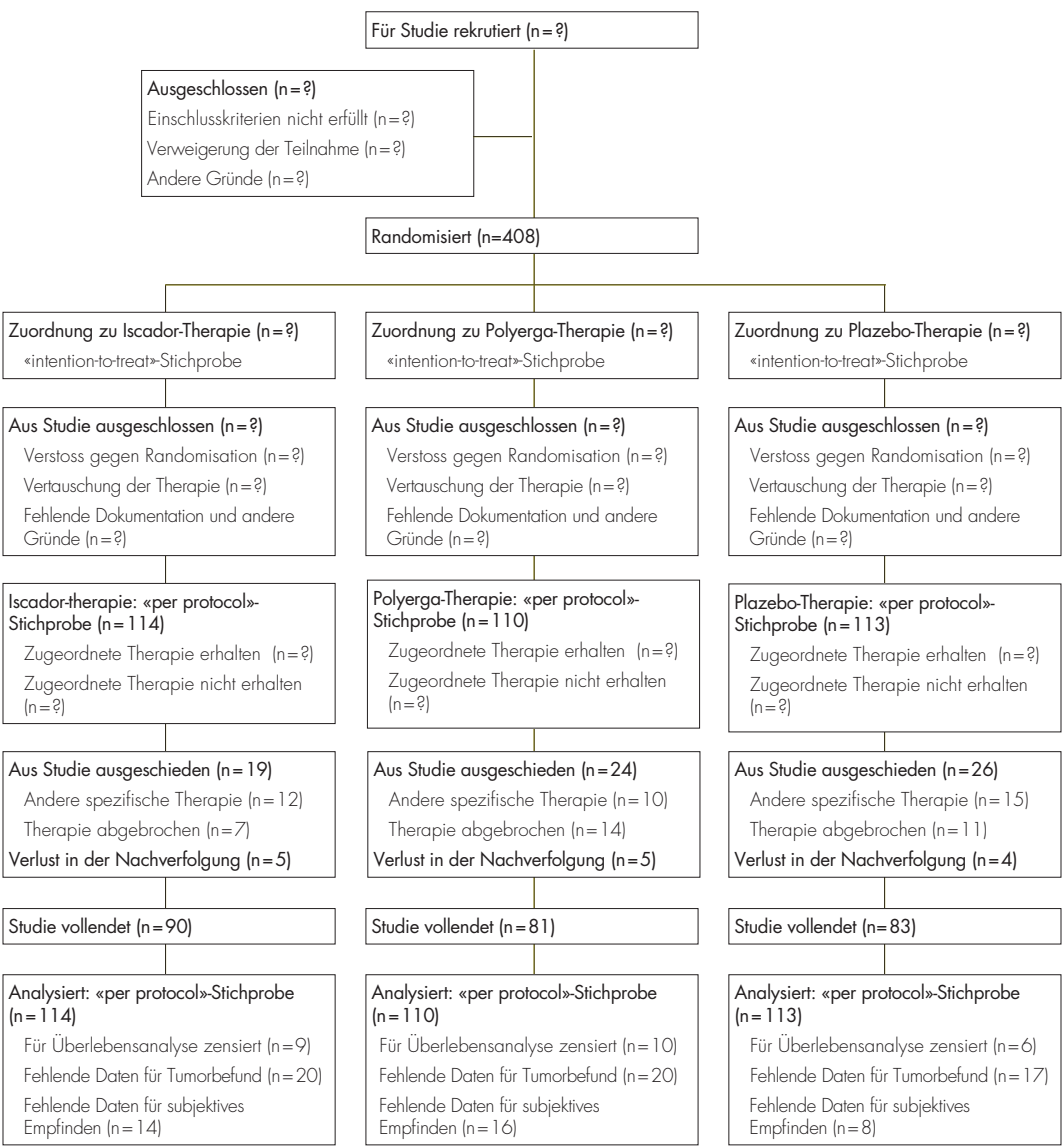
Die Studie war als eine multizentrische dreiarmlige randomisierte unverblindete Studie zur Prüfung der Wirksamkeit von Iscador und Polyerga (Antitumorglucosamin) im Vergleich mit Placebo (Multivitaminpräparat BVK Roche) geplant. Als Zielparame-ter der Wirksamkeit dienten Tumorremissionen, subjektives Empfinden (durch behandelnden Arzt erfragte Lebensqualität) und Überlebenszeit. Als Patienten kamen inoperable Fälle mit nichtkleinzelligem Bronchialkarzinom ohne Indikation für initiale Radiotherapie oder Chemotherapie in Frage, mit histologisch gesicherter Diagnose und schriftlichem Einverständnis zur Studienteilnahme.

Patientenrekrutierung und Studienverlauf

Die Patienten wurden in 4 Kliniken in Deutschland rekrutiert. Über die Anzahl der angefragten Patienten ist nichts bekannt. Die ursprünglich vorgesehene Fallzahl von $n = 560$ musste wegen Rekrutierungsproblemen (S. 8: «Schwierigkeiten bei der Auswahl von für die Studie geeigneter und kooperationsbereiter Patienten») zweimal nach unten korrigiert werden auf 300 bis 350 Patienten. Schlussendlich wurden 408 Patienten in die Studie rekrutiert und randomisiert. Direkt nach der Randomisation mussten 71 Patienten ausgeschlossen werden, unter anderem wegen Non-compliance gegen die zugewiesene Therapie (32) und Zuweisungsfehlern in der Therapie (20), fehlender Dokumentation (8) sowie weiterer Probleme (11). Damit blieben insgesamt 337 Patienten mit ordnungsgemäßer Studiendurchführung («per protocol»-Stichprobe, siehe Figur 3). Wie sich diese ausgeschlossenen Patienten auf die drei Studienarme verteilten, wurde nicht berichtet. Diese reduzierte Patientenzahl wurde der endgültigen biometrischen Auswertung zugrunde gelegt. Dies ist ein massiver Verstoss gegen die Prinzipien der Auswertung und Darstellung einer randomisierten klinischen Studie (Windeler 1993, Moher et al. 2001), da sie die durch die Randomisation erzwungene Strukturgleichheit der Therapiegruppen zu Studienbeginn («intention-to-treat»-Stichprobe) zerstört.

Grosse Probleme bereitete die lange Studiendauer wegen der ungewöhnlich langsamen Rekrutierungsrate. Dadurch kam es zu ver-

schiedenen Protokollverletzungen. Die Fortschritte der konventionellen Medizin führten zu diversen Therapiewechseln (insgesamt 37) und damit zum Ausscheiden aus der Studie (aber nicht aus der Auswertung). Zum Abbruch der Therapie kam es in total 32 Fällen, insbesondere wegen Tumorprogressionen oder dem Auftreten von Metastasen. Weitere 14 Patienten konnten nicht mehr nachverfolgt werden («drop-outs») (für die Verteilung dieser Zahlen auf die Studienarme, siehe Figur 3).



Figur 3: Überprüfung der Vollständigkeit eines Studienberichts: Flussdiagramm der Patienten und deren Status im Studienverlauf in den drei Behandlungsgruppen (Iscador, Polyerga, Placebo=Polyvitaminpräparat) in einer klinischen Studie des nichtoperablen nichtkleinzelligen Bronchialkarzinoms (Daten aus Dold et al. 1991).

Resultate

In der Überlebenszeit zeigte sich kein signifikanter Vorteil von Iscador gegenüber der Kontrollbehandlung, nur ein Trend (Median: 9.1 vs. 7.6 Monate, $p=0.24$). Beim Tumorbefund (nach dem Urteil des behandelnden Arztes) war die Iscador-Gruppe im Vergleich zu Placebo im Trend besser (einseitiger Test, $p=0.08$), ebenfalls bei den Remissionen (einseitiger Test, $p=0.10$); beim subjektiven Empfinden zeigte sich ein signifikant besserer Wert (einseitiger Test, $p=0.018$).

Hinsichtlich der Nebenwirkungen kam es in der Iscador-Gruppe nur zu den üblichen kurzfristigen Rötungen und Schmerzen in der Injektionsgegend für 1 bis 2 Tage und manchmal zu kurzfristigen Fieberreaktionen. Das Medikament musste in keinem Fall wegen Unverträglichkeit abgesetzt werden.

Diskussion

Gemäss der Analyse von Kienle und Kiene (2003, S.455–456; siehe auch Kienle und Kiene 2004c) produzierte diese Studie höchstwahrscheinlich falsch-negative Ergebnisse bezüglich Überlebenszeit und Tumorbefund. Denn in der Placebo-Gruppe gab es im wesentlichen ebenso häufig Tumorremissionen wie in den beiden Behandlungsgruppen. Es liegt nahe, hier eine Nivellierung des Studienergebnisses aufgrund verschiedener Ko-Interventionen in der Placebogruppe zu vermuten (Behandlungsbias, «performance bias»). In diesem Zusammenhang auffällig an dieser Studie ist die nur summarische Dokumentation der Ko-Interventionen («performance bias»), über deren Resultate darüber hinaus in der Publikation nichts berichtet wird. Eine weitere Ursache für die Nivellierung der Studienergebnisse könnte in den häufigen Therapiewechseln und -abbrüchen in der Placebogruppe zu suchen sein.

Hinzu kommen Unsicherheiten bezüglich der Vergleichbarkeit der Therapiegruppen, da nur die «per protocol»-Stichprobe analysiert wurde. Ein statistischer Vergleich mit der «intention-to-treat»-Stichprobe fand nicht statt, es wurde also keine Sensitivitätsanalyse durchgeführt.

Insgesamt muss das Resultat dieser Studie mit vielen Fragezeichen versehen werden, was unter anderem auch durch die vielen offenen Positionen über den Status der Patienten in den einzelnen Behandlungsgruppen im Studienverlauf zum Ausdruck kommt (Figur 3). Die in der Originalpublikation vorhandenen Informationen geben darüber keine befriedigende Auskunft.

Verzerrungsquellen und Studiendesigns

Quellen für Verzerrungen (Bias) und Validität

Systematische Fehler (Bias) können bei klinischen Studien in verschiedenen Phasen ihrer Planung und Durchführung auftreten (Figur 4 und Kasten 3).

Die Minimierung von Zufallsfehlern sowie die Vermeidung von systematischen Fehlern im Design-Stadium einer klinischen Studie, deren sorgfältige Kontrolle während der Durchführung der Studie sowie angemessene Methoden zum Ausgleich dieser Fehler in der statistischen Auswertung gehören zu den Fundamenten einer validen Studie. Verschiedene Methoden zur Vermeidung dieser Verzerrungen sind in Tabelle 1 angeführt. Hinzu kommt die angemessene Berücksichtigung und Behandlung von Störgrößen (confounder, z.B. Alter: ältere Patienten sind in der Regel kranker in jeder Indikation) sowie die Präzision und Verlässlichkeit der Messungen.

Designs für vergleichende Studien

Von den vielen möglichen Varianten vergleichender Studien sollen die einfachsten Varianten vorwärts longitudinaler kontrollierter Studien, kurz die kontrollierten Kohortenstudien, herausgegriffen werden (siehe dazu die Begriffsbestimmungen in Kasten 2). Dabei wird vorausgesetzt, dass die jeweiligen Gruppen Teil der Studie sind und nicht aus anderen Studien oder Datenregistern herbeigezogen werden (also keine Studien mit externen oder historischen Kontrollgruppen). Studien mit einem (oder mehreren) Therapiewechseln, die sogenannten «cross-over»-Studien, werden hier nicht weiter diskutiert, da sie eine eigene Betrachtung notwendig machen und für die Anwendung von Mistelpräparaten, insbesondere von Iscador, in der Regel nicht in Betracht kommen.

Kontrollierte vorwärts longitudinale experimentelle klinische Studien mit prospektiver Datenerfassung

In einem kontrollierten vorwärts longitudinalen klinischen Versuch wird eine experimentelle Forschungsfrage im Zeitverlauf untersucht durch eine vergleichende Betrachtung verschiedener Gruppen aus derselben Population, die mit verschiedenen Interventionen behandelt werden. Es sind protokollierte Studien mit vorwärts longitudinaler Forschungsrichtung und mit einer prospektiven Datenerfassung. Sie umfassen vorgeplante Auswahlkriterien für die ins Auge gefasste Patientenpopulation, einen vorgeplanten Zuweisungsmechanismus für die Verteilung der Patienten auf die zu vergleichenden Interventionen (Therapien), vorgeplante Interventionen, definierte Zielgrößen und prognostische Variablen sowie ein vorgeplantes Analyseprozedere.

Basiert der Zuweisungsmechanismus auf einem Zufallsverfahren, so handelt es sich um eine *randomisierte Studie*. Der entscheidende *Vorteil* einer randomisierten Studie ist die (mit einer gewissen Wahrscheinlichkeit) garantierte statistische Vergleichbarkeit der Studiengruppen vor Interventionsbeginn. Dies vereinfacht sowohl die Analyse nach Abschluss der Studie wie auch die Interpretation der Ergebnisse, falls andere Verzerrungen ausgeschlossen werden können oder vernachlässigbar sind.

Zu den prinzipiellen Nachteilen von randomisierten Studien gehört die mangelhafte externe Validität, das heisst die fragliche Verallgemeinerbarkeit solcher Studien auf die Alltagspraxis und durchschnittliche Patientenpopulationen. Die externe Validität ist insbesondere beeinträchtigt durch stark standardisierte und fixierte, für den klinischen Alltag untypische Behandlungsschemata, die besondere Auswahl von Studienpatienten, Studienärzten und Studienzentren (meist Universitätsspitäler), kurze Studien-dauer und auf die Beobachtung des Therapieverlaufs reduzierte

Komponenten der Forschung	Stichprobe	Selektion der Interventionsgruppe	Anwendung der Intervention	Erfassung der Zielgrößen	Follow-up	Analyse
	<div>Intendierte Stichprobe</div> <div>Aktuelle Stichprobe</div> <div>Zielpopulation</div>	<div>Gruppe A</div> <div>Gruppe B</div>	<div>Intervention A</div> <div>Intervention B</div>	<div>Zielgrößen für Gruppe A</div> <div>Zielgrößen für Gruppe B</div>	<div>Follow-up</div> <div>Follow-up</div>	<div>Verbleibende Fälle in Gruppe A</div> <div>Verbleibende Fälle in Gruppe B</div>
Quellen für Verzerrung		Selektionsbias, «selection bias»	Behandlungsbias, «performance bias»	Erfassungsbias, «detection bias»	Verlustbias, «attrition bias», «follow-up bias»	«intention-to-treat», «per protocol»
Quellen der Validität	Externe Validität: unverzerrte Stichprobe	Interne Validität: Strategien zur Vermeidung oder Minimierung der Verzerrung im Designstadium und im Analysestadium Strategien zur Vermeidung von «confounding» und «effect modification» Strategien für den Umgang mit der Adäquatheit der Variablen und der Präzision der Messungen				

Figur 4: Beurteilung der Qualität klinischer Studien: Quellen von Verzerrungen und Validität in den verschiedenen Phasen einer klinischen Studie (nach Feinstein 1985).

Arzt-Patient-Verhältnisse und somit die Gefahr falsch-negativer Ergebnisse (Abel 1997, 1999; Heusser 1999; Kiene 2001). Die Verallgemeinerbarkeit randomisierter Studien wird immer wieder zurecht in Frage gestellt und deshalb ausdrücklich eine Ergänzung solcher Studien durch Datenbankstudien gefordert, um die Wirksamkeit und Sicherheit einer Therapie bei in solchen Studien nicht behandelten Patientenkategorien oder Indikationen zu prüfen (Dieppe et al. 2004).

Zu den praktischen Schwierigkeiten der Durchführung einer randomisierten Studie gehören (Jack et al. 1990, Black 1996, Ross et al. 1999, Prescott et al. 1999, Gerhard et al. 2004): Therapiepräferenzen von Ärzten und Patienten, welche die Rekrutierung stark beeinträchtigen, ferner selbstbewusste und entscheidungsfreudige Patient(inn)en, ausserhalb der vorgesehenen Intervention erhältliche Therapie- und Präparatevielfalt und Therapiegewohnheiten. Nicht zu vergessen sind die hohen Kosten einer solchen Studie, da es einer grossen Anzahl von Teilnehmern bedarf, bis es zum hinreichenden Nachweis der Wirksamkeit hinsichtlich einer bestimmten Zielgrösse und für eine bestimmte Indikation kommt. (Für alternative Designs randomisierter Studien, welche einige dieser Probleme abschwächen könnten, siehe etwa Gaus 2004.)

Insbesondere zur Vermeidung oder Minimierung des Behandlungsbias und des Erfassungsbias wird in randomisierten Studien die Maskierung, üblicherweise Verblindung genannt, von behandelnden Ärzten und Patienten hinsichtlich der angewandten Therapie verwendet. Dies setzt voraus, dass es eine zusätzliche Behandlung durch einen wirksamkeitsneutralen Stoff (Plazebo) gibt, der in allen Aspekten der therapeutischen Intervention, hier insbesondere

Selektionsbias («selection bias») tritt auf, wenn die Zielgrösse beeinflusst ist durch systematische Differenzen zwischen den Behandlungsgruppen aufgrund der Zuteilungsprozedur.

Behandlungsbias («performance bias»), das heisst Verzerrungen der Zielgrösse aufgrund unterschiedlicher Behandlungen der Studiengruppen tritt auf, wenn die Behandlung nicht genügend standardisiert oder standardisierbar ist, wenn die Behandlung nicht sachgerecht und unvollständig durchgeführt wird oder wenn zusätzliche Behandlungen ausser den im Studienprotokoll vorgesehenen appliziert werden (Ko-Interventionen).

Beobachtungsbias oder **Erfassungsbias** («detection bias») liegt vor, wenn die Zielgrössen in den verschiedenen Behandlungsgruppen auf verschiedene Arten gemessen werden.

Nachverfolgungsbias oder **Verlustbias** («attrition bias», «transfer bias», «follow-up bias»), tritt auf, wenn es zu Verzerrungen der Zielgrösse kommt aufgrund unterschiedlicher Anzahl von Austritten von Patienten aus der Studie während des Studienverlaufs in den verschiedenen Behandlungsgruppen.

Kasten 3: Systematische Fehler bei klinischen Studien.

Quellen für systematische Fehler	Studien-typ	Designphase	Analysenphase
Selektionsbias («selection bias»)	ES	Randomisation «allocation concealment» Stratifizierung	Multivariate Adjustierung der Strata
	BS	Stratifizierung Paarweises Matching	Multivariate Adjustierung Matched Pair Analyse Matching gemäss «propensity scores» Sensitivitätsanalysen
Behandlungsbias («performance bias»)	ES	Standardisierte Interventionen und Beschränkung der erlaubten Ko-Interventionen Maskierung/Verblindung der Patienten, Ärzte und Pflegenden	
	BS	Definition und Auflistung der möglichen Ko-Interventionen Messungen aller Ko-Interventionen	
Beobachtungsbias, Erfassungsbias («detection bias»)	ES	Gut definierte und standardisierte Messung der Zielgrössen Maskierung/Verblindung der Personen, welche die Zielgrössen erfassen	
	BS	Gut definierte und standardisierte Messung der Zielgrössen	Multivariate Adjustierung für unverblindete Patienten
Nachverfolgungsbias, Verlustbias («attrition bias», «follow-up bias»)	ES,BS	Komplettes «follow-up» Erfassung der Gründe für Studienaustritt Wohldefinierte Patientenpopulation mit gutem Zugang zu Informationen über deren Zustand Population von Individuen mit überdurchschnittlichem Krankheitszustand oder -risiko (solche sind eher an einem «follow-up» interessiert)	
Analysenbias	ES		«intention-to-treat»-Analyse
	BS		Vergleich von «per protocol»- und «intention-to-treat»-Analyse Sensitivitätsanalysen Ersetzung von fehlenden Daten durch angemessene Schätzungen

Tabelle 1: Fehlerquellen und deren Verhinderung/Behebung in klinischen Studien: Strategien zur Minimierung von systematischen Fehlern (Verzerrung, Bias) und Maximierung der internen Validität in kontrollierten experimentellen Studien (ES) oder beobachtenden Studien (BS) in der Designphase und der Analysephase.

der subkutanen Injektion von Iscador, äquivalent ist, ausser natürlich im spezifischen Inhaltsstoff auf Mistelbasis. Aufgrund der bekannten Lokalreaktionen bei Iscador tritt in den meisten Fällen bei gut orientierten Patienten und erfahrenen Ärzten eine Entblindung ein, da die bisher eingesetzten Placebos nicht hinreichend vergleichbare Reaktionen zeigten wie die Iscador-Therapie (siehe dazu Rostock und Huber 2004). Ausserdem ist es verständlicherweise schwierig, krebserkrankte Patienten davon zu überzeugen, sich selbst über eine längere Zeit regelmässig zwei bis dreimal wöchentlich zu spritzen, wenn damit gerechnet werden muss, dass die Ampullen keinerlei wirksame Stoffe enthalten.

Ein *nichtrandomisierter* kontrollierter klinischer Versuch ist eine protokollierte Studie auf der Grundlage einer experimentellen Forschungsfrage mit vorwärts longitudinaler Forschungsrichtung und prospektiver Datenerfassung, in welcher vordefinierte Therapien vergleichend untersucht werden, die allerdings durch die Ärzte und/oder Patienten frei gewählt werden können. Auch hier ist natürlich das Ziel, die Studienergebnisse, das heisst die Veränderungen der Zielgrösse (z.B. progressionsfreies Intervall, Lebensqualität, Überlebenszeit) auf die unterschiedlichen Interventionen in den verschiedenen Gruppen zurückzuführen. Dies ist in der Regel kein Problem, wenn der Effekt genügend stark ist. Die Vorteile einer solchen Studie liegen neben der Möglichkeit des Vergleichs in der prospektiven Datenerfassung mit einem direkten Einfluss auf die Datenquantität und -qualität, die Praxisnähe der vordefinierten Behandlungsmethoden, die bestenfalls repräsentative Patientenauswahl (weitgefasste Einschlusskriterien, wenig Ausschlusskriterien).

Der Hauptnachteil solcher vergleichender Studien ist die Schwierigkeit bei der Abschätzung der Baseline-Vergleichbarkeit (Strukturgleichheit der Behandlungsgruppen vor Behandlungsbeginn) als Grundlage zur Interpretation der Veränderung der Zielgrösse als Therapieeffekt, das heisst die Abschätzung der Auswirkungen des Selektionsbias. Zu diesem Zweck muss eine möglichst vollständige Erfassung aller in Frage kommenden prognostischen Faktoren durchgeführt werden, die dann für eine eventuelle Stratifizierung (z.B. Gliederung der rekrutierten Patienten und der statistischen Auswertung nach Tumorstadien) und statistische Adjustierung (rechnerischer Ausgleich von Strukturungleichheiten) verwendet werden können. Hinzu kommt, dass die statistischen Methoden nicht in allen Fällen eine angemessene Adjustierung garantieren können, also immer ein nicht eliminierbarer (systematischer) Rest von Gruppenungleichheit zurückbleibt. Dies hat zur Folge, dass die Interpretation solcher Studien, insbesondere bei nicht besonders starken Effekten, problematisch ist (siehe Deeks et al. 2003, Reeves 2004b). Einem eventuell fraglichen Nutzen einer solchen

Studie stehen die hohen Kosten einer prognostischen Datenerfassung gegenüber, die, wegen der notwendigen Berücksichtigung aller prognostischen Faktoren und des eventuell langjährigen Zeithorizontes, sehr aufwändig sein kann und in der Regel aufwändiger ist als ein randomisierter klinischer Versuch. Auf jeden Fall muss eine solche Studie mit denselben Qualitätsstandards wie eine randomisierte Studie durchgeführt werden (siehe dazu Reeves und Gaus 2004). Aus den genannten Gründen (Nutzen vs. Kosten) kann es sinnvoll sein, das weiter unten erwähnte und in dem anschliessenden Aufsatz (Ziegler 2004d) ausführlicher dargestellte *vorwärts longitudinale Studiendesign mit retrospektiver Datenerfassung* (auch «retrolektive Studie» genannt) in Betracht zu ziehen.

Kontrollierte vorwärts longitudinale Beobachtungsstudien (epidemiologische Kohortenstudien)

In einer kontrollierten vorwärts longitudinalen Beobachtungsstudie, auch epidemiologische Kohortenstudie genannt, wird eine beobachtende Forschungsfrage durch eine vergleichende Betrachtung verschiedener Expositionen in verschiedenen Gruppen aus einer Population im Zeitverlauf untersucht. Die Studie hat vordefinierte Selektionskriterien für die Population, die Expositionen, die prognostischen Faktoren und die Zielgrössen. Ausserdem ist in einem Protokoll das Analyseverfahren festgelegt.

Eine solche Studie kann nun eine prospektive oder eine retrospektive Datenerfassung haben (siehe Figur 1). Im ersteren Fall handelt es sich um eine klassische nichtrandomisierte Beobachtungsstudie oder vergleichende prospektive epidemiologische Kohortenstudie. Im letzteren Fall um eine sogenannte «retrolektive» Studie oder eine vergleichende retrospektive epidemiologische Kohortenstudie. Auf die Besonderheiten solcher vorwärts longitudinaler Studien mit retrospektiver Datenerfassung wird im folgenden Aufsatz näher eingegangen (Ziegler 2004d).

Die oben angeführten Schwierigkeiten nichtrandomisierter vergleichender experimenteller Studien gelten auch für prospektive Beobachtungsstudien, sodass sich auch hier die Frage nach dem Kosten-Nutzen-Verhältnis stellt. Es ist offensichtlich, dass man mit vergleichenden Beobachtungsstudien einen guten Einblick in die tatsächliche Behandlungspraxis (Beobachtungen des tatsächlichen individualisierten Behandlungsverlaufs) gewinnen kann. Auf der anderen Seite steht diese Vielfalt in der Stichprobe der Studienpatienten und der therapeutischen Massnahmen einer eindeutigen Interpretation eines statistisch geschätzten Einflusses der letzteren auf die Zielgrösse (etwa dem progressionsfreien Intervall oder den Nebenwirkungen der onkologischen Basistherapie) im Sinne einer Wirkung der Prüftherapie (Iscaador-Behandlung) im Wege. Es ist

nicht leicht, abzuschätzen, ob nun die Veränderungen in der Zielgrösse, falls vorhanden, tatsächlich auf die unterschiedliche Behandlung zurückzuführen sind oder auf unterschiedliche Ausgangsbedingungen sowie unterschiedliche Krankheitsverläufe in der Therapie- und Kontrollgruppe.

Übersicht zu vergleichenden Studien mit Iscador seit 1990

Fragestellung und Datenerfassung

In der nachfolgenden Übersicht werden alle seit 1990 durchgeführten und vollständig publizierten prospektiven vergleichenden klinischen Studien mit einer Iscador-Therapiegruppe und einer Kontrollgruppe dargestellt und bewertet. Ich stütze mich dabei auf die neusten und vollständigsten Arbeiten zur systematischen Erfassung und Bewertung von Mistelstudien (Kienle et al. 2003a, siehe auch Kienle und Kiene 2004b, 2004c).

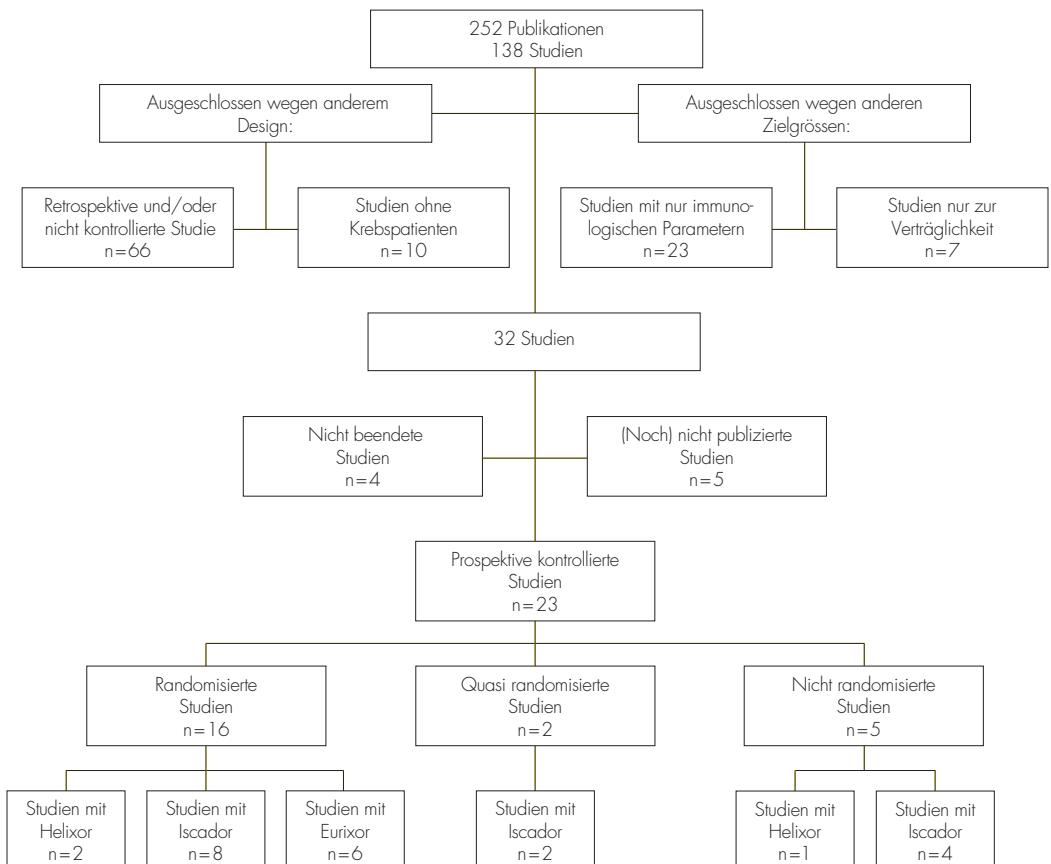
Fragestellung der Übersicht von Kienle et al. 2003a: Zeigen prospektiv kontrollierte Mistelstudien verlässliche Wirkungen auf das Überleben, auf die Tumorremission oder auf die Lebensqualität von Krebspatienten? Das alle Mistelstudien umfassende systematische Review beruhte auf folgender *Suchstrategie*: Es wurden alle bekannten Datenbanken wie MEDLINE, EMBASE, CANCERLIT, COCHRANE CONTROLLED TRIAL REGISTER, MEDIKAT, AMED, BIOS etc. bis zum Stand im Dezember 2002 durchsucht anhand folgender Stichwörter: mistletoe, viscum, Mistel, Misteltherapie, Mistelextrakt, und alle Handelsnamen für Mistelextrakte. Darüber hinaus wurden die Referenzen der gefundenen Publikationen geprüft sowie Experten und Hersteller von Mistelpräparaten für zusätzliche Informationen angefragt. Die *Auswahlkriterien* enthielten folgende *Einschlusskriterien*: (1) Prospektive kontrollierte klinische Studien mit Randomisierung und ohne Randomisierung; (2) Studienpopulation mit Krebs, inklusive «cervical intraepithelial neoplasia» (CIN); (3) Interventionsgruppe wurde mit Mistelpräparaten behandelt; (4) Messung von klinisch relevanten Zielparametern (z. B. Gesamtüberleben, krankheitsfreies Überleben, Remission, Zeit bis Relaps, Lebensqualität QoL, Reduktion von Nebenwirkungen während Chemotherapie); (5) Studie beendet bis Dezember 2002; (6) Publikation als Manuskript oder Abstrakt oder Konferenz-Report; (7) Alle Sprachen. Dabei galten folgende *Ausschlusskriterien*: Reine Toxizitäts- und Verträglichkeitsstudien (Phase I), ausschliessliche Messung von immunologischen Parametern oder Probanden ohne Krebserkrankung.

Von den ursprünglich gefundenen 138 Studien in 252 Publikationen blieben nach konsequenter Anwendung der Auswahlkriterien nur 23 Studien übrig (siehe Figur 5). Davon gab es 14 Studien mit Iscador. Für die vorliegende Darstellung wurden noch weitere

Einschränkungen vorgenommen: Vollständige Publikation ab 1990, Zielgrösse («outcome») Überlebenszeit und Lebensqualität, keine Berücksichtigung der Tumorresponse. Am Schluss blieben nur noch 6 Studien übrig, 4 randomisierte Studien (Tabellen 2 und 3) und 2 nichtrandomisierte Studien (Tabellen 4 und 5).

Qualitätsbeurteilung

Die *Qualitätsbeurteilung* ist ein entscheidender Bestandteil dieser Übersicht und beruht im wesentlichen auf einer Analyse der weiter oben dargestellten Verzerrungsquellen sowie einiger zusätzlicher Punkte, welche vor allem die Qualität der Berichterstattung und die Datenqualität betreffen. Die *Qualitätskriterien* wurden vor Beginn der Beurteilung der Studien aufgestellt (Kasten 4) und die



Figur 5: Auswahl von Studien für einen systematischen Review klinischer Studien mit Mistelpräparaten bis 2002: Auswahlverfahren für prospektive kontrollierte Studien mit Mistelpräparaten von der Literaturrecherche bis zur Erfüllung aller Einschlusskriterien (Daten aus Kienle et al. 2003a; bei quasi-randomisierten Studien werden systematische Verfahren der Zuteilung in die Behandlungsgruppen verwendet, wie etwa alternierende Zuweisungen im Falle zweier Gruppen).

Analyse wurde von zwei unabhängig arbeitenden Teams durchgeführt. Es ergab keine wesentlichen Differenzen in der Studienbeurteilung; allfällige Unterschiede wurden durch Diskussion bereinigt. Die *Qualitätskriterien* selbst wurden übernommen vom «National Health Service Center for Reviews and Dissemination» (Khan et al. 2001) und aus einem früheren Review von Mistelstudien (Kleijnen und Knipschild 1994).

Das *Rating* wurde nach folgenden vier Kategorien eingeteilt:
 + hinreichend erfüllt, (+) teilweise erfüllt, (-) wenig erfüllt, (-) nicht erfüllt.

Diskussion

Die hauptsächlichen Resultate der randomisierten und nichtrandomisierten Studien von Grossarth-Maticsek finden sich in den Tabellen 2 und 4 und die dazugehörige Qualitätsbeurteilung in den Tabellen 3 und 5. Für eine ausführlichere Darstellung und Diskussion dieser Studien verweise ich auf das vorangehende Heft von Mistelsteinn (Grossarth-Maticsek et al. 2003) oder auf die Originalliteratur (Grossarth-Maticsek et al. 2001a, 2001b, 2001c, 2004). In diesem Zusammenhang soll nur festgehalten werden, dass sich diese Studien den Zielgrössen der langfristigen Überlebenszeit und der kurzfristigen Veränderung Lebensqualität (Selbstregulation) widmen, nicht aber Fragen der Verträglichkeit und Sicherheit. Eine entscheidende Innovation dieser Studien besteht darin, dass

- A. Massnahmen gegen Selektionsbias («selection bias»), insbesondere Randomisation
- B. Minimierung der Heterogenität durch Prästratifikation oder Matching
- C. Massnahmen gegen Beobachtungsbias («observer bias», «detection bias»), insbesondere durch Verblindung (Maskierung) von Patienten, Prüfärzten, «outcome assessor»
- D. Massnahmen gegen Behandlungsbias («performance bias») durch Standardisierung der Behandlung, Dokumentation aller Ko-Interventionen, Verblindung von Patienten und Prüfärzten
- E. Massnahmen gegen Messfehler durch Standardisierung der Messung der Zielgrösse («outcome assessment»)
- F. Massnahmen gegen Verlustbias («attrition bias», «follow-up bias»): weniger als 10% Datenverlust, Auswertung nach «intention-to-treat» und «per protocol», Sensitivitätsanalysen
- G. Effektmessungen relevant und gut beschrieben
- H. Hinreichend präzise Beschreibung der Interventionen, der Patientencharakteristiken, der Krankheit und der früheren Therapien
- I. Klare Beschreibung des Studiendesigns
- J. Präzise Darstellung der Resultate
- K. Datenqualität durch GCP-ICH-Richtlinien garantiert, insbesondere durch Monitoring

Kasten 4: Qualitätskriterien zur Beurteilung klinischer Studien (nach Kienle et al. 2003a).

innerhalb desselben Kollektivs randomisierte und nichtrandomisierte Studien durchgeführt wurden, wodurch der Stellenwert sowohl der randomisierten Studien (hinsichtlich externer Validität oder Generalisierbarkeit) als auch der nichtrandomisierten Studien (hinsichtlich Validität bei gleichgerichteten Studienresultaten) erheblich verbessert werden kann. Unerwünschte Nebenwirkungen wurden in dieser Studie nicht dokumentiert.

Studie	Outcome	Lokalisation	Stadium	n	Resultat für Iscadorgruppe	Intervention
Grossarth 2001a	Mittleres Überleben (57.5 vs. 28.9 Monate)	Brust	IIIA, IIIB	34	Signifikant länger	Standard + Iscador (17) Standard (17)
Dold 1991	Medianes Überleben (9.1 vs. 7.6 Monate)	Lunge	alle Stadien	408	Trend zur Verlängerung	Iscador (114) Vitamine (113)
	Zunahme der Lebensqualität (59% vs. 45%)	Lunge	alle Stadien	370	Signifikante Zunahme	Iscador (114) Vitamine (113)
Grossarth 2001a	Mittleres Überleben (42 vs. 29 Monate)	Brust, Lunge, Rektum, Kolon, Magen	alle Stadien	78	Signifikant länger	Standard + Iscador (39) Standard (39)
Salzer 1991	Medianes Überleben (33 vs. 31 Monate)	Lunge	I bis IV	210	Trend zur Verlängerung	Operation + Iscador (87) Operation (96)

Tabelle 2: Resultate von klinischen Studien mit Iscador: Prospektive vergleichende randomisierte Studien seit 1990

Studie	Qualitätskriterien											n	Verlustrate insgesamt
	A	B	C	D	E	F	G	H	I	J	K		
Grossarth 2001a	+	+	-	(-)	+	+	+	(-)	+	+	-	34	0 %
Dold 1991	+	+	-	-	+	(-)	+	(+)	+	+	(-)	408	17 %
Grossarth 2001a	+	+	-	(-)	+	(-)	+	(-)	+	+	-	78	20 %
Salzer 1991	+	(+)	-	(-)	(+)	(-)	+	(+)	(+)	+	-	210	16 %

Tabelle 3: Qualitätsbeurteilung von klinischen Studien mit Iscador: Qualität der prospektiv vergleichenden randomisierten Studien mit Iscador von Tabelle 2 (nach Kienle et al. 2003a).

Die Resultate der randomisierten Studie von Salzer et al. 1991 finden sich in Tabelle 2 und die Qualitätsbeurteilung in Tabelle 3. Für eine ausführliche Diskussion dieser Studie verweise ich auf Kienle und Kiene 2003b (S. 538–539). Die Studie von Dold et al. 1991 wurde weiter oben bereits ausführlich dargestellt und diskutiert.

Die nichtrandomisierte Studie von Schuppli 1990 zu Melanom-Patienten wird im anschliessenden Aufsatz näher besprochen (Ziegler 2004d). Auf die im systematischen Review Kienle et al. 2003a noch nicht enthaltene Studie Kleeberg et al. 2004 wird ebenfalls dort eingegangen.

Bezüglich der Studienqualität ist auffällig, dass die meisten Studien Qualitätsmängel aufwiesen, die sich durch hinreichendes Problembewusstsein im Prinzip hätten überwinden lassen. Es gibt keine in der Natur von klinischen Studien mit Iscador liegende Sachverhalte, welche diese Mängel begründen. So gab es insbesondere kaum Anstrengungen, um den Beobachtungs- oder Erfassungsbias («detection bias») zu minimieren oder durch vollständige Dokumentation der Ko-Interventionen den Behandlungsbias («performance bias») zu berücksichtigen. Wie bereits weiter oben

Studie	Outcome	Lokalisation	Stadium	n	Resultat	Präparat
Grossarth 2001a	Mittleres Überleben (50.8 vs. 36.6 Monate)	Brust, Lunge, Rektum, Kolon, Magen	alle Stadien	792	Signifikant länger	Standard + Iscador (396) Standard (396)
Schuppli 1990	5-Jahres Überleben (86 vs. 72%)	Melanom	keine Angaben	198	Trend zur Verlängerung	Operation + Iscador (84) Operation + BCG (114)

Tabelle 4: Resultate von klinischen Studien mit Iscador: Prospektive vergleichende nichtrandomisierte Beobachtungsstudien mit Iscador seit 1990

Studie	Qualitätskriterien											n	Verlustrate insgesamt
	A	B	C	D	E	F	G	H	I	J	K		
Grossarth 2001a	(+)	+	-	(-)	+	+	+	-	+	+	-	792	3.5%
Schuppli 1990	(+)	-	-	(-)	+	(-)	+	(-)	-	-	-	198	-

Tabelle 5: Qualitätsbeurteilung von klinischen Studien mit Iscador: Qualität der prospektiven vergleichenden nicht-randomisierten Beobachtungsstudien mit Iscador von Tabelle 4 (nach Kienle et al. 2003a).

erwähnt, kommt im Zusammenhang einer Misteltherapie eine Verblindung in der Regel nicht in Frage, sodass das Fehlen der Verblindung nicht als Qualitätsmangel einer Mistelstudie eingestuft werden muss. Ein weiteres, allerdings nicht so leicht überwindbares Problem ist der Datenverlust; aber auch hier sind sicher Verbesserungen möglich, insbesondere auch bei der Berichterstattung. Kaum zu entschuldigen sind die häufig mangelhaften Berichterstattungen über Patientencharakteristiken und die Details der Intervention (Iscador-Sorten, Therapieverlauf, Therapiepausen etc.).

Für alle bisherigen Studien gab es keine Kontrolle der Datenqualität gemäss den Richtlinien der «Good Clinical Practice» (GCP), was natürlich zum Teil mit dem Beginn dieser Studien in den 70er und 80er Jahren zu tun hat. Ausserdem sind bei diesen klinischen Studien die Informationen zur Verträglichkeit und Sicherheit der Iscador-Therapie sehr spärlich (siehe dazu die Übersicht in Stein und Berg 2000).

Ein vollständiges Flussdiagramm über den Rekrutierungsprozess und den genauen Status der Patienten während des Ablaufs der Studie, aufgeschlüsselt nach allen Therapiegruppen, gemäss den Richtlinien der CONSORT-Arbeitsgruppe (Moher et al. 2001) lässt sich aus keiner dieser Studien gewinnen.

Schlussfolgerung

Die Ergebnisse vergleichender klinischer Studien mit Iscador sind in der Regel ermutigend, wenn auch noch nicht hinreichend qualitativ und quantitativ abgesichert. Vom Gesichtspunkt konventioneller Studiendesigns sind dies die am besten abgesicherten Ergebnisse.

Die einzigen Studien zur langfristigen Wirksamkeit sind diejenigen von Grossarth-Maticsek. Langfristige und vergleichende *klinische Studien* zur Lebensqualität und zur Verträglichkeit und Sicherheit der Iscador-Therapie gibt es bisher nicht. Allerdings gibt es von Seiten der Ärzte und der Patienten jahrzehntelange positive Erfahrungen mit dieser Therapie. Unter diesem Gesichtspunkt erscheinen die genannten Studienergebnisse in einem ganz anderen Licht: Sie reflektieren in ihren Ergebnissen nicht nur die blosse Anwendung von Iscador, sondern sind auch Ausdruck des breiten Umfeldes, innerhalb welchem die Anwendung von Iscador zu meist positiven Auswirkungen mit wenig störenden Nebenwirkungen gekommen ist. Trotzdem muss festgehalten werden, dass insgesamt bezüglich aller genannten Studienziele weitere Studien von hoher Qualität und Aussagekraft sowie mit langfristigem Horizont vonnöten sind. Dies betrifft sowohl vergleichende Studien (Effektforschung) als auch auf hohem Standard dokumentierte Einzelfälle und Fallserien (Bedingungsforschung).

Weitere Studien auf hohem Qualitätsstandard sind vor allem aus zwei Gründen wichtig. Erstens muss dafür gesorgt werden, dass Mistelpräparate im allgemeinen und Iscador im besonderen weiterhin öffentlich zugänglich bleiben, das heisst es muss gegenüber Behörden und Krankenkassen mit den üblichen Methoden gezeigt werden, dass die Iscador-Therapie langfristig verträglich, sicher und wirksam ist. Es muss in diesem Zusammenhange auch daran gearbeitet werden, verschiedene Forschungsdesigns und -strategien zu verfolgen, die einander ergänzen (siehe dazu Heusser 2001, van Haselen 2004). Von einer bisher noch wenig ausgeschöpften Möglichkeit, in gewissen Grenzen verlässliche Studienresultate mit grossen Patientenzahlen zu bekommen, soll in der nachfolgenden Arbeit berichtet werden.

Wie am Anfang dieses Aufsatzes erwähnt, hat diese klinische Forschung zweitens aber auch die Aufgabe, das in der individuellen klinischen und ambulanten ärztlichen Erfahrung verankerte und in Falldarstellungen und Fallserien teilweise dokumentierte Vertrauen in die Sicherheit und Wirksamkeit, ja Notwendigkeit der Misteltherapie im allgemeinen und der Iscador-Therapie im besonderen zu stärken und weiter in die Zukunft zu tragen.

Literatur

- Abel U. (1997) Randomisation in klinischen Studien. Empirisch begründet oder nur ein Dogma? *Internist* **38**, S. 318–324.
- Abel U. (1999) The role of randomization in clinical studies: myths and beliefs. *Journal of Clinical Epidemiology* **52** (6), S. 487–497.
- Black N. (1996) Why we need observational studies to evaluate the effectiveness of health care. *British Medical Journal* **312**, S. 1215–1218.
- Deeks J. J., Dinnes J., D’Amico R., Sowden A., Sakarovich C., Song F., Petticrew M., Altman D. G. (2003) Evaluating non-randomised intervention studies. *Health Technology Assessment* **7** (27), S. 1–173.
- Dieppe P., Bartlett C., Davey P., Doyal L., Ebrahim S. (2004) Balancing benefits and harms: the example of non-steroidal anti-inflammatory drugs. *British Medical Journal* **329**, S. 31–34.
- Dold U., Edler L., Mäurer C. M., Müller-Wening D., Sakellariou B., Trendelenburg F., Wagner G. (1991) *Krebszusatztherapie beim fortgeschrittenen nicht-kleinzelligen Bronchialkarzinom – Multizentrische kontrollierte Studie zur Prüfung der Wirksamkeit von Iscador und Polyerga*. Stuttgart, Thieme.
- Feinstein A. (1985) *Clinical Epidemiology: The Architecture of Clinical Research*. Philadelphia, W. B. Saunders Company.
- Gaus W (2004) Randomized Trials versus Observational Studies: Is the Comparative Observational Study a Poor Compromise or a Perfect Synthesis? *Forschende Komplementärmedizin* **11** (suppl 1), S. 3–4.
- Gerhard I., Abel U., Loewe-Mesch A., Huppmann S., Kuehn J. J. (2004) Problematik randomisierter Studien in der Komplementärmedizin dargestellt am Beispiel der Misteltherapie bei Patientinnen mit Mammakarzinom. *Forschende Komplementärmedizin* **11** (3), S. 150–157.
- Grossarth-Maticek R., Kiene H., Baumgartner S., Ziegler R. (2001a) Use of Iscador, an extract of European mistletoe (*Viscum album*), in cancer treatment: prospective nonrandomized and randomized matched-pairs studies nested within a cohort study. *Alternative Therapies in Health and Medicine* **7** (3), S. 57–78.
- Grossarth-Maticek R., Kiene H., Baumgartner S., Ziegler R. (2001b) Addendum to Iscador article. *Alternative Therapies in Health and Medicine* **7** (4), S. 26.
- Grossarth-Maticek R., Kiene H., Baumgartner S., Ziegler R. (2001c) Verlängerung der Überlebenszeit von Krebspatienten unter Misteltherapie (Iscador) – Ergebnisse einer epidemiologischen Langzeitstudie [Teil I]. *Schweizerische Zeitschrift für Ganzheitsmedizin* **13** (4), S. 217–225.
- Grossarth-Maticek R., Baumgartner S., Ziegler R. (2003) Verlängerung der Überlebenszeit von Krebspatienten unter Misteltherapie (Iscador). *Mitteilenn* **4**, S. 52–67.
- Grossarth-Maticek R., Kiene H., Baumgartner S., Ziegler R. (2004) Synergieeffekte von Selbstregulation und Misteltherapie (Iscador) auf die Überlebenszeit bei Krebspatienten – Ergebnisse einer epidemiologischen Langzeitstudie, Teil II. *Schweizerische Zeitschrift für Ganzheitsmedizin* **16** (2), S. 81–89.

- Heusser P. (1999) Probleme von Studiendesigns mit Randomisation, Verblindung und Placebogabe. *Forschende Komplementärmedizin* **6**, S. 89–102.
- Heusser P. (2001) Kriterien zur Beurteilung des Nutzens von komplementärmedizinischen Methoden. *Forschende Komplementärmedizin* **8**, S. 14–23.
- Horneber M., Büschel G., Kaiser G., Kappauf H., Wilhelm M., Gallmeier W. M. (2003) Unkonventionelle Verfahren: Chancen für die Arzt-Patient-Beziehung. *Der Onkologe* **9** (12), S. 1335–1342.
- Hulley S. B., Cummings S. R., Browner W. S., Grady D., Hearst N., Newman T. B. (2001) *Designing Clinical Research: An Epidemiologic Approach*. Philadelphia, Lippincott Williams & Wilkins, 2001 (2nd edition).
- Jack W. J., Chetty U., Rodger A. (1990) Recruitment to a prospective breast conservation trial: Why are so few patients randomised? *British Medical Journal* **301**, S. 83–85.
- Khan K. S., ter Riet G., Glanville J., Sowden A. J., Kleijnen J. (2001) *Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for those Carrying Out or Commissioning Reviews*. York, University of York, NHS Centre for Reviews and Dissemination: CRD Report Number **4** (2nd edition).
- Kiene H. (2001) *Komplementäre Methodenlehre der klinischen Forschung: Cognition-based Medicine*. Springer, Berlin/Heidelberg.
- Kienle G. S., Kiene H. (2003a) Kriterien für die Erstellung therapeutischer Einzelfallberichte in der Onkologie. *Der Merkurstab* **56** (1), S. 2–5.
- Kienle G. S., Kiene H. (2003b) *Die Mistel in der Onkologie. Fakten und konzeptionelle Grundlagen*. Stuttgart, Schattauer.
- Kienle G. S., Kiene H. (2004a) Ärztliche Urteilskraft und Formalismus in der modernen Medizin. *Erfahrungsheilkunde* **53**, S. 127–133.
- Kienle G. S., Kiene H. (2004b) Klinische Studien zur Misteltherapie onkologischer Erkrankungen. *Deutsche Zeitschrift für Onkologie* **36**, S. 17–24.
- Kienle G. S., Kiene H. (2004c) Klinische Studien zur Misteltherapie der Krebserkrankung – Eine Übersicht. *Erfahrungsheilkunde* **53**, S. 193–209.
- Kienle G. S., Berrino R., Büssing A., Portalupi E., Rosenzweig S., Kiene H. (2003a) Mistletoe in cancer: A systematic review on controlled clinical trials. *European Journal of Medical Research* **8**, S. 109–119.
- Kienle G. S., Karutz M., Matthes H., Matthiessen P. F., Petersen P., Kiene H. (2003b) Evidenzbasierte Medizin: Konkurs der ärztlichen Urteilskraft? *Deutsches Ärzteblatt* **100** (33), S. A2142–A2146.
- Kienle G. S., Hamre H. J., Portalupi E., Kiene H. (2004) Improving the quality of therapeutic reports of single cases and case series in oncology – criteria and checklist. *Alternative Therapies in Health and Medicine* **10** (5), S. 68–72.
- Kleeberg U., Suciú S., Bröcker E B., Ruiter D. J., Chartier C., Liénard D., Marsden J., Schadendorf D., Eggermont A. M. for the EORTC Melanoma Group in cooperation with the German Cancer Society (DKG) (2004) Final results of the EORTC 18871/DKG 80-1 randomised phase III trial: rIFN- α 2b versus rIFN- γ versus

- ISCADOR M versus observation after surgery in melanoma patients with either high-risk primary (thickness > 3 mm) or regional lymph node metastasis. *European Journal of Cancer* **40**, S. 390–402.
- Kleijnen J., Knipschild P. (1994) Mistletoe treatment for cancer: review of controlled trials in humans. *Phytomedicine* **1**, S. 255–260.
- Moher D., Schulz K. F., Altman D. G. (2001) The CONSORT Statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Journal of the American Medical Association* **285**, S. 1987–1991.
- Porzolt F., Gaus W., Hart D. (1996) Der Therapieoptimierungsvergleich (TOPV): Ein Verfahren zum Vergleich der Qualität therapeutischer Massnahmen. *Forum Deutsche Krebsgesellschaft* (11), S. 88–95.
- Prescott R. J., Counsell C. E., Gillespie W. J., Grant A. M., Russell I. T., Kiauka S., Colthart I. R., Ross S., Shepherd S. M., Russell D. (1999) Factors that limit the quality, number and progress of randomised controlled trials. *Health Technology Assessment* **3** (20), S. 1–134.
- Reeves B. C. (2004a) A framework for classifying study designs to evaluate health care interventions. *Forschende Komplementärmedizin* **11** (suppl 1), S. 13–17.
- Reeves B. C. (2004b) Reasons for caution when evaluating health care interventions using non-randomised study designs. *Forschende Komplementärmedizin* **11** (suppl 1), S. 40–45.
- Reeves B. C., Gaus W. (2004) Guidelines for reporting non-randomised studies. *Forschende Komplementärmedizin* **11** (suppl 1), S. 46–52.
- Ross S., Grant A. M., Counsell C. E., Gillespie W. J., Russell I. T., Prescott R. J. (1999) Barriers to participation in randomised controlled trials: A systematic review. *Journal of Clinical Epidemiology* **52** (12), S. 1143–1156.
- Rostock M., Huber R. (2004) Randomized and double blind studies – demands and reality as demonstrated by two examples of mistletoe research. *Forschende Komplementärmedizin* **11** (suppl 1), S. 18–22.
- Sackett D. L., Richardson W. S., Rosenberg W., Haynes R. B. (1999) *Evidenzbasierte Medizin* (deutsche Ausgabe von R. Kunz und L. Fritzsche). München, Zuckschwerdt.
- Salzer G., Danmayr E., Wutzlhofer F., Frey S. (1991) Adjuvante Iscador-Behandlung operierter nicht kleinzelliger Bronchuskarzinome. Ergebnisse einer randomisierten Studie. *Deutsche Zeitschrift für Onkologie* **23** (4), S. 93–98.
- Schneider B. (2001) Analysis of therapeutic efficacy in observational cohort studies. *Cancer Chemotherapy and Pharmacology* **47** (Supplement), S. S35–S37.
- Schuppli R (1990) Die adjuvante Behandlung des malignen Melanoms mit Iscador c. Hg. In: Jungi W. F., Senn H. J. (Hrsg.), *Krebs in der Alternativmedizin, Band II*. Berlin: Springer, S. 84–87.
- Stein G. M., Berg P. A. (2000). Adverse effects during therapy with mistletoe extracts. In: Büssing A. (ed.), *Mistletoe – The Genus Viscum*. Amsterdam, Harwood Academic Publishers, S. 195–208.

- Steiner R. (1894/1918), *Die Philosophie der Freiheit*, Rudolf Steiner Verlag, Dornach ¹⁶1995.
- Steiner R. (1910) *Die Offenbarungen des Karma* (GA 120), Rudolf Steiner Verlag, Dornach ⁸1992.
- Steiner R. (1920) *Geisteswissenschaft und Medizin* (GA 312), Rudolf Steiner Verlag, Dornach ⁷1999.
- Van Haselen R. (2004) The importance of considering research design in a wider context. *Forschende Komplementärmedizin* **11** (suppl 1), S. 53–55.
- Windeler J. (1993) Das Intention-to-treat-Prinzip in klinischen Arzneimittelprüfungen. *Arzneimitteltherapie* **11** (4), S. 103–111.
- Ziegler R. (2004a) Zur historischen Entwicklung der klinischen Wirksamkeitsforschung. In: Fintelmann V. (Hrsg.), *Onkologie auf anthroposophischer Grundlage* (Mayer, Stuttgart), Abschnitt **6.1** (3. Lieferung), S. 1–31.
- Ziegler R. (2004b) Erkenntnismethodische und ethische Grundlagen der klinisch-therapeutischen Wirksamkeitsforschung. In: Fintelmann V. (Hrsg.), *Onkologie auf anthroposophischer Grundlage* (Mayer, Stuttgart), Abschnitt **6.2** (3. Lieferung), S. 1–49.
- Ziegler R. (2004c) Elements of Therapeutic Research Structure. *Forschende Komplementärmedizin* **11** (suppl 1), S. 5–12.
- Ziegler R. (2004d) Neuere klinische Studien mit Iscador zum Mammakarzinom und malignen Melanom. *Mitteiln* **5**, S. 74–99.