

Aus der klinischen Ökonomik, Universitätsklinikum Ulm

Leiter: Prof. Dr. med. Franz Porzsolt

**Spezifität der adjuvanten 5-FU/Leucovorin Chemotherapie
bei kolorektalem Karzinom**

Eine Medline©-Literaturrecherche randomisierter klinischer Studien zur
adjuvanten Chemotherapie bei kolorektalem Karzinom

Dissertation

zur Erlangung des Doktorgrades der Humanbiologie (Dr. biol. hum.)
an der Medizinischen Fakultät der Universität Ulm

vorgelegt von
Joachim Brade
aus Mannheim

2007

Mit Genehmigung der Medizinischen Fakultät
der Universität Ulm

Amtierender Dekan: Prof. Dr. Klaus-Michael Debatin

1. Berichterstatter: Prof. Dr. Franz Porzsolt

2. Berichterstatter: Prof. Dr. Marko Kornmann

Tag der Promotion: 21.01.2008

Betreuung durch die Leiterin der Abteilung für Medizinische Statistik an der
Medizinischen Fakultät Mannheim der Universität Heidelberg: Dr. sc. hum. Christel Weiß

Inhaltsverzeichnis

1. Einleitung	1
2. Krankheitsbild des kolorektalen Karzinoms: klinische Fakten.....	5
2.1 Symptomatik, Ätiologie und Diagnostik	5
2.2 Abgrenzung der adjuvanten Chemotherapie	6
2.3 Wirkungsweise der Chemotherapie.....	8
2.4 Therapieablauf und Nebenwirkungen	9
2.5 Empirische Evidenz zur Chemotherapie bei KKK , historischer Überblick	10
3. Material und Methoden	13
3.1 Medline©-Literaturrecherche	13
3.2 Selektion	14
3.3 Operationalisierung der Arbeitshypothese	15
3.4 Qualitative Beurteilung der Publikationen	16
4. Ergebnisse	18
4.1 Vergleich eines Chemotherapieprotokolls mit einer Nulltherapie	18
4.2 Vergleich eines Chemotherapieprotokolls mit einer alternativen aktiven Therapie	22
4.3 Qualitative Gesamtbetrachtung	26
5. Diskussion.....	29
5.1 Methodische Forderungen für kontrollierte klinische Studien	29
5.2 Methodische Durchführung klinischer Studien zur Überlebenszeit.....	31
5.2.1 Statistische Methoden.....	31
5.2.2 Messkonzept und Kausalität.....	39

5.2.3 Der p-Wert als mathematisch-statistisches Entscheidungskriterium.....	42
5.3 Wirkungsebenen einer medizinischen Therapie und Art der Vergleichsgruppe.....	45
5.4 Ganzheitliche Wirkungsweise einer medizinischen Intervention	47
5.5 Interaktion zwischen Arzt und Patient: Kommunikationstheoretische Aspekte	49
5.6 Art der Verblindung.....	50
5.7 Randomisierung.....	52
5.8 Multiple Kontrollgruppen.....	53
6. Folgerungen und Empfehlungen.....	54
6.1 Assay Sensitivity und Kontrollgruppenwahl.....	54
6.2 Add-on Studien.....	54
6.3 Behandlungsgleichheit, Randomisierung und Verblindung.....	55
6.4 Berichtswesen.....	55
6.5 Leitlinien.....	56
7. Zusammenfassung und Schlussbemerkungen	57
8. Quellenverzeichnis.....	61
8.1 Verzeichnis der im Text verwendeten Lehrbücher, Leitlinien und Datenbanken.....	61
8.2 Verzeichnis der im Text verwendeten Publikationen.....	63
9. Danksagung.....	69

Abkürzungsverzeichnis

AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
CIFU	Continuous Infusion of Fluorouracil
DNS	Desoxyribonukleinsäure
FA	Folinic Acid (Folinsäure)
FFCD	Fondation Française de Cancerologie Digestive
5-FU	5-Fluorouracil
GIVIO	Gruppo Interdisciplinare Valutazione
ICH	International Conference of Harmonization
IFN	Interferon
IMPACT	International Multicentre Pooled Analyses of Colon Cancer Trials
ITT	Intention to Treat Analysekollektiv
i.v.	intravenous, intravenös
KRK	Kolorektales Karzinom
LEV	Levamisole
LV	Leucovorin (Handelsbezeichnung für Folinsäure)
NCIC-CTG	National Cancer Institute of Canada Clinical Trials Group
NIH	National Institute of Health
NNT	Number Needed to Treat
PFÜZ	Progressionsfreie Überlebenszeit
PVI	Protracted Venous Infusion
PP	Per Protocol Analysekollektiv
RFÜZ	Rezidivfreie Überlebenszeit
s.c.	subcutane, subkutan
TNM	Tumour / Node / Metastasis (Klassifikationsschema)
UICC	Union Internationale Contre le Cancer (Klassifikationsschema)
UFT	Tegafur / Uracil
ÜZ	Gesamtüberlebenszeit

1. Einleitung

Im Vergleich zu den Anfängen der Chemotherapie Beginn der 50er Jahre hat die Anzahl möglicher Behandlungsvarianten erheblich zugenommen. Die Anzahl zugelassener Behandlungsmittel hat sich in diesem Zeitraum etwa verfünffacht [17]. Nach Maßgabe der Ergebnisse klinischer Studien sind die in Leitlinienkonferenzen entworfenen Standards mehrfach aktualisiert und neuen Erkenntnissen angepasst worden.

Vor der Einführung von Behandlungsstandards trat die adjuvante Chemotherapie kostenspezifisch kaum in Erscheinung, da sie nicht flächendeckend eingesetzt wurde. Im Vergleich bewegen sich heute die Behandlungskosten pro Quartal und Patient je nach Behandlungsort (onkologische Praxis, teilstationär, stationär) und Applikationsschema (oral, Infusion) zwischen 1.600 € und 10.000 € [32].

Eine ökonomisch effiziente medizinische Versorgungsleistung und der mit dem Therapieeinsatz verbundene Ressourceneinsatz - abgebildet in den direkten Behandlungskosten – setzen eine nachweisbare pharmakologische Wirksamkeit sowie einen messbaren Nutzen für die betroffenen Patienten voraus.

In diesem Zusammenhang wurde unlängst durch eine Langzeitstudie am Münchner Krebsregister die Diskussion um die Nachhaltigkeit von Therapiestandards bei Krebserkrankungen neu angeregt [54]. Die systematische Auswertung der Daten zu den entsprechenden Krankheitsakten von Krebspatienten seit 1980 wurde in Medienberichten dahingehend gedeutet, dass sich die Erfolgsaussichten bei metastasierten Bronchial-, Kolon-, Mamma- oder Prostatakarzinomen in den vergangenen 20 bis 25 Jahren kaum verändert hat [17]. Nicht zuletzt aufgrund der Schlussfolgerungen aus dieser Langzeitstudie erscheint die Entwicklung systematischer Behandlungsstandards über die vergangenen fünf Jahrzehnte nicht vollständig konsistent im Hinblick auf den therapeutischen Nutzen.

Es bleibt eine Wissenslücke bestehen, welche sich in einem erkenntnistheoretischen Konflikt zwischen klinisch-medizinischen, klinisch-statistischen und klinisch-ökonomischen Expertisen manifestiert. Klinische Standards werden fast ausschließlich von medizinischen Fachkreisen gesetzt. Rationale und effiziente Entscheidungen im Bereich der Gesundheitsversorgung erfordern auch die Berücksichtigung ökonomischer Nebenbedingungen bei der Allokation und Finanzierung von Leistungen sowie die Nachhaltigkeit von Forschungsergebnissen klinischer Studien durch die Optimierung statistisch-mathematischer Methoden.

Mit den verfügbaren Informationen aus wissenschaftlichen Publikationen zu diesem Themenbereich soll erörtert werden, inwieweit sich das Paradigma der adjuvanten Chemotherapie belegen lässt. Die heute im Rahmen von Therapieempfehlungen formulierten Konzepte stützen sich schwerpunktmäßig auf die Ergebnisse klinischer Studien Ende der 80er und Beginn der 90er Jahre, während Untersuchungen aus jüngerer Vergangenheit dieses Bild einer substanziellen therapeutischen Wirksamkeit kontrastieren.

Vor diesem Hintergrund wird in der vorliegenden Arbeit zunächst auf der Basis einer MedLine©-Recherche die pharmakologische Spezifität eines definierten Chemotherapieprotokolls am Beispiel eines häufigen und bösartigen Karzinoms erörtert. Ausgangspunkt der Überlegungen ist, dass eine spezifisch-wirksame Behandlungsstrategie sich insbesondere im Vergleich zu alternativen, möglicherweise neueren Therapieansätzen als vorteilhaft zeigen sollte.

Die Analyse der Spezifität einer Behandlung bietet sich insbesondere im Zusammenhang mit der adjuvanten Zielausrichtung an, denn hierbei steht eine dauerhafte Heilung des Patienten im Vordergrund. Als Indikationsgebiet wird die Diagnose einer Krebserkrankung des Dickdarms und/oder des Mastdarms (kolorektales Karzinom) ausgewählt - eine Erkrankung, die weltweit jährlich bei einer Inzidenz von 1.000.000 etwa 500.000 Todesfälle verursacht und neben anderen malignen Krebserkrankungen wie der Prostata, Blase, Brust, Lunge oder Magen zu den häufigsten zählt [6]. In Deutschland liegt die Neuerkrankungsrate etwa bei 60.000 Personen jährlich [13,19].

Die Intention der zugrunde liegenden Arbeit ist nicht, den therapeutischen Nutzen der adjuvanten Chemotherapie grundsätzlich in Frage zu stellen. Es wird vielmehr erörtert, ob eine in klinischen Studien beobachtete Wirkung möglicherweise nicht spezifisch ist und daher auch nicht ausschließlich auf pharmakologische Faktoren zurückzuführen ist.

In Kapitel 2 wird kurz auf Ätiologie, Diagnostik und Symptomatik des Krankheitsbildes bei kolorektalem Karzinom eingegangen. Zunächst wird das Konzept der adjuvanten Chemotherapie begrifflich-inhaltlich abgegrenzt, die pharmakologische Wirkungsweise und der allgemeine Therapieablauf werden skizziert. Schließlich wird die empirische Evidenz zur Wirksamkeit adjuvanter Chemotherapie aufgearbeitet und daraus resultierende Therapieempfehlungen sowie Richtlinien werden dargestellt.

Kapitel 3 beinhaltet die Vorgehensweise im Rahmen der MedLine© Recherche. Die zugrunde liegende Suchstrategie wird erläutert im Hinblick auf die Definition der Suchbegriffe, die betrachtete Zeitperiode sowie die Selektionskriterien. Eine Übersicht zum qualitativen Status der betrachteten Studien wird entworfen.

Kapitel 4 beschreibt die selektive Auswahl randomisierter klinischer Studien zu diesem Thema inhaltlich nach Ergebnissen. Für ein definiertes Standardchemotherapieprotokoll werden zunächst die Effekte bezüglich eines Vergleichs mit einer Nulltherapie illustriert. Die Frage nach der Spezifität der medizinischen Intervention wird über einen Abgleich der ersten Ergebnisse mit denen aus Studien zu alternativen aktiven Therapieansätzen beleuchtet.

Eine Diskussion der gefundenen Gesamtergebnisse und deren Implikationen findet sich in Kapitel 5. Designprinzipien wie Art der Verblindung, Randomisierung und Kontrollgruppenwahl werden problematisiert. Ein beobachteter ganzheitlicher Effekt einer medizinischen Maßnahme wird im Zusammenhang mit Wirkungsebenen auf physikalisch-chemischer und psychosozialer Art diskutiert. Dabei wird auch die kommunikative Beziehung zwischen Arzt und Patient als Einflussfaktor erörtert. Grundsätzliche Probleme der Messtechnik, die Rolle des Kausalitätsprinzips und Vor- und Nachteile einzelner zur Disposition stehender statistischer Methoden werden behandelt.

Kapitel 6 schließlich fasst Folgerungen und Empfehlungen zusammen. Den Schwerpunkt bilden Aspekte des Studiendesigns wie Kontrollgruppen- und Vergleichstherapiewahl sowie Prinzipien zur Behandlungsgleichheit, Randomisierung und Verblindung. Mögliche Konsequenzen für das Berichtswesen und die Anwendung von Leitlinien werden dargestellt.

2. Krankheitsbild des kolorektalen Karzinoms: klinische Fakten

2.1 Symptomatik, Ätiologie und Diagnostik

Das Risiko, an Darmkrebs zu erkranken, nimmt mit steigendem Alter zu. Männer erkranken im Durchschnitt mit 64, Frauen mit 67 Jahren. Nur 10% der kolorektalen Karzinome entstehen bei Patienten unter 50 Jahren [3]. Erstgradige Verwandte von Patienten mit Darmkrebs haben ein erhöhtes Risiko, ebenfalls am kolorektalen Karzinom zu erkranken. Ein überdurchschnittlich hohes Darmkrebsrisiko weisen Patienten mit familiärer, adenomatöser Polyposis oder mit hereditärem, nicht-polypösem kolorektalem Krebsyndrom auf. Auch chronisch-entzündliche Darmerkrankungen erhöhen möglicherweise das Risiko für Darmkrebs. Man nimmt außerdem an, dass eine Ernährung, die reich an Früchten und Gemüse ist, aber nur wenig tierische Fette enthält, das Risiko für Darmkrebs vermindern könnte.

Durch regelmäßige Vorsorgeuntersuchungen, Untersuchung des Stuhls auf nicht-sichtbares Blut (Hämokulttest) sowie Darmspiegelungen (Koloskopie) soll ein vorhandenes Karzinom frühzeitig entdeckt und somit das Risiko, an Darmkrebs zu versterben, gesenkt werden. Bei der Darmspiegelung können verdächtige Areale biopsiert und feingeweblich untersucht werden. Um die Ausdehnung der Erkrankung festzustellen, sind weitere radiologische Untersuchungen wie Oberbauchsonographie, Röntgenthorax, Computertomographien oder Kernspinuntersuchungen einsetzbar.

Zu den Symptomen, die zur Entdeckung dieser Erkrankung bei den Patienten führen, gehören besonders Änderungen der Stuhlgangsgewohnheiten sowie (sichtbares) Blut im Stuhl. Bei manchen Patienten kommt es durch den Tumor zum Darmverschluss. Begleitsymptome können unter anderem abdominelle Schmerzen, Gewichtsverlust oder Blutarmut sein [18].

2.2 Abgrenzung der adjuvanten Chemotherapie

Die Therapie des Darmkrebses erfolgt abhängig von der Ausdehnung der Erkrankung durch Operationen mit anschließenden Bestrahlungen und/oder Chemotherapie. Eine Heilung vom kolorektalen Karzinom wird als möglich angesehen, wenn sämtliche Tumoranteile durch operativ entfernt werden können. Daher besteht im Fall eines Karzinoms ohne Fernmetastasen ein kurativer Therapieansatz in einer blockweisen Resektion des Primärtumors (R0-Resektion) mit angrenzendem Lymphabflussgewebe. Für Patienten mit metastasiertem Karzinom muss im Einzelfall entschieden werden, ob aufgrund des Allgemeinzustandes eine Primärtumor und/oder Metastasenresektion durchgeführt werden kann.

Bei der Chemotherapie des kolorektalen Karzinoms unterscheidet man grundsätzlich die adjuvante, neoadjuvante und die palliative Chemotherapie.

Bei Patienten mit Metastasenbildung, die wegen ihrer Größe operativ nicht zu entfernen ist, kann eine neoadjuvante Chemotherapie angesetzt werden, um eine Verkleinerung der Metastasen zu erreichen, die sich anschließend operativ entfernen lassen.

Palliative Chemotherapie kommt bei Patienten in Frage, bei denen der Darmtumor und seine Metastasen durch eine Operation nicht vollständig entfernt werden können. In der Regel ist keine Heilung von der Tumorerkrankung mehr möglich.

Eine adjuvante Chemotherapie kann bei Patienten durchgeführt werden, bei denen der Tumor durch die Operation möglicherweise vollständig entfernt worden ist, die jedoch ein deutliches Risiko für einen Rückfall haben. Durch die adjuvante Chemotherapie sollen im Körper verbliebene Tumorzellen, die während der Operation und durch die bildgebenden Verfahren nicht sichtbar waren, abgetötet werden.

Schwerpunkt der folgenden Ausführungen ist die adjuvante Therapie. Speziell mit dieser werden eine kurative Zielsetzung und damit spezifische Therapiewirkung assoziiert.

Zur Charakterisierung des Schweregrades der Tumorerkrankung existieren verschiedene Klassifikationsschemata (Tabelle 1). Eine Indikation zur adjuvanten Chemotherapie wird in Abhängigkeit des klinischen Stadiums der Tumorerkrankung angesetzt.

Tabelle 1: Indikation zur adjuvanten Chemotherapie nach Tumorstadium [18]

UICC – Stadium	TNM – Klassifikation	Schema nach Dukes	Indikation
I	T1 N0 M0 / T2 N0 M0	A / B1	Nein
II	T3 N0 M0 / T4 N0 M0	B2	Ja / Nein
III	T1-T4 N1 M0 / T1-T4 N2 M0	C	Ja
IV	T1-T4 N1-N2 M1	D	entfällt

Die TNM - Klassifikation (Abbildung 1) kombiniert Charakteristika der Tumorerkrankung bezüglich Wachstum und Infiltrationsgrad des Primärtumors (T), den Befall mit regionären Lymphknoten und das Auftreten von Fernmetastasen (M).

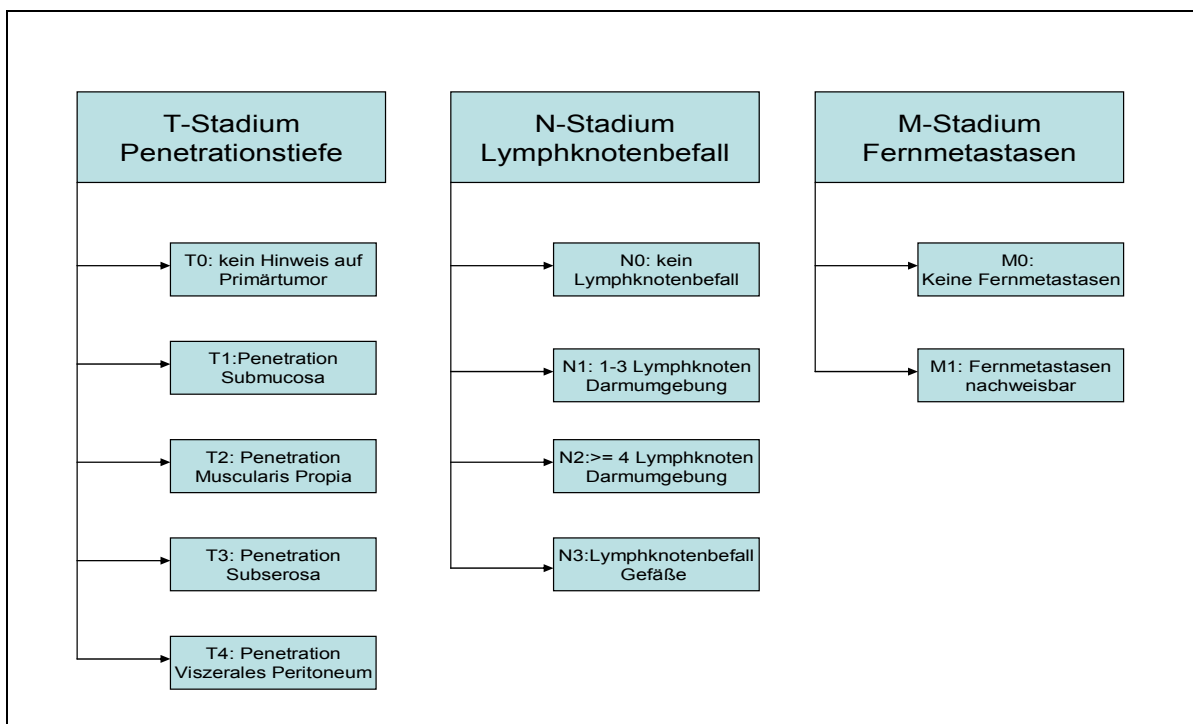


Abbildung 1: Aufschlüsselung TNM Klassifikation, eigene Darstellung nach [18]

Es besteht laut Leitlinien Konsens dahingehend, dass adjuvante Chemotherapie bei Patienten mit Kolonkarzinom im Stadium III klinisch sinnvoll und effektiv ist, bei Patienten im Stadium II wird eine Anwendung außerhalb von klinischen Studien nicht grundsätzlich empfohlen [1]. Zu den Kolonkarzinomen zählen Tumore, deren aboraler Rand mehr als 16 cm von der Anokutanlinie entfernt ist.

Bei Patienten mit Rektumkarzinom gilt für das Stadium III im allgemeinen eine adjuvante Radiochemotherapie. Die Behandlung von Stadium II Patienten wird ebenfalls kontrovers beurteilt [2]. Als Rektumkarzinom zählt ein Tumor, dessen aboraler Rand weniger als 16 cm von der Anokutanlinie entfernt ist.

Etwa 57,5% aller Darmtumoren treten im Kolon auf, 42,5% im Rektum, wobei der Schwerpunkt der Lokalisation im Rektum und Colon sigmoidum ist (67%) [19].

2.3 Wirkungsweise der Chemotherapie

In der Krebstherapie finden Medikamente Anwendung, welche in die Stoffwechselvorgänge der Krebszelle eingreifen. Ziel ist eine nachhaltige Störung der Proteinbiosynthese, was den Tod der Krebszelle herbeiführen soll. Hierbei geht man im Rahmen der medikamentösen Behandlung davon aus, dass sich Tumorzellen schneller als normales Gewebe teilen und deshalb empfindlicher auf therapieinduzierte Wachstumsstörungen reagieren.

Eine Substanz, welche seit den 50er Jahren die Grundlage vieler Chemotherapieprotokolle bildet und bei den folgenden Ausführungen einen zentralen Stellenwert einnimmt, ist Fluorouracil (5-FU). Sie wird der Gruppe der Antimetaboliten zugerechnet, welche das Lesen der Erbinformation aus der DNS im Zellstoffwechsel der Tumorzellen durch Einbinden eines „falschen“ Enzyms stört. Bei der Zellteilung wird die DNS zum Ablesen entspiralisiert. Dabei werden Basen, die den eigentlichen genetischen Code enthalten, freigegeben. An diesen lagern sich wieder passende Basen an.

Es bildet sich ein Überträgermolekül, welches nach Abschluss des Leseprozesses aus dem Zellkern ins Zellplasma wandert und die Bildung von körpertonotwendigen Eiweißverbindungen steuert. Die natürlichen Basen leiten sich chemisch von den Substanzen Purin und Pyrimidin ab. Fluorouracil ist mit den DNS-Basen chemisch verwandt und bewirkt – als „falsche“ Base in die DNS eingebracht – über den Prozess des Eiweißaufbaus den Tod der Krebszelle [9].

Neben dieser Gruppe existieren weitere Subgruppen von Zytostatika, die sich nach dem jeweiligen Angriffspunkt unterscheiden lassen, wobei jedoch der Wirkungsmechanismus prinzipiell ähnlich ist.

Durch eine Kombination verschiedener Zytostatika mit unterschiedlichen Angriffspunkten (Polychemotherapie, multimodale Chemotherapie) kann versucht werden, eine Steigerung der Wirkung beziehungsweise eine Reduzierung der Nebenwirkungen zu erreichen.

2.4 Therapieablauf und Nebenwirkungen

Die Chemotherapie ist als Intervallbehandlung in Form von Zyklen konzipiert. Behandlungsphasen wechseln sich mit Behandlungspausen ab. Während der Behandlungsphase werden die Chemotherapeutika meist intravenös verabreicht. Während der Pausen kann der Körper eventuell angegriffenes gesundes Gewebe regenerieren. In der Regel werden mehrere Zyklen durchgeführt. Zytostatika sind dann wirksam, wenn sich die betreffende Zelle in einer Teilungsphase befindet. Durch die Anwendung in mehreren Zyklen sollen auch solche Zellen erfasst werden, die sich während eines vorangegangenen Zyklus gerade in der Ruhephase befanden und so nicht durch den vorangegangenen Zyklus erfasst wurden. Ziel der Behandlung ist eine möglichst vollständige Remission mit anschließender Stabilisierung des erreichten Niveaus.

Die Zytostatika schädigen neben den Tumorzellen stets auch die gesunden Zellen. Sie besitzen in diesem Sinne eine relative Spezifität, indem Tumorzellen stärker geschädigt werden als das gesunde Gewebe. Abgesehen von einer möglichen akuten allergischen Reaktion erklären sich die Nebenwirkungen einer Chemotherapie weitgehend durch ihre

zellschädigende Wirkung auf Gewebe mit raschem Zellumsatz: Blutbildendes Knochenmark, lymphatisches Gewebe, Schleimhaut des Verdauungstraktes, Keimdrüsen, Haut und Hautanhangsgebilde.

Die im Laufe einer Chemotherapie möglichen Nebenwirkungen sind unter anderem abhängig von der Art der eingesetzten Zytostatika, ihrer Dosis sowie der Verabreichungsdauer. Eine Rolle spielt auch die allgemeine psychische und körperliche Verfassung der Betroffenen.

Zu den wichtigsten Nebenwirkungen zählen in diesem Zusammenhang [9]:

- Übelkeit und Erbrechen
- Haarausfall
- Schleimhautveränderungen und -entzündungen
- Störung der Blutgerinnung
- Erhöhtes Infektionsrisiko
- Müdigkeit
- Entzündungen im Bereich des Verdauungstraktes (Stomatitis, Enterokolitis)
- Schädigung verschiedener Organe wie Leber, Lunge, Nieren, Herz
- Schädigung der Fortpflanzungsfähigkeit etwa durch Erbgutveränderungen

2.5 Empirische Evidenz zur Chemotherapie bei KRK , historischer Überblick

Vor fünf bis sechs Jahrzehnten wurden erste Behandlungsversuche mit zytotoxischen Substanzen unternommen. In der Folgezeit untersuchte man zunehmend deren Wirkungsweise im Zusammenhang mit verschiedenen Arten von Krebserkrankungen. Bereits in den 60er Jahren gab es Therapieansätze mit der noch heute verwendeten Basissubstanz 5-Fluorouracil. Insbesondere im Verlauf der 80er Jahre intensivierte sich die klinische Forschung zur Wirksamkeit der Chemotherapie bei kolorektalem Karzinom mit einer zunehmenden Anzahl klinischer Studien und den entsprechenden Publikationen.

Zunächst jedoch blieb der therapeutische Stellenwert der Chemotherapie umstritten, waren die beobachteten Effekte doch eher marginal [33]. 1989 wurden die Studienergebnisse zu einer Therapie basierend auf 5-Fluorouracil (5-FU) und Levamisole veröffentlicht [36,41]. Eine Kombinationstherapie aus 5-FU und Levamisole zeigte sich als hoch wirksam hinsichtlich rezidivfreier Überlebenszeit und allgemeiner Überlebenszeit bei Stadium III Patienten. Bei Stadium II Patienten wurden Aussagen zur Wirksamkeit weiterhin kontrovers diskutiert [23].

In den folgenden Jahren untersuchte man - additives bis synergistisches Zusammenwirken einzelner Substanzen vermutend - verschiedene 5-FU basierende Modulationen, die in einen weiteren Therapieansatz mündeten [34]: 5-FU in Kombination mit Leucovorin (Folinsäure) - 5-FU / LV. Konsequenterweise wurde darauf abgezielt, die Wirksamkeit durch Erweiterung mit zusätzlichen Komponenten – wie beispielsweise Interferon, Mitomycin oder Oxaliplatin – zu erhöhen [56,22,28]. Diese Therapiemodulationen wurden im Allgemeinen jedoch nicht als Standard etabliert. Eine Ausnahme bildet die Therapievariante mit Oxaliplatin. Im Rahmen einer klinischen Studie aus dem Jahr 2004 gelang es, einen entsprechend ausgeprägten Wirksamkeitseffekt zu beobachten [22]. In überarbeiteten aktuellen Leitlinien wird deshalb dieses Schema als Therapiemöglichkeit angezeigt [1,2,3].

Abweichend von den 5-FU basierenden Behandlungen wurden auch immuntherapeutische Behandlungsstrategien untersucht [55]. Immuntherapeutische Behandlungen können bei Patienten mit empfindlicher Antikörperreaktion problematisch sein, daher hat sich diesbezüglich noch kein Standard durchgesetzt. Auch im Bereich der Gentherapie wird nach möglichen Alternativen zur herkömmlichen Chemotherapie geforscht.

5-FU wird im Allgemeinen intravenös verabreicht. Dagegen bietet eine orale Verabreichung Vorteile im Hinblick auf Verträglichkeit und Handhabung [48,40]. Gebräuchlich sind hierbei Medikationen wie Tegafur oder Uracil.

Der Antimetabolit 5-FU ist seit Jahrzehnten die Basiskomponente einer adjuvanten Chemotherapie.

Gegenwärtig stellt das intravenös verabreichte mit Folinsäure modulierte 5-FU Chemotherapieprotokoll immer noch den Eckpfeiler in der adjuvanten Behandlung bei Patienten mit kolorektalem Karzinom nach erfolgreichem operativen Eingriff dar [57]. Fluorouracil und Leucovorin werden in verschiedenen Dosierungen und Zyklen verwendet. Eine Behandlungsdauer über sechs Monate wird als Standard in der adjuvanten Therapie bei lymphknotenpositiven Stadium III Patienten angesetzt [38]. Weder eine Erhöhung der Behandlungsdauer (auf 12 Monate) noch eine Erhöhung der Folinsäuredosis scheint einen zusätzlichen therapeutischen Nutzen zu erbringen [27,42,34,46].

3. Material und Methoden

3.1 Medline©-Literaturrecherche

Ausgangspunkt ist eine MedLine© – Literaturrecherche anhand der Suchbegriffe “colorectal carcinoma”, “adjuvant chemotherapy” and “randomized clinical trial” im Zeitraum zwischen 1960 und 2006. Es wurden ausschließlich Studien an Menschen („Humans“) betrachtet. Einschränkungen bezüglich des Publikationstyps wurden zunächst nicht getroffen. Mit Verwendung der entsprechenden Eingabemasken für den Suchalgorithmus ergaben sich insgesamt 446 Treffer unter den definierten Suchbegriffen. Das Ergebnis ist in Tabelle 2 abgebildet.

Tabelle 2: Ergebnis MedLine© – Suche [PubMed]

Limits: Entrez Date from 1960/01/01 to 2006/12/31, Publication Date from 1960/01/01 to 2006/12/31, Humans

Result:

[446](#)

Translations:

adjuvant chemotherapy	"chemotherapy, adjuvant"[MeSH Terms] OR adjuvant chemotherapy[Text Word]
colorectal carcinoma	("colorectal neoplasms"[TIAB] NOT Medline[SB]) OR "colorectal neoplasms"[MeSH Terms] OR colorectal carcinoma[Text Word]
randomized clinical trial	"randomized controlled trial"[Publication Type] OR "randomized controlled trials"[MeSH Terms] OR "randomized clinical trial"[Text Word] OR "randomised clinical trial"[Text Word]
Humans[Mesh]	"humans"[MeSH Terms]

Database:

PubMed

User query:

randomized clinical trial and colorectal carcinoma and adjuvant chemotherapy AND (Humans[Mesh]) AND ("1960/01/01"[PDat]:"2006/12/31"[PDat]) AND ("1960/01/01"[EDat]:"2006/12/31"[EDat])

3.2 Selektion

Die spezifische Auswahl der Literaturdatenbankeinträge erfolgte durch manuelle Charakterisierung der Studien auf Basis der Abstracts beziehungsweise Volltexte. Aus der Gesamtheit von 446 identifizierten Einträgen wurden in einem ersten Schritt 196 Referenzen ausgeschlossen. Die Kriterien für einen Ausschluss beinhalteten zunächst semantische Aspekte: Eine Veröffentlichung sollte in englischer Sprache erfolgt sein. 105 in der jeweiligen Landessprache veröffentlichte Studien erfüllten diese Anforderung nicht. Hinsichtlich formaler Aspekte wurde gefordert, dass es sich um eine originäre Studie als wissenschaftliche Publikation handelte und nicht als „Letter“ oder „Comment“ typisiert war. Hiervon waren 22 Publikationen betroffen. Schließlich wurden inhaltliche Aspekte angesetzt: es sollte sich um eine Therapiestudie mit Wirksamkeit als primäres Studienziel und entsprechend definierten Hauptzielgrößen handeln. Aufgrund dieser Anforderung wurden 69 weitere Einträge ausgeschlossen. Es verblieben somit 250 Einträge, 156 Einzelstudien und 94 Übersichtsarbeiten als thematische Grundlage für die Diskussion.

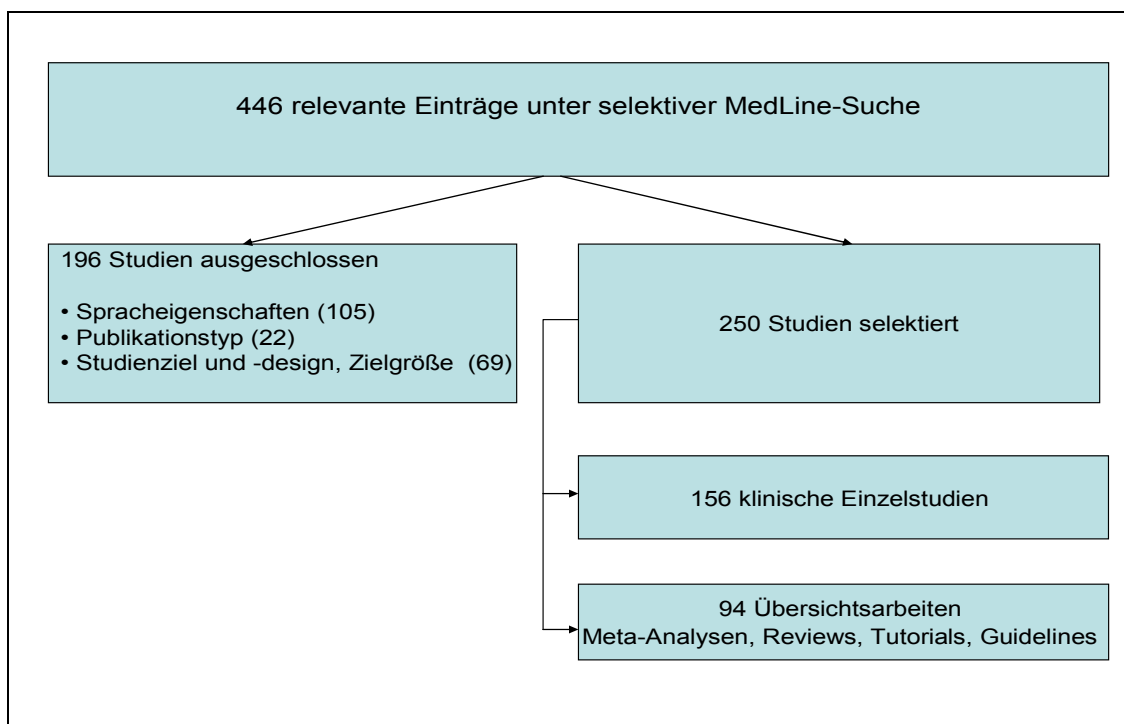


Abbildung 2: Flussdiagramm der Medline© - Recherche

3.3 Operationalisierung der Arbeitshypothese

Die Konkretisierung der Arbeitshypothese erfolgte anhand eines in den AMWF - Leitlinien und vergleichbaren Richtlinien definierten Standardtherapieansatzes zur Behandlung des Kolonkarzinoms. In Deutschland werden derzeit drei spezifische Schemata als Leitlinie in der Chemotherapie-Anwendung empfohlen [1,2,3]. Alle beinhalten eine pharmakologische Variante der Kombination aus 5-FU und Folinsäure beziehungsweise Leucovorin:

Variante 1

20 mg/m² Folinsäure (rasche intravenöse Infusion) plus 425 mg/m² 5-FU (Bolusinjektion innerhalb von weniger als 5 min) in Woche 1, 4 und 8 jeweils von Tag 1-5, 3 weitere Zyklen alle 5 Wochen, insgesamt 6 Zyklen

Variante 2

einmal pro Woche über insgesamt 6 Wochen 500 mg/m² Folinsäure (2 Stunden Infusion), anschließend 500 mg/m² 5-FU (1 Stunden Infusion), alle 8 Wochen insgesamt 4 Zyklen

Variante 3

Folinsäure (200 mg/m² als 2 Stunden Infusion, Tag 1 und 2) plus 5-FU (400mg/m² als Bolus, danach 600 mg/m² als 22 Stunden Infusion; Tag 1 und 2) in Kombination mit Oxaliplatin (85 mg/m² als 2 Stunden Infusion; Tag 1)
1 Zyklus umfasst 2 Wochen, insgesamt 12 Zyklen

Die Fragestellung nach der Spezifität der adjuvanten Chemotherapie wurde unter Bezugnahme auf das in den Empfehlungen erstgenannte Schema untersucht. Da dieser Standard bereits 1990 in der NIH Konsensuskonferenz [10] formuliert worden ist, wird aufgrund der Erfahrungen über einen Anwendungszeitraum von mehr als 15 Jahren ein entsprechendes Informationspotenzial in der Literaturdatenbank vermutet. Die Limitierung der zur Verfügung stehenden 250 Einträge aus Abschnitt 3.2 auf das definierte Chemotherapieprotokoll ergab eine selektive Auswahl von 10 Studien, welche den weiteren Ausführungen zugrunde gelegt wurden [25,26,29,31,34,39,44,47,49,50].

3.4 Qualitative Beurteilung der Publikationen

Insgesamt zehn Publikationen, vier Meta-Analysen basierend auf neun randomisierten individuellen Studien sowie sechs weitere Einzelstudien wurden einer detaillierten Betrachtung unterzogen. Zu diesem Zweck wurden die Studien in einem Bewertungsschema eingeordnet [35]. Das Schema enthält Aspekte im Hinblick auf das Studien-Design, insbesondere Randomisierung, Verblindung, Behandlungs- und Beobachtungsgleichheit sowie Kriterien zur statistischen Methodik wie Power-Berechnung und Anwendung multivariater Verfahren. Darüber hinaus wurden auch Eigenschaften der Studie wie „Intention-to-Treat“ Analyse und Handhabung des „Follow-up“ in die Beurteilung mit einbezogen. Eine Übersicht findet sich in Tabelle 3 auf der folgenden Seite.

Im Gesamtergebnis gleichen sich die Studien hinsichtlich qualitativer Aspekte. Die Anzahl und Art erfüllter Kriterien ist fast immer identisch. Eine qualitative Beurteilung von Studien anhand solcher Schemata hat zunächst jedoch nur formalen Charakter, denn die Einordnung bezüglich der verschiedenen Aspekte basiert in erster Linie auf den eigenen Angaben der Autoren.

Um Ergebnisse gezielt nachvollziehen zu können, müssten zumindest stichprobenartig von den Autoren Studiendaten zwecks Replikation der Datenselektion, -transformation und Rechenalgorithmen beispielsweise bei einem verantwortlichen Statistiker im entsprechenden Verlag eingereicht werden. Zu spezifischen Aspekten des Designs wie Art der Randomisierung, Zeitpunkt und Vollständigkeit der Patienteneinwilligung, Verblindungstechnik und Gewährleistung der Behandlungs- und Beobachtungsgleichheit wäre im Einzelfall eine detailliertere Darstellung zweckmäßig, da genau jene Designkriterien das in Untersuchungen beobachtete Wirksamkeitsergebnis beeinflussen können. Diese Problematik wird im nachfolgenden Diskussionsteil aufgegriffen.

Tabelle 3: Qualitatives Bewertungsschema [35]

Kriterium	Studiendesign		
	Erfüllt	Nicht erfüllt	Unklar
Einwilligungserklärung	alle		
Randomisierung	alle		
Keine genauen Angaben	26, 31, 34, 39, 47		
Einfach	44		
Stratifiziert	25, 29, 49, 50		
Verblindete Gruppenzuteilung			alle
Verblindung (Patienten)			alle
Verblindung (Prüfärzte)			alle
Verblindung (Analysten)			alle
Intention-to-Treat Analyse	25, 26, 29, 34, 39, 47, 49, 50		29, 44
Follow-Up Zeit ausreichend	26, 31, 34, 39, 44, 47, 50	29	25, 49
Beschreibung loss to follow-up	alle		
Vergleichbare Baseline Eigenschaften	alle		
Behandlungsgleichheit	alle		
	Statistische Methodik		
Power Berechnung	25, 26, 31, 34, 39, 44, 47, 49, 50	29	
Anwendung multivariater Verfahren	25, 26, 31, 34, 39, 47, 49, 50	29, 44	
	Validität		
Konsistenz	25, 26, 29, 31, 34,		44
Vermeidung systematischer Fehler	39, 47, 49, 50		

4. Ergebnisse

4.1 Vergleich eines Chemotherapieprotokolls mit einer Nulltherapie

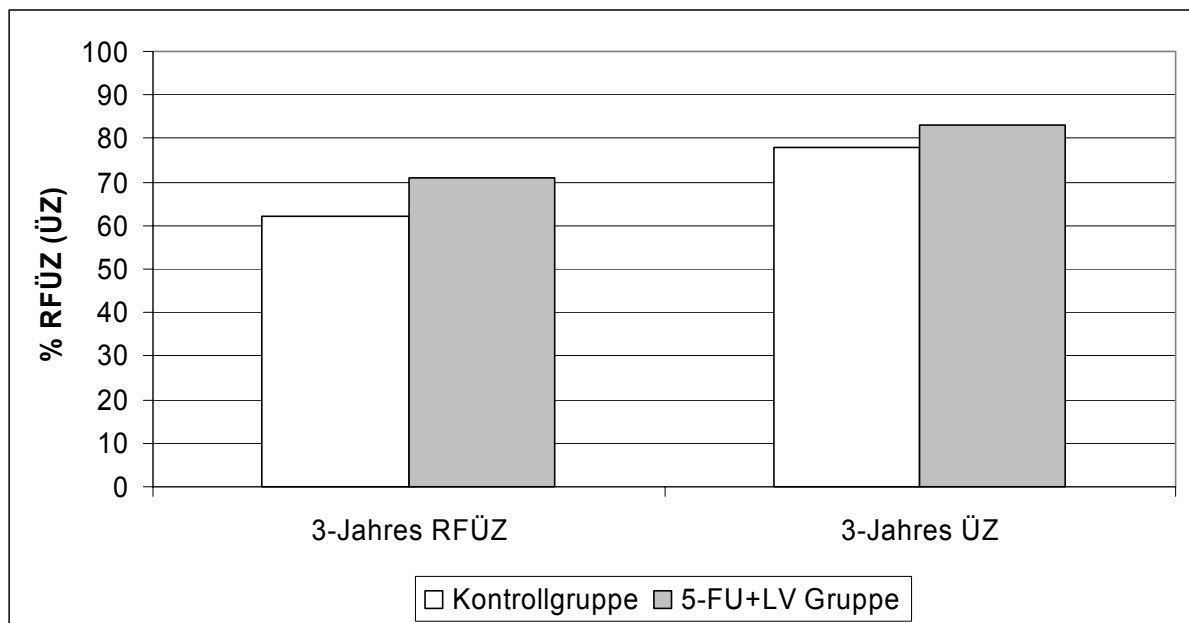
Zunächst wird eingegangen auf die Ergebnisse einer Meta-Studie (IMPACT) aus dem Jahre 1995 [34]. Unter Verwendung von Daten aus drei verschiedenen Quellen - GIVIO, NCIC-CTG, FFCD - betrachteten die Autoren vergleichbare Behandlungsschemata (FU 370-400 mg/m² i.v. + 200 mg/m² Folinsäure x 5 Tage alle 28 Tage über 6 Monate). In vorangegangenen Einzelstudien waren diese jeweils als Experimentalgruppe einer Kontrollgruppe mit Nulltherapie gegenübergestellt. In dieser erhielten die Patienten ausschließlich eine kurative Operation und keine weiteren adjuvanten Folgetherapien. Insgesamt konnten im Rahmen dieser Meta-Analyse 1493 Patienten evaluiert werden. Im Ergebnis zeigt sich ein statistisch signifikanter Unterschied zwischen der Experimentalgruppe und der Kontrollgruppe: Durch den Einsatz adjuvanter Chemotherapie stieg der Anteil der nach drei Jahren noch lebenden Patienten um 5%-Punkte, der Anteil der nach 3 Jahren noch rezidivfreien Patienten um sogar 9%-Punkte. Der Effekt bezüglich der primären Hauptzielgröße „Gesamtüberlebenszeit“ (Overall Survival), im folgenden abgekürzt als ÜZ, ist somit deutlich weniger ausgeprägt als der Effekt der sekundären Hauptzielgröße „rezidivfreie Überlebenszeit“ (Disease Free Survival), im folgenden abgekürzt als RFÜZ.

Tabelle 4: Behandlungseffekte 3-Jahres RFÜZ & ÜZ [34]

	5 FU / LV	Nulltherapie	P-Wert	Hazard Ratio	NNT
3-Jahres RFÜZ	71%	62%	0,03	0,67 (0,56 – 0,80)	11
3-Jahres ÜZ	83%	78%	0,02	0,77 (0,62 – 0,96)	20

Studienperiode: 1982-1993, Fallzahl: 1493, Median Follow-Up: 3,3 Jahre

Hazard Ratio definiert als 5-FU / LV vs Beobachtung, 95%-Konfidenzintervall in Klammern, NNT = Number Needed To Treat, gerundet, eigene Berechnungen



*Abbildung 3: Behandlungseffekte, eigene Darstellung
3-Jahres rezidivfreie Überlebenszeit / Gesamtüberlebenszeit*

2001 wurden in einer Folgestudie die langfristigen (10 Jahres-) Effekte analysiert, wobei sich der therapeutische Vorteil einer 5-FU/LV Behandlung im Vergleich zur Nulltherapie-Kontrollgruppe bestätigte [39].

Eine neuere übergreifende Analyse aus dem Jahr 2004 basierend auf den Individualdaten randomisierter klinischer Einzelstudien ergab ebenfalls einen hoch signifikanten Einfluss auf die Überlebenszeit bezüglich 6 verschiedener Protokollvariationen auf der Basis von 5-FU mit LV beziehungsweise einer Therapie-Modulation mit Levamisole [21,50]. In die Betrachtung wurden verschieden-dosierte Kombinationstherapien aus Fluorouracil und Folinsäure (Leucovorin) sowie eine Therapie bestehend aus Fluorouracil und Levamisole einbezogen. Die Schemata sind in der folgenden Tabelle 5 aufgeführt. Allen in der Meta-Analyse verwendeten Einzelstudien lag eine Gegenüberstellung der aktiven Behandlung als Experimentalgruppe mit einer Nulltherapie-Kontrollgruppe zugrunde. Die zu vergleichenden Gruppen waren hinsichtlich der Verteilung üblicherweise erhobener demographischer und klinischer Basisparameter (Alter, Geschlecht, Lymphknotenbefall, T-Klassifikation, Histologie, Lokalisation) ausgewogen.

Tabelle 5: Übersicht Chemotherapieprotokolle Meta-Analyse [31, 50], Seite 1798

Nr	Modulation mit Levamisole	Dauer
1	FU 450 mg/m ² i.v. x 5 Tage alle 28 Tage plus Levamisole 50 mg oral dreimal täglich x 3 Tage alle 14 Tage	12 Monate
Modulation mit Levamisole		
2	FU 425 mg/m ² i.v. plus Leucovorin 20 mg/m ² i.v. x 5 Tage alle 28 Tage	6 Monate
3	FU 400 mg/m ² i.v. plus Leucovorin 200 mg/m ² i.v. x 5 Tage alle 28 Tage	12 Monate
4	FU 400 mg/m ² i.v. plus Leucovorin 200 mg/m ² i.v. x 5 Tage alle 28 Tage	6 Monate
5	FU 370 mg/m ² i.v. plus Leucovorin 200 mg/m ² i.v. x 5 Tage alle 28 Tage	6 Monate
6	FU 340 mg/m ² i.v. plus Leucovorin 200 mg/m ² i.v. x 5 Tage alle 28 Tage	6 Monate

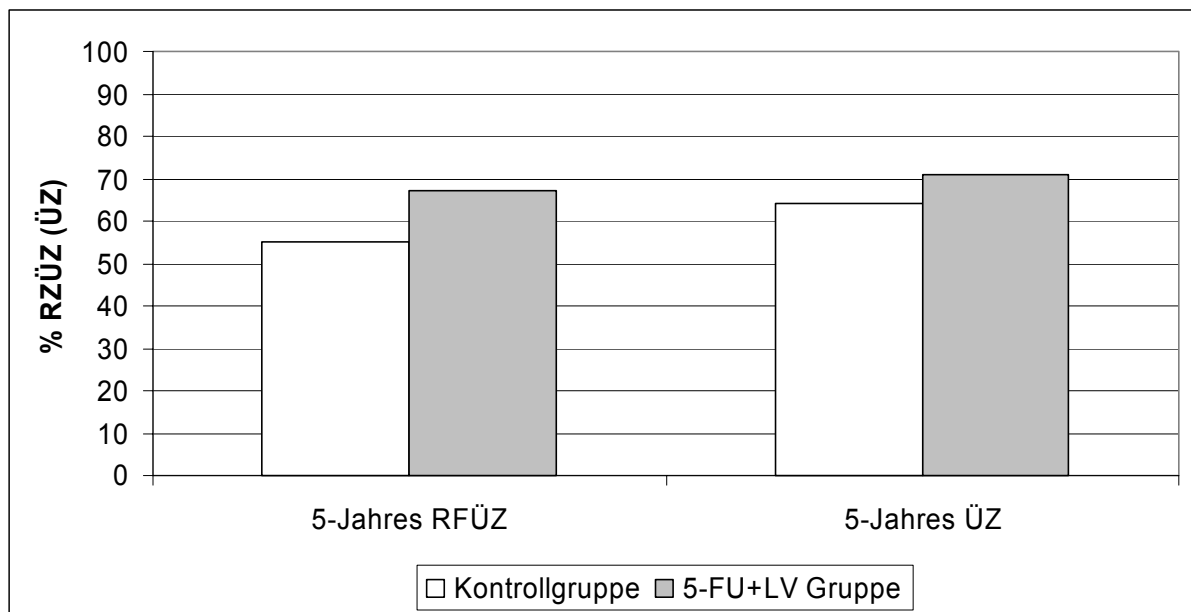
Insgesamt ist in der Gruppe mit Chemotherapie die Überlebensrate um 7%-Punkte signifikant erhöht einhergehend mit einem um 12%-Punkte verringertem Rezidivrisiko. Der Therapievorteil einer 5-FU Behandlung wird durch die entsprechenden Hazard Ratios (0,74 respektive 0,70) unterstrichen.

Tabelle 6: Behandlungseffekte 5-Jahres Zeithorizont [31, 50], Seite 1800

	5 FU / LV	Nulltherapie	P-Wert	Hazard Ratio	NNT
5-Jahres RFÜZ	67%	55%	< ,0001	0,70 (0,63 – 0,78)	8
5-Jahres ÜZ	71%	64%	< ,0001	0,74 (0,66 – 0,83)	14

Studienperiode: 1982 – 1993, Fallzahl: 3302, Median Follow-Up: 5,3 - 8,5 Jahre

Hazard Ratio definiert als 5-FU / LV vs Beobachtung, 95%-Konfidenzintervall in Klammern, NNT = Number Needed To Treat, gerundet, eigene Berechnungen



*Abbildung 4: Behandlungseffekte, eigene Darstellung
5-Jahres rezidivfreie Überlebenszeit / Gesamtüberlebenszeit*

Es ergibt sich folgendes Gesamtbild: Bezüglich des Zielkriteriums „rezidivfreie Überlebenszeit“ unterscheiden sich Experimentalgruppe und Kontrollgruppe signifikant. Der beobachtete Unterschied rezidivfreie Überlebenszeit schwankt zwischen neun und 12 Prozentpunkten. Hinsichtlich der Zielgröße „Gesamtüberlebenszeit“ bewegen sich die gemessenen Effekte zwischen fünf und sieben Prozentpunkten. Die Wirksamkeit in Bezug auf diese Zielgröße scheint somit weniger stark ausgeprägt. Möglicherweise wirkt die Studientherapie schwerpunktmäßig auf Metastasen und weniger stark auf die Gesamtüberlebenszeit. Selbst wenn die beiden Zielgrößen korreliert sind, bleibt die eigentlich relevante Zielgröße die Gesamtüberlebenszeit. Auch ist in diesem Zusammenhang nicht eindeutig, welchen Stellenwert nicht-pharmakologische Faktoren, insbesondere Placebo-Effekte, haben, da der Abgleich der in der Experimentalgruppe beobachteten Ergebnisse ausschließlich über eine Beobachtungsgruppe (Nulltherapie-Kontrollgruppe) erfolgt.

4.2 Vergleich eines Chemotherapieprotokolls mit einer alternativen aktiven Therapie

Unter Berücksichtigung der vorangegangenen Ausführungen würde man bei einem Vergleich gegen eine Kontrollgruppe mit ebenfalls aktiver chemotherapeutischer Behandlung eine ähnliche therapeutische Überlegenheit erwarten, sofern die beobachtete Wirksamkeit spezifisch ist. Um diese Annahme näher zu untersuchen, wird ein Chemotherapieprotokoll in der Ausgestaltung „FU 425 mg/m² i.v. + Leucovorin 20 mg/m² i.v. x 5 Tage alle 28 Tage über 6 Monate“ exemplarisch betrachtet. Bei diesem Anwendungs- und Dosierungsschema handelt es sich um den ältesten Standard, welcher in teilweise leicht abgewandelter Protokollform weit verbreitet ist und auch in aktuellen Leitlinien noch empfohlen wird [1, 2,3].

Genau dieser Therapieansatz war im vorangegangenen Abschnitt anhand von Ergebnissen aus Analysen mit aggregierten Daten („Pooled Data“) untersucht und als signifikant wirksam beurteilt worden [31, 34, 39, 50]. Auch in Einzelstudien mit wesentlich geringerer Fallzahl, beispielsweise in einer klinischen Studie aus dem Jahre 1997, zeigten sich statistisch signifikante Wirksamkeitseffekte für den Vergleich des Therapieansatz mit einer Nulltherapie Kontrollgruppe [43]: Für das aktive Schema ergab sich ein signifikanter Behandlungsvorteil bezüglich der rezidivfreien und allgemeinen Überlebenszeit (p=0.004 beziehungsweise p=0.02).

Tabelle 7: 5-Jahres Behandlungseffekte 5-FU / LV vs Nulltherapie [43]

	5-FU / LV	Nulltherapie	P	Relatives Risiko	NNT
5-Jahres RFÜZ	74%	58%	0,004	1,19-2,60	6
5-Jahres ÜZ	74%	63%	0,02	1,06-2,31	9

Studienperiode: Feb 1988 – Aug 1989, Fallzahl: 309, Median Follow-Up: 6 Jahre

95%-Konfidenzintervall für relatives Risiko definiert als Nulltherapie vs 5-FU / LV

NNT = Number Needed To Treat, gerundet, eigene Berechnungen

In einer Studie aus dem Jahr 2003 wird dieses Chemotherapieprotokoll andererseits als Kontrollgruppe im Vergleich zu einer alternativen aktiven Behandlungsgruppe 5-FU PVI („Protracted Venous Infusion“: 5-Fluorouracil 300 mg/m² lang-andauernde venöse Infusion über 12 Wochen) untersucht [47]. Bei der rezidivfreien Überlebenszeit hat die Experimentalgruppe mit PVI 5-FU einen signifikanten Wirksamkeitsvorteil gegenüber der Kontrollgruppe mit 5-FU / LV. Bezüglich der Gesamtüberlebenszeit waren die beobachteten Unterschiede jedoch nicht signifikant. Die Behandlung in der Experimentalgruppe PVI 5-FU war mit einer geringeren Toxizität und einer höheren gemessenen Lebensqualität verbunden.

Tabelle 8: 3-Jahres Behandlungseffekte 5-FU / LV vs PVI 5-FU [47]

	PVI 5-FU	5-FU / LV	P	Hazard Ratio	NNT
3-Jahres RFÜZ	80,0%	68,6%	0,023	1,44 (1,01 – 2,06)	9
3-Jahres ÜZ	87,9%	83,2%	0,764	2,25 (0,91 – 5,60)	21

Studienperiode: nicht definiert, Fallzahl: 692, Median Follow-Up: 1,7 Jahre

Hazard Ratio definiert als 5-FU / LV vs PVI 5-FU; 95%-Konfidenzintervall in Klammern

NNT = Number Needed to Treat, gerundet, eigene Berechnungen

Eine Folgestudie in 2005 [26] bestätigt diese Ergebnisse im Wesentlichen.

Tabelle 9: 5-Jahres Behandlungseffekte PVI 5-FU vs 5-FU / LV [26]

	5-FU / LV	PVI 5-FU	P	Hazard Ratio	NNT
5-Jahres RFÜZ	66,7%	73,3%	0,100	1,25 (0,99 – 1,62)	15
5-Jahres ÜZ	71,5%	75,7%	0,083	1,26 (0,97 – 1,64)	24

Studienperiode: 1993 – 2003, Fallzahl: 801, Median Follow-Up: 5,3 Jahre

Hazard Ratio definiert als 5-FU / LV vs PVI 5-FU; 95%-Konfidenzintervall in Klammern

NNT = Number Needed to Treat, gerundet, eigene Berechnungen

Die Therapie mit PVI erweist sich tendenziell als therapeutisch überlegen.

Ein ähnliches Bild ergab sich im Rahmen einer weiteren Studie aus dem Jahr 2005. Hier wurde als Experimentalgruppe eine Behandlung mit 250 mg/m² FU Dauerinfusion (CIFU = continuous-infusion fluorouracil) über 56 Tage, alle 9 Wochen über 3 Zyklen verwendet. Die Kontrollgruppe bestand aus einer leicht abgewandelten 5-FU / LV Standardbehandlung. Zusätzlich wurde das Schema in beiden Gruppen um eine Verabreichung von Levamisole modular erweitert (LEV 50 mg/m² alle 8 Stunden x 3 Tage, Wiederholung alle 14 Tage, 6 Zyklen). Die Ergebnisse deuten auf keine wesentlichen Unterschiede zwischen den Behandlungsgruppen hin [44].

Tabelle 10: 5-Jahres Behandlungseffekte CIFU vs 5-FU / LV [44]

	CIFU 5-FU	5-FU / LV	P	Hazard Ratio	NNT
5-Jahres RFÜZ	63%	61%	0,65	0,95 (0,77-1,16)	50
5-Jahres ÜZ	69%	70%	0,18	0,86 (0,69-1,08)	100

Studienperiode: 1994 - 1999, Fallzahl: 940 Patienten, Median-Follow Up: 6,5 Jahre
Hazard Ratio definiert als 5 FU / LV vs CIFU; 95%-Konfidenzintervall in Klammer
NNT = Number Needed to Treat, gerundet, eigene Berechnungen

Auch in einer vorangegangenen Studie aus dem Jahr 2002 [29], in welcher das betrachtete 5-FU / LV Standardschema als Kontrollgruppe gegen eine Kombinationstherapie aus 5-FU + alfa-2b Interferon (5-FU 600 mg/m² 24h Infusion x 5 Tage; 5-FU 600 mg/m² i.v. 1 x Woche IFN 5 MU s.c. 3 x Woche über 6 Monate) verglichen wurde, waren keine statistisch signifikanten therapeutischen Unterschiede beobachtbar.

Tabelle 11: Behandlungseffekte 5-FU / IFN vs 5-FU / LV [29]

	5-FU / IFN	5-FU / LV	P	NNT
RFÜZ	57%	54%	0,435	33
ÜZ	57%	59%	0,245	50

Studienperiode: Jan 1991-Jun 1998, Fallzahl: 322 Patienten, Median-Follow Up: 4,0 Jahre
NNT = Number Needed to Treat, gerundet, eigene Berechnungen

Im Zusammenhang mit einem bezüglich der Ausgangslage leicht abgewandelten klinischen Szenario zur Behandlung des metastasierenden kolorektalen Karzinoms ohne eine vorangegangene makroskopische Entfernung des Primärtumors zeigte sich ebenfalls kein spezifischer therapeutischer Wirksamkeitsvorteil für das 5-FU - Folinsäure Standardschema.

In einem Studiendesign mit drei Behandlungsgruppen wurde eine 3-Komponenten Kombinationstherapie bestehend aus Irinotecan, Fluorouracil und Folinsäure (Irinotecan 125 mg/m² i.v. Infusion über 90 Minuten + Folinsäure 20 mg/m² i.v. Bolus + 5-Fluorouracil 500 mg/m² i.v. Bolus Tag 1,8,25,22 – alle 6 Wochen, 6 Zyklen) mit dem in den vorangegangenen Abschnitten betrachteten Standardprotokoll und einer Monotherapie ausschließlich bestehend aus Irinotecan (Irinotecan 125 mg/m² i.v. Infusion über 90 Minuten Tag 1,8,25,22 – alle 6 Wochen, 6 Zyklen) verglichen. Schwerpunkt der Analyse waren die beiden Kombinationstherapien mit den Zielgrößen „progressionsfreie Überlebenszeit“ und „Gesamtüberlebenszeit“ [49].

Tabelle 12: Behandlungseffekte Irinotecan / LV / 5-FU vs 5-FU / LV [49]

In Monaten	Irinotecan + 5-FU / LV	5-FU / LV	P	NNT
Mediane PFÜZ	7,0	4,3	0,004	37
Mediane ÜZ	12,6	14,8	0,04	45

Studienperiode: Mai 1996-Mai 1999, Fallzahl: 683 Patienten

NNT = Number Needed to Treat, gerundet, eigene Berechnungen

Für die Experimentalgruppe bestehend aus der multimodalen Behandlung Irinotecan + 5-FU / LV zeigte sich ein geringfügiger Wirksamkeitsvorteil bei vergleichbarer Toxizität und Lebensqualität. Der deutlichste Unterschied wurde hinsichtlich der Response-Rate – auf Computer-Tomographie sichtbare Reduktion des betroffenen Areals um mindestens 50% - beobachtet (50% vs 28%, p<0,001).

In einer weiteren Studie, 2002, wurde besagtes 5-FU / LV Chemotherapieprotokoll einer oral applizierbaren 5-FU Behandlungsvariante (300 mg/m² UFT + 20 mg/m² LV oral über 28 Tage alle 35 Tage wiederholt) gegenübergestellt [25]. Neben „Gesamtüberlebenszeit“, „Response-Rate“, „Toxizität“ und „Lebensqualität“ wird als Zielgröße ebenfalls „progressionsfreie Überlebenszeit“ betrachtet. Bei vergleichbaren Response-Raten (10,5% für UFT / LV vs 9% für 5-FU / LV) lässt sich hinsichtlich der Wirksamkeitsparameter zur Überlebenszeit keine der beiden Behandlungsarten favorisieren. Bezüglich des Toxizitätsprofils finden sich jedoch leichte Vorteile beim UFT / LV Schema.

Tabelle 13: Behandlungseffekte UFT / LV vs 5-FU / LV [25]

In Monaten	UFT / LV	5-FU / LV	P	Hazard Ratio	NNT
Mediane PFÜZ	3,3	3,4	0,591	0,94 (0,75-1,18)	100
Mediane ÜZ	12,2	10,3	0,226	1,14 (0,92-1,42)	53

Studienperiode: Mai 1996 – Juli 1997 , Fallzahl: 380 Patienten

Hazard Ratio definiert als 5-FU / LV vs UFT / LV; 95%-Konfidenzintervall in Klammer

NNT = Number Needed to Treat, gerundet, eigene Berechnungen

4.3 Qualitative Gesamtbetrachtung

Die Wirksamkeit des betrachteten 5-FU / LV Standardprotokolls zeigt sich in erster Linie im Studiendesign gegen eine Kontrollgruppe mit Nulltherapie signifikant überlegen. In Studien, in denen genau dieses Schema seinerseits als Kontrollgruppe gegen eine Experimentalgruppe mit einem alternativen Behandlungsschema verglichen wird, zeigen sich heterogene Ergebnisse von äquivalenter oder teilweise höherer Wirksamkeit seitens der alternativen Behandlung. Eine spezifische Wirksamkeit für das 5-FU / LV Chemotherapieprotokoll kann in diesem Zusammenhang insofern nicht aufgezeigt werden.

In der folgenden Tabelle 14 ist das Gesamtergebnis der im vorherigen Abschnitt betrachteten Studien qualitativ abgebildet.

Die Vergleichsoperatoren deuten dabei die Relation der Wirksamkeitsergebnisse zwischen den Therapiegruppen hinsichtlich des Zielgrößenbündels Wirksamkeit, Verträglichkeit und Toxizität an, die Größe der Operatoren zeigt das Ausmaß des Effekts an.

Tabelle 14: Qualitative Gesamtergebnisse 5-FU / LV Schema

<p>5-FU / LV i.v.</p> <p>(5-FU 425 mg/m² i.v. + 20 mg/m² Folinsäure X 5 Tage alle 28 Tage über 6 / 12 Monate)</p>	\gg	<p>Beobachtung</p> <p>Nulltherapie</p>
	$< (=)$	<p>5-FU p.v.i.</p> <p>(300 mg/m²/Tag 5-FU p.v.i. über 12 Wochen)</p>
	$= (<)$	<p>UFT + LV oral</p> <p>(300 mg/m² UFT + 20 mg/m² LV oral über 28 Tage alle 35 Tage wiederholt)</p>
	$(>) =$	<p>5-FU + IFN s.c.</p> <p>(600 mg/m² 5-FU 24h Infusion X 5 Tage, anschl. 600 mg/m² 5-FU i.v. 1 X Woche + 3 X IFN 5MU s.c. über 6 Monate)</p>
	$= (<)$	<p>Irinotecan + 5-FU / LV</p> <p>(Irinotecan 125 mg/m² i.v. Infusion über 90 Minuten + Folinsäure 20 mg/m² i.v. Bolus + 5-Fluorouracil 500 mg/m² i.v. Bolus Tag 1,8,25,22 – alle 6 Wochen, 6 Zyklen)</p>
	$= (<)$	<p>5-FU c.i.</p> <p>(250 mg/m²/Tag 5-FU Dauerinfusion X 56 Tage alle 9 Wochen über 3 Zyklen)</p>

Im Vergleich zum natürlichen postoperativen Krankheitsverlauf, gemessen anhand einer Kontrollgruppe mit Nulltherapie, ist das betrachtete 5-FU / LV Therapieprotokoll signifikant wirksam. In Studien mit Kontrollgruppen unter kompetitiven Therapieansätzen kann eine vorteilhafte Stellung konsistent über alle Vergleiche nicht beobachtet werden. Die Ergebnisse bedeuten nicht zwangsläufig, dass das betrachtete Chemotherapieprotokoll unwirksam ist. Es könnte theoretisch folgendes Beziehungsgeflecht zwischen den Wirksamkeitsniveaus bei den verschiedenen Therapiestrategien gelten:

Alternative Therapie > 5-FU / FA >> Nulltherapie

Spezifität der betrachteten Therapie setzt jedoch voraus, dass das betrachtete folinsäurehaltige 5-FU Protokoll sich sowohl im Vergleich zur Nulltherapie (Beobachtung) als auch im Vergleich zur alternativen Behandlung als signifikant vorteilhaft erweist:

5-FU / FA > Alternative Therapie > Nulltherapie

In der Gesamtbetrachtung wird deutlich, dass verschiedene Chemotherapie-Protokolle ähnliche Messergebnisse erzeugen. Somit ist der beobachtete Effekt nicht spezifisch [45]. Es erscheint daher zweckmäßig, den therapiespezifischen Wirkungsmechanismus näher zu erörtern.

Trotz der hier vorliegenden Ergebnisse hat sich besagtes Therapieprotokoll weltweit als Standard durchgesetzt. Es wird jedoch zunehmend angezweifelt, ob ein (multimodales) Bolus 5-FU Schema aufgrund des möglichen Nebenwirkungsprofils und der unter epidemiologischen Gesichtspunkten limitierten Wirkung auch zukünftig einen therapeutischen Standard im (neo-) adjuvanten oder palliativen Ansatz darstellen kann, wenn die Wertigkeit der Chemotherapie nicht eindeutig belegt werden kann [52].

5. Diskussion

5.1 Methodische Forderungen für kontrollierte klinische Studien

Anhand kontrollierter klinischer Studien soll empirisch unter Verwendung konfirmatorischer statistischer Verfahren, das heißt auf Power-Berechnung basierender fallzahlenorientierter induktiver Testverfahren und multivariater Ansätze, aufgezeigt werden, dass ein bestimmtes definiertes therapeutisches Verfahren unter Praxisbedingungen wirksam und anwendbar ist.

Es existiert eine Reihe von Vorschlägen, Richtlinien und gesetzlichen Vorschriften (Deklaration von Helsinki, Deutsches Arzneimittelgesetz, Richtlinien der „Good Medical Practice“, Richtlinien der „International Conference on Harmonization“) zur methodischen Durchführung klinischer Studien.

Eine Studie bezieht sich zunächst auf eine über Ein- und Ausschlusskriterien genau bezeichnete Patientenpopulation. Über die Betrachtung dieser untersuchten Gruppe von Patienten hinaus ist dabei die Gültigkeit der Ergebnisse in Bezug auf alle potenziellen Patienten entscheidend. Man spricht in diesem Zusammenhang hinsichtlich der Übertragbarkeit der Ergebnisse von externer Validität. Das therapeutische Vorgehen sollte exakt beschrieben werden und nachvollziehbar sein. Gibt es für die Zielgruppe von Patienten eine Standardtherapie oder sonstige nachweislich wirksame Behandlung, so sollte die Wirkung der neuen Therapie auch mit der Wirkung dieser Standardtherapie verglichen werden. Die Güte einer klinischen Studie per se, das bedeutet, die Konsistenz der Ergebnisse ohne systematische Fehler oder logische Widersprüche, bezeichnet sich als interne Validität.

Der Erfüllungsgrad im Hinblick auf interne und externe Validität kann anhand verschiedener Kriterien abgeschätzt werden:

- Kontrollgruppenvergleich
- Genau definiertes Patientenkollektiv (Einschluss- und Ausschlusskriterien)
- Prospektives Studiendesign
- Randomisierung von Patienten
- Verblindung von Patienten, Prüfarzt und Analyst
- Kontrolle der Einhaltung der Therapiebedingungen durch die Patienten und Ärzte
- Auswahl für den Patienten relevanter klinischer Zielkriterien
- Ausführung konfirmatorischer, statistischer Analysemethoden

In diesem Zusammenhang wurde durch die Cochrane Collaboration© eine umfangreiche Prüfliste mit Kriterien zur Beurteilung der Studienqualität entwickelt. Anhand dieser lassen sich qualitative Aussagen bezüglich der Validität beispielsweise nach der Anzahl der erfüllten Kriterien ableiten. Ein kompakter Kriterienkatalog zur Abschätzung der Studienqualität bezüglich der in dieser Arbeit verwendeten Publikationen wurde in Kapitel 3 vorgestellt. Auf diese Weise können Publikationen auf systematische Fehler hin überprüft werden. Eine formale Erfüllung wichtiger Kriterien gewährleistet im Einzelfall jedoch nicht automatisch eine fehlerfreie Operationalisierung in der statistischen Analyse. Die konkrete Quantifizierung der Güte einer klinischen Studie anhand eines validen Scoring- und Berechnungssystems gestaltet sich schwierig. Die nachfolgenden Erörterungen sollen aufzeigen, dass im Zusammenhang mit definierten Qualitätskriterien im Einzelfall weitere Subkriterien wichtig sein können, die aber in der Durchführung klinischer Studien möglicherweise nicht optimal umgesetzt werden.

5.2 Methodische Durchführung klinischer Studien zur Überlebenszeit

5.2.1 Statistische Methoden

Klinische Studien zur Wirksamkeit erzeugen auf Basis einer Stichprobe einen genau definierten, mathematisch-statistischen Optimalitätskriterien genügenden Schätzwert bezüglich eines wahren, unbekanntem Effekts. Es existieren formale Eigenschaften bezüglich Studienaufbau und statistischer Methodik, welche den systematischen Versuchsfehler minimieren. Unter diesen Voraussetzungen wird die Abweichung der geschätzten Wirksamkeit einer medizinischen Behandlung vom wahren, unbekanntem Effekt bei zunehmendem Stichprobenumfang gering sein. Die zufällige Abweichung kann über ein adäquates Wahrscheinlichkeitsmodell abgebildet werden.

Studien zur klinischen Wirksamkeit von Chemotherapie basieren auf der Beobachtung der rezidivfreien Überlebenszeit (Disease Free Survival). Sie ist definiert als die Zeit von der Randomisierung bis zu einem definierten Ereignis wie Rezidiv, Zweitumor oder Tod ohne Rückfall. Zusätzlich wird die Gesamtüberlebenszeit (Overall Survival) gemessen. Diese berechnet sich aus der Zeitspanne von der Randomisierung bis zum Tod jeglicher Ursache. Als Nebenzielgrößen werden häufig Toxizität, Lebensqualität oder je nach Ausgestaltung der Studie weitere Parameter wie Ansprechrate oder progressionsfreie Überlebenszeit definiert. Messtechnisch ist die Erfassung einer rezidivfreien Überlebenszeit insofern problematisch, als dass diese von der Zuverlässigkeit und Genauigkeit der Untersuchungsmethoden in der Nachsorge abhängt. Rezidivfrei in diesem Sinne bedeutet nur, dass ein Rezidiv nicht beobachtet wurde, nicht aber unbedingt, dass es nicht vorlag. Insbesondere bei einem Design, welches nicht doppelblind durchgeführt wird, kann sich die Erwartungshaltung der Ärzte in einer unterschiedlichen Intensität der Nachbeobachtung bei den zu vergleichenden Gruppen auswirken: Eine Kontrollgruppe mit Nulltherapie wird möglicherweise einem genaueren Screening unterzogen, da von vornherein ein erhöhtes Rezidivrisiko vermutet wird. So werden in der Folge möglicherweise bei der Kontrollgruppe Rezidive mit höherer Wahrscheinlichkeit beobachtet.

Damit vergrößert sich jedoch auch der rechnerische Unterschied in der Effektgröße zwischen Kontrollgruppe und Experimentalgruppe. Falls eine Beobachtungsgleichheit in den Vergleichsgruppen nicht gewährleistet ist, könnte ein therapeutischer Vorteil in der Experimentalgruppe auch durch Unterschiede im Follow-Up verursacht sein.

Nicht zu guter Letzt handelt es sich bei den mit statistischen Schätzmethoden berechneten Effekten um Durchschnittswerte, die langfristig, mit wachsendem Stichprobenumfang eine gute Näherung bezüglich des unbekanntem wahren Effekts ergeben. In der ärztlichen Praxis können sich für einen patientenspezifischen Einzelfall davon abweichende Ergebnisse einstellen.

5.2.1.1 Univariate Verfahren

Im Rahmen von kontrollierten Therapiestudien werden für quantitative Zielgrößen wie Überlebenszeit methodisch-rechnerisch insbesondere gruppenvergleichende Verfahren für Lageparameter (Mittelwert, Median) zur Signifikanzprüfung beobachteter Unterschiede verwendet. Grundsätzlich unterscheidet man zwischen parametrischen, bezüglich der Zielgröße eine bestimmte Verteilungsform voraussetzenden Verfahren und nichtparametrischen, verteilungsfreien Verfahren. In Abhängigkeit von der Anzahl der zu vergleichenden Therapiegruppen umfassen die parametrischen Modelle den t-Test für zwei unabhängige Stichproben sowie die einfache Varianzanalyse für mehr als zwei Gruppen. Analog dazu beinhalten die nichtparametrischen Modelle den U-Test von Mann, Whitney und Wilcoxon für den Zweigruppen-Fall sowie den Kruskal-Wallis Test für den Vergleich von mehr als zwei Gruppen.

Über parametrische Verfahren lässt sich der zu überprüfende Effekt in einer Testhypothese genau quantifizieren, inhaltliche Folgerungen lassen sich besser konkretisieren. Ferner kann in diesem Modellrahmen ein Wertebereich als Konfidenzintervall angegeben werden, um das wahre Ausmaß des Effekts beziehungsweise des Unterschiedes zwischen verschiedenen Therapiegruppen abzuschätzen.

Nachteilig wirkt sich aus, dass die Modellannahmen zumindest approximativ erfüllt sein sollten und im Einzelfall gegebenenfalls rechnerische Transformationen der Zielgrößen beispielsweise durch Logarithmierung erforderlich sein können.

Sofern die zugrunde liegende Messgröße ordinal-skaliert ist (beispielsweise ein Therapieerfolg kategorisiert in abgestufte Wirkungsniveaus) oder stark asymmetrisch verteilt ist, bieten sich aufgrund der schwächeren Modellvoraussetzungen nichtparametrische Modelle zur Überprüfung der klinischen Hypothesen an. Der dabei angewandte Rechenalgorithmus basiert allerdings nicht mehr auf den Beobachtungsdaten selbst, sondern auf rangtransformierten Daten einer der Größe nach geordneten Stichprobe.

Die Übersetzung eines statistischen Signifikanzergebnisses in klinische Schlussfolgerungen gestaltet sich dadurch weniger eindeutig als im Fall parametrischer Modelle. Weiterhin steht dem Vorteil weniger restriktiver Modellannahmen bei den nichtparametrischen Verfahren der Nachteil eines konstruktionsbedingt weniger trennscharfen Signifikanzergebnisses gegenüber: Der Unterschied beispielsweise zwischen zwei Therapieverfahren operationalisiert in einer quantitativ messbaren Erfolgsgröße wird hierbei in den zugrunde liegenden Hypothesen nicht mehr genau quantifiziert.

Im Ergebnis kann nur noch ein Unterschied statistisch gestützt und die Richtung aufgezeigt werden. In einer Gesamtwürdigung zur Anwendung von Lokalisationstests ist auch anzumerken, dass es sich nur um punktuelle Betrachtungen zu mittleren beziehungsweise medianen Unterschieden in den Zielgrößen handelt und insofern eine Aussage zur therapeutischen Dynamik im engeren Sinne nicht ermöglicht.

Tabelle 15: Gruppenvergleichende Testverfahren zu Lageparametern

	Normalverteilung	Verteilung unbekannt
2 Stichproben	Zweistichproben t-Test	U-Test
Mehr als 2 Stichproben	Einfache Varianzanalyse	Kruskal-Wallis-Test

Ein univariates Standardinstrument für qualitative Zielgrößen, wie beispielsweise in Überlebenszeitstudien die dichotome Zielgröße mit den Ausprägungen „5-Jahre überlebt“ versus „5 Jahre nicht überlebt“, sind χ^2 -Test und Fishers exakter Test.

Man interessiert sich in diesem Zusammenhang schwerpunktmäßig für Häufigkeiten bezüglich bestimmter Ereigniskonstellationen der Zielgröße mit definierten Faktoren. Im Vordergrund steht die Frage, ob ein in der Stichprobe beobachtetes Häufigkeitsmuster zufällig entstanden ist oder ob man einen systematischen Zusammenhang zwischen den betrachteten Größen vermuten kann. Zur Quantifizierung eines Effekts, beispielsweise als Unterschied von Überlebensraten in verschiedenen Therapiegruppen eignet sich der Binomialtest zum Vergleich zweier relativer Häufigkeiten. Nachteilig bei diesen Verfahren wirkt sich aus, dass die Ergebnisse zum einen mit veränderten Häufigkeiten in unterschiedlichen Stichproben stark variieren können. Möglicherweise lässt sich ein komplexer klinischer Sachverhalt nicht auf einer derart abstrakten qualitativen Datenskala abbilden.

Eine klinische relevante Bedeutung im Zusammenhang mit Therapiestudien hat insbesondere auch der Verlauf von Überlebenszeiten. Ein Standardinstrument hierzu ist die Überlebenszeitanalyse nach Kaplan-Meier. Hierbei wird eine Überlebenszeitkurve empirisch, auf Basis der beobachteten Todesfälle ermittelt. Die resultierende Kurve gibt die Wahrscheinlichkeit an, mindestens bis zu einem definierten Zeitpunkt zu überleben. In diesem Modellrahmen kann auch die Bedeutung von Einflussfaktoren auf den Verlauf der Überlebenszeitkurve untersucht werden. Dazu muss der betrachtete Faktor als nominale oder kategoriale Größe vorliegen. Es werden die Kurven getrennt nach jeder Kategorie berechnet und anschließend in einem gemeinsamen Schaubild abgetragen. Die statistische Überprüfung eines Unterschiedes im Verlauf erfolgt anhand des Logrank Tests.

Eine ausführliche Abhandlung dieser Methoden anhand von medizinischen Beispielen findet sich in [14,20].

5.2.1.2 Multivariate Verfahren

Problematisch bei der Anwendung univariater Methoden ist die Beschränkung auf jeweils einen zentralen Parameter ohne Berücksichtigung relevanter Nebenfaktoren (Confounder-Größen). Sofern diese bekannt sind oder valide Surrogatgrößen verwendet werden, können stratifizierte Analysen getrennt nach Subgruppen durchgeführt werden.

Da bei der Analyse klinischer Zusammenhänge eine Reduzierung der Sachhypothesen auf einen, genau bestimmbar, zentralen Parameter oft nicht möglich ist, werden bei diesen Untersuchungen häufiger multivariate Ansätze eingesetzt. Solche Methoden lassen Aussagen bezüglich einer Zielgröße unter simultaner Berücksichtigung mehrerer Einflussfaktoren und deren Interaktionen zu. Insbesondere kann eine genauere Bestimmung des Erklärungspotenzials einzelner Größen erfolgen. Diese Effekte werden im univariaten Ansatz möglicherweise nicht zuverlässig geschätzt. Jedoch erfordern diese Methoden deutlich höhere Fallzahlen als univariate Verfahren. Insbesondere dann, wenn diese Verfahren als Prognoseinstrument zur Vorhersage eines Zielereignisses gegeben spezifischer Werte für die prädiktiven Faktoren genutzt werden, ist eine konsistente Schätzung der Modellparameter unerlässlich.

Hervorzuheben unter Berücksichtigung der bisherigen Ausführungen ist das multivariate logistische Regressionsmodell [14,16]. Hier wird die Eintrittswahrscheinlichkeit eines Zielereignisses (beispielsweise der Überlebenswahrscheinlichkeit) in einem nicht-linearen Modell abgebildet. Über die Logit-Transformation wird der logarithmierte Odds, ein Verhältnis der Wahrscheinlichkeit des Eintretens eines Ereignisses zur Wahrscheinlichkeit des Nicht-Eintretens als lineare Funktion verschiedener prognostischer Faktoren formuliert. Die Modellkoeffizienten errechnen sich über ein iteratives numerisches Optimierungsverfahren. Dabei erfolgt die Bestimmung der Parameter derart, dass die Wahrscheinlichkeit, Daten wie in der Stichprobe beobachtet zu erzeugen, maximiert wird. Als Lösung erhält man die Maximum-Likelihood-Schätzer, welche die Veränderung der Odds anzeigen, wenn sich die entsprechende unabhängige Variable um eine Einheit erhöht.

Anschaulicher speziell im Fall dichotomer exogener Größen (beispielsweise Novumtherapie und Standardtherapie) ist das Odds Ratio ein Maß für die Vervielfachung der Eintrittswahrscheinlichkeit des Zielereignisses unter dem Verhältnis der beiden Ausprägungen des Faktors. In Therapiestudien kann so beispielsweise die Abhängigkeit des Überlebensstatus von demografischen, anamnestischen und therapeutischen Größen ausgedrückt werden kann. Die Güte der Anpassung des Modells an die Daten wird unter anderem mit einem speziellen Test (Likelihood-Quotienten-Test) überprüft.

Als Erweiterung der univariaten Methoden zum Vergleich von Erwartungswerten (t-Test und einfache Varianzanalyse) ist das multivariate varianzanalytische Modell anwendbar. Im Allgemeinen Linearen Modell können mittlere Unterschiede bezüglich einer Zielgröße als Funktion verschiedener klinisch-relevanter Faktoren erklärt werden. Die Besonderheit dieses Ansatzes liegt darin, eine simultane Schätzung des Einflusses einzelner Variablen unter Berücksichtigung von Kreuzeffekten als Interaktion einer Erklärungsgröße mit einer anderen zu ermöglichen. Rechnerisch beruht diese Technik auf den faktorenspezifischen Quadratsummenzerlegungen, der Summe der quadrierten Abweichungen der Beobachtungswerte vom jeweiligen Bezugsmittelwert [5,14,16].

Die Cox-Regression stellt eine Erweiterung der Überlebenszeitanalyse nach Kaplan-Meier dar [5,16]. Im Rahmen dieses Ansatzes kann die Wirkung einer Therapie auf die Überlebenszeit adjustiert um den Einfluss weiterer relevanter klinischer Faktoren abgebildet werden. Eine bestimmte Verteilungsform der Zielgröße „Überlebenszeit“ wird dabei jedoch nicht gefordert. Insofern handelt es sich um einen semiparametrischen Ansatz. Schwerpunkt ist die Schätzung der Hazard-Funktion in Abhängigkeit verschiedener Einflussfaktoren. Dabei bezeichnet der Hazard die Wahrscheinlichkeit, innerhalb eines Zeitintervalls das Zielereignis (z.B. Tod) zu realisieren unter der Voraussetzung, dass dieses bis zu Beginn dieses Zeitintervalls noch nicht eingetreten war. Das Standardmodell der Cox-Regression geht von der Annahme zeitkonstanter, proportionaler Hazardraten aus, das bedeutet, der Quotient der Hazards beispielsweise zweier Therapiegruppen wird als zeitkonstant betrachtet. Die Annahme zeitkonstanter Therapieeffekte jedoch könnte sich als ungeeignete Modellannahme erweisen.

Eine derartige Situation entsteht dann, wenn solche Effekte nicht gleichmäßig über die Zeit wirken, sondern beispielsweise nur während einer Frühphase. Eine anwendungsbezogene Einführung mit einschlägigen medizinischen Studienbeispielen ist beispielsweise in [16] gegeben.

Jedes einzelne methodische Verfahren kann in Bezug auf eine konkrete Studien- und Stichprobenkonstellation Vor- und Nachteile mit entsprechenden Implikationen für die Interpretation der Ergebnisse aufweisen. Daher erscheint es unter Berücksichtigung der komplexen klinischen Zusammenhänge nicht zweckmäßig, bestimmte Verfahren selektiv einzusetzen, es sei denn, dies ist durch die Art der Zielgrößen, Fallzahl, Definition der Fragestellung oder besonderen Eigenschaften des Studiendesigns bedingt.

Wenn möglich sollte das insgesamt verfügbare Methodenpotenzial ausgeschöpft werden, die Ergebnisse im Gesamtzusammenhang dargestellt und interpretiert werden. Zur Betrachtung von Langzeiteffekten erscheint es sinnvoll, sequentiell nach bestimmten zeitlichen Fixpunkten, beispielsweise nach drei, fünf und 10 Jahren, Analysen durchzuführen.

Gelegentlich wird auf der Basis einer hohen Korrelation zwischen der 3-Jahres-Überlebensrate und der 5-Jahres-Überlebensrate argumentiert, bereits in einem relativ überschaubaren Beobachtungszeitraum von drei Jahren valide Ergebnisse zu erhalten [51]. Sofern keine technischen, ökonomischen oder juristischen Gründe zwingend dagegen sprechen, wäre jedoch die Ausweitung der Analysen auf einen längeren Beobachtungszeitraum grundsätzlich anzuwenden, um die Dynamik eines Wirksamkeitseffekts zu betrachten. Zweiseitiges Testen speziell beim Vergleich zweier Therapien ist unter einem konservativen Ansatz grundsätzlich empfehlenswert. Beim Vergleich einer aktiven neuen Therapie mit einer Nulltherapie, bei welcher die relative Risikoreduktion in einer bestimmten Richtung antizipiert werden kann, kann jedoch auf zweiseitigem Niveau im Grenzfall der Wirksamkeitsunterschied statistisch insignifikant sein, obgleich in Wahrheit ein Therapieunterschied vorhanden ist. Hier könnte ein einseitiges Testen zweckmäßig sein, sollte im Einzelfall jedoch begründet werden.

Ein besonderes Problem im Zusammenhang mit Überlebenszeitstudien sind Zensierungen verursacht durch Studienabbruch einzelner Patienten oder Ausbleiben des Zielereignisses zum Beobachtungsende. Bei zensierten Daten liegt nur die reine Beobachtungszeit als Information vor. Sofern diese als Überlebenszeit interpretiert würde, ließen sich sämtliche Daten zur Analyse verwenden. Wenn jedoch das Ausmaß zensierter Werte am Gesamtvolumen der Patientendaten einen hohen Anteil annimmt, könnte dies eine Unterschätzung des wahren Effekts zur Folge haben. Zensierungen aufgrund der Stichtagswahl zum Beobachtungsende lassen sich durch Ausweitung der Beobachtungsperiode vermeiden. Zensierungen aufgrund Studienabbruchs sind an dieser Stelle problematischer. Über geeignet definierte Analysegruppen (Per-Protocol Patienten) wird ein derartiges Patientengut aus den Berechnungen ausgeschlossen. Eine standardisierte Handhabung solcher Fälle erscheint in diesem Zusammenhang zweckmäßig.

Analysestrategien im Hinblick auf ein definiertes Patientenkollektiv spielen für die Interpretierbarkeit der gefundenen Ergebnisse und deren Verallgemeinerung eine erhebliche Rolle. Im Rahmen kontrollierter klinischer Studien wird eine Therapiewirkung für eine selektive Patientengruppe unter „Laborbedingungen“ beobachtet. Die Randomisierung soll hinsichtlich definierter Eigenschaften strukturgleiche Gruppen sicherstellen. Es kann aufgrund von Therapieprotokollverletzungen oder unerwünschten Arzneimittelwirkungen dazu kommen, dass Patienten nicht in der ihnen zu Beginn der Studie zugewiesenen Gruppe verbleiben. In diesem Zusammenhang werden speziell zwei verschiedene Analysekollektive gebildet [20].

Beim Ansatz des „Intention-to-Treat“ (ITT) werden alle Patienten in der Gruppe, in welche sie anfangs randomisiert zugeteilt worden sind, für die Analyse berücksichtigt. Hierbei bleibt die Strukturgleichheit erhalten, therapiespezifische Unterschiede lassen sich jedoch weniger trennscharf identifizieren.

Das „Per-Protocol Prinzip“ (PP) schließt alle nicht protokollgemäß behandelten Patienten von der Analyse aus. Die Strukturgleichheit ist nicht mehr gewährleistet, vorhandene Unterschiede treten deutlicher hervor, werden jedoch möglicherweise überschätzt.

Grundsätzlich sollten nach den Empfehlungen des ICH [8] beide Ansätze ITT und PP verfolgt werden, um eine über die Bedingungen des kontrollierten Experiments hinausgehende realistische Einschätzung der Wirksamkeit zu erzeugen.

5.2.2 Messkonzept und Kausalität

Die Durchführung klinischer Therapiestudien basiert auf der Vermutung, dass ein direkter Zusammenhang zwischen der in den Studien untersuchten medizinischen Intervention (adjuvante Chemotherapie) und der Zielgröße (Gesamtüberlebenszeit, rezidivfreie Überlebenszeit) besteht. Aus einem beobachteten Effekt wird auf den wahren Effekt geschlossen. Wenn jedoch die Zielgröße nicht ausschließlich über den vermuteten Transmissionsmechanismus in der Regel physikalisch-chemischer Art bestimmt wird, sondern über weitere, nicht direkt erfassbare Faktoren, wird auch bei Auswahl der geeigneten statistisch-mathematischen Methoden ein Zusammenhang nicht optimal geschätzt werden können. Im Ergebnis entstehen dann möglicherweise fehlerhafte Schlussfolgerungen.

Für die Bewertung eines kausalen Zusammenhangs zwischen einer medizinischen Intervention (Testtherapie) und einer Zielgröße ist eine Reihe von Prinzipien unabdingbar [24]. Ursprünglich als Minimalanforderungen an den Nachweis eines Kausalzusammenhangs in epidemiologischen Untersuchungen konzipiert, umfasst dieser Katalog neun Aspekte:

1. vorgelagerte zeitliche Beziehung zwischen einer relevanten Exposition oder Intervention und dem interessierenden beobachtbaren Ergebnis
2. ausreichende, klinisch-relevante Stärke des Zusammenhangs mit geeigneten Methoden gemessen
3. Dosis-Wirkungs-Beziehung als Zusammenhang zwischen Intensität der Exposition (Intervention) und dem Niveau des Ergebnisses
4. Konsistenz als Reproduzierbarkeit des Zusammenhangs unter verschiedenartigen Studienbedingungen und Methoden

5. Plausibilität als konsensfähiges Verständnis zum Wirkungsmechanismus klinisch-biologischer Prozesse)
6. Abwägung alternativer kompetitiver Erklärungsansätze
7. experimentellen Charakter mit der Möglichkeit zur Veränderbarkeit der Versuchsbedingungen
8. Spezifität als Zusammenhang zwischen einem bestimmten Faktor und einem spezifischen Ergebnis
9. Kohärenz im Sinne einer Kompatibilität mit richtungsweisender Theorie

Für die Beurteilung der Kausalität im Rahmen der vorliegenden Fragestellung zur Spezifität einer Behandlung in Therapiestudien sind nicht alle Punkte gleichermaßen relevant oder anwendbar.

So ist Punkt 1 bei Therapiestudien grundsätzlich erfüllt, denn die medizinische Intervention geht der Beobachtung beziehungsweise statistischen Erfassung des Zielereignisses naturgemäß voraus. Die Aspekte Konsistenz und Spezifität überschneiden sich teilweise, denn eine spezifische Wirksamkeit findet sich konsistent unter verschiedenen Studienbedingungen hinsichtlich Art der Kontrollgruppe oder Art des betrachteten Patientenkollektivs. Auch lassen sich die Punkte 5 und 6 inhaltlich zusammenfassen: Ein bestimmter kausaler Zusammenhang erscheint plausibel, sofern das beobachtete Ergebnis durch keine alternativen Erklärungsmodelle abgebildet werden kann. Der Kriterienkatalog für die vorliegende Untersuchung kann im Zusammenhang mit der betrachteten Thematik auf fünf Hauptpunkte reduziert werden [11]:

- Ausreichende biologische Plausibilität
- Definition einer hinreichenden Größe des zu beobachtenden Effekts
- Nachweis einer Dosis-Wirkungsbeziehung
- Nachweis einer therapiespezifischen Wirksamkeit
- Abwägung alternativer Erklärungsansätze

Für die im Rahmen dieser Arbeit betrachteten Therapieansätze gilt der Aspekt zur biologischen Plausibilität als gesichert. Aufgrund theoretischer Überlegungen und Laboruntersuchungen zu mikrobiologischen Prozessen lässt sich eine pharmakologische Wirkung der medikamentösen Behandlung mit Fluorouracil und Leucovorin vermuten [42].

Ein klinisch relevantes Ausmaß in Bezug auf den zu erwartenden klinischen Effekt kann nur als teilweise erfüllt angesehen werden: Es fehlt eine übergreifende stringente Vorgehensweise bestehend aus der Berechnung konfirmatorischer Statistiken mit Fallzahlenschätzung unter Zugrundelegung einer genau definierten Effektgröße und anschließender Diskussion der klinischen Bedeutung der Ergebnisse. Auch kleine Überlebensvorteile der Testtherapie verbunden mit statistisch signifikanten Ergebnissen werden gelegentlich als inhaltlich relevant und wirksamkeitsspezifisch gewertet. Dieses Kriterium wird zusätzlich dadurch aufgeweicht, dass zunehmend die Möglichkeit angedacht wird, anstelle des traditionellen 5-Jahres-Zeithorizontes für die primäre Zielgröße „Gesamtüberlebenszeit“ die „rezidivfreie Überlebenszeit“ über einen 3-Jahres-Zeithorizont als stellvertretende Hauptzielgröße einzusetzen. Kostenorientierte Überlegungen und Untersuchungen zur statistischen Korrelation zwischen den Zielgrößen können ein derartiges Vorgehen rechtfertigen [51]. Eine indikationsgebietübergreifende Diskussion diesbezüglich findet sich in [30], wo die Autoren eine grundlegende Abwägung unter Berücksichtigung praktischer, ökonomischer und inhaltlicher Gesichtspunkte vornehmen. Im Ergebnis ist davon auszugehen, dass jede der Zielgrößen einen klinisch-inhaltlichen Schwerpunkt bildet und für eine abschließende Bewertung die kombinierte Betrachtung unerlässlich ist.

Eine Dosis-Wirkungs-Beziehung ist ebenfalls nur bedingt nachvollziehbar. Speziell im Hinblick auf das Standard-Chemotherapieprotokoll fehlen Anzeichen dafür, dass über ein Basisniveau hinausgehend weder Höhe der 5-FU-Dosis noch Ausmaß der verabreichten Folsäure eine zusätzliche Therapiewirkungen erzeugen [27,42,34].

Eine spezifische Wirksamkeit des betrachteten Chemotherapie-Protokolls mit 5-FU / LV konsistent unter verschiedenen Studienbedingungen, das heißt unter Berücksichtigung verschiedenartiger Kontrollgruppen und Subkollektiven von Patienten, kann anhand der hier dargestellten Ergebnisse und Ausführungen im Ergebnisteil (Kapitel 4.1 – 4.3) nicht aufgezeigt werden.

Sämtliche Studien zur therapeutischen Wirksamkeit der adjuvanten Chemotherapie basieren auf der Überzeugung eines vollständig oder mehrheitlich pharmakologischen Zusammenhangs. Alternative Transmissionsmechanismen bezüglich des Einflusses nicht-pharmakologischer, psychosozialer Kontextfaktoren werden nicht oder nur unzureichend in Betrachtung bezogen. Die Bedeutung solcher Kontext-Faktoren wird konditioniert über die Art der Versuchsanordnung (Art der Kontrollgruppe und Art der Verblindung). Auch wenn prinzipiell eine pharmakologische Wirkung nicht in Frage gestellt wird, so zeigt sich, dass die beobachteten beziehungsweise gemessenen Ergebnisse einen kausalen Zusammenhang nicht uneingeschränkt stützen.

5.2.3 Der p-Wert als mathematisch-statistisches Entscheidungskriterium

Den Schwerpunkt der statistischen Argumentation in einer klinischen Studie bildet der p-Wert. Ein kleiner p-Wert (in der Regel $< 0,05$) wird als signifikantes Ergebnis gedeutet. Das bedeutet, der in der Stichprobe beobachtete Effekt wird als allgemein, über die konkrete Stichprobe hinaus gültig betrachtet. In Bezug auf die Fragestellung eines therapiespezifischen Wirksamkeitsunterschiedes zwischen zwei Behandlungsvarianten würde ein kleiner p-Wert als Wirksamkeitsunterschied interpretiert werden.

Die Bestimmung des p-Wertes basiert auf der rechnerischen Entwicklung eines statistischen Tests. Unter Vorgabe der Fragestellung erfolgt die Formulierung zweier komplementärer Hypothesen (Null- und Alternativhypothese). In therapievergleichenden Studien wird bei einer zweiseitigen Fragestellung ohne nähere Spezifizierung der Richtung eines Unterschiedes in der Nullhypothese eine Gleichheit, in der Alternativhypothese eine Ungleichheit der Studientherapien postuliert. Ziel ist die Falsifizierung der Nullhypothese.

Aus einer Zufallsstichprobe wird anhand der Stichprobendaten das für die Fragestellung relevante Effektmaß (arithmetisches Mittel, Stichprobenvarianz, relative Häufigkeit) geschätzt. Nach Maßgabe des Skalenniveaus und der Verteilung der Zielgröße lässt sich das Effektmaß bei Gültigkeit der Nullhypothese durch eine geeignete Normierung in eine Zufallsgröße (Teststatistik) transformieren. Damit kann für diese Prüfgröße ein theoretischer Wert, den man bei Richtigkeit der Nullhypothese erwarten würde, ermittelt werden. Je weiter dieser von dem durch die Stichprobendaten errechneten tatsächlichen Prüfgrößenwert abweicht, desto unwahrscheinlicher ist die Nullhypothese. Man verwirft die Nullhypothese zugunsten der Alternativhypothese.

Das Entscheidungskriterium zur Beurteilung der „statistischen Glaubwürdigkeit“ der Nullhypothese erfolgt anhand eines Schwellenwertes, in der folgenden Abbildung 5 als vertikale Linie mit dem Abszissenwert $t_{n-1;1-\alpha/2}$ beziehungsweise $-t_{n-1;1-\alpha/2}$ gekennzeichnet. Dieser teilt die Gesamtfläche unter der Dichtefunktion der Prüfverteilung in einen Annahmehereich mit der Fläche $1-\alpha$ und einen Ablehnungsbereich mit der Fläche α . Dieser kritische Wert wird durch die Wahl eines geeigneten Niveaus bezüglich des Fehlers 1. Art (α) sowie durch das zugrunde liegende Wahrscheinlichkeitsmodell eindeutig bestimmt. Das α -Niveau wird unter Abwägung der Konsequenzen einer Fehlentscheidung vor Versuchsbeginn festgelegt - in der Regel 5% [20]. Bei einer zweiseitigen Fragestellung wird die Nullhypothese im unteren 2,5%-Bereich (die Prüfgröße nimmt einen extrem niedrigen Wert an) und im oberen 2,5%-Bereich (die Prüfgröße nimmt einen extrem hohen Wert an) verworfen.

Dies bedeutet, dass bei wiederholter Durchführung des Experiments der Anwender davon ausgehen darf, in höchstens 5% der Fälle die Nullhypothese fälschlicherweise zu verwerfen. Die Ablehnung der Nullhypothese beinhaltet jedoch konstruktionsbedingt auch die Möglichkeit eines Entscheidungsirrtums: Die Nullhypothese wird fälschlicherweise verworfen. Die Wahrscheinlichkeit hierfür ist klein (kleiner als α), aber existent.

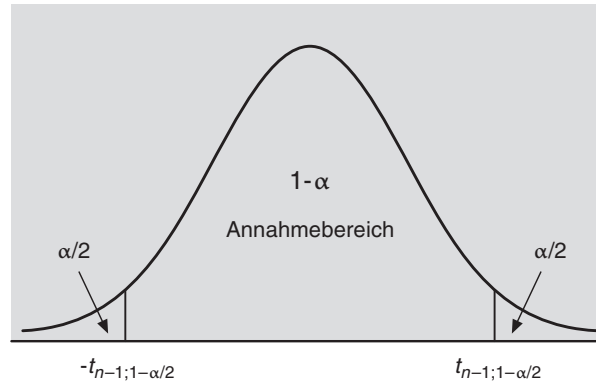


Abbildung 5: Prüfverteilung: Annahme- und Verwerfungsbereich [20]

In der Anwendung statistischer Software wird ein p-Wert berechnet, anhand dessen die Entscheidung für eine der Hypothesen getroffen wird. Dieser quantifiziert die Wahrscheinlichkeit, dass die Prüfgröße den beobachteten oder einen extremeren Wert annimmt beziehungsweise dass diese wertmäßig im Ablehnungsbereich liegt (in Abbildung 5 jenseits der $t_{n-1;1-\alpha/2}$ Kennzeichnung). Damit lässt sich die bekannte Entscheidungsregel ableiten: Die Nullhypothese wird beibehalten, sofern der p-Wert das vorgegebene Signifikanzniveau α übersteigt. Die Nullhypothese wird verworfen, sofern der p-Wert das α -Niveau unterschreitet.

Es ergeben sich zwei wichtige Aspekte bei der Beurteilung der statistischen Ergebnisse klinischer Studien im Zusammenhang mit dem p-Wert.

Der α -Fehler wird subjektiv vom Anwender unter Abwägung der Konsequenzen einer Fehlentscheidung auf ein bestimmtes Niveau festgelegt. Anhand dieses variablen Kriteriums erfolgt nach Maßgabe des p-Wertes eine Entscheidung zugunsten der Alternativ- oder Nullhypothese. Es existiert somit keine Signifikanz im absoluten Sinne, da der Anwender das Kriterium individuell aufgrund spezifischer Besonderheiten einzelner Studien anpassen kann. In wissenschaftlichen Untersuchungen hat sich ein Niveau von 5% als kritischer Wert durchgesetzt. Dies dient zwar einer Vereinheitlichung in unterschiedlichen Studien, aber die Festlegung auf einem bestimmten Niveau kann nicht den unterschiedlichen Bedingungen verschiedener Studien insbesondere hinsichtlich Messniveau der Zielgrößen und Stichprobenumfang gerecht werden.

Die rechnerische statistische Signifikanz bezieht sich auf eine in der Stichprobe gemessene Effektgröße. Statistische Tests, insbesondere parametrische, reagieren sensibel auf die Höhe des Stichprobenumfangs. Ein hoher Stichprobenumfang erhöht in der Regel die Wahrscheinlichkeit, Unterschiede aufzudecken und die Alternativhypothese anzunehmen. Sofern die statistische Signifikanz nicht mit einer entsprechenden klinischen Relevanz assoziiert ist, ist auch das Gesamtergebnis nicht schlüssig. In Studien zum Nachweis therapeutischer Überlegenheit sollte daher grundsätzlich die Methodik unter Berücksichtigung eines klinischen relevanten Mindestunterschiedes durchgeführt werden.

5.3 Wirkungsebenen einer medizinischen Therapie und Art der Vergleichsgruppe

Die medizinische Intervention beim kolorektalen Karzinom spielt sich auf zwei Ebenen ab: Nach erfolgreichem chirurgischen Eingriff (R0-Resektion) in einer ersten Stufe wird dem Patienten auf der zweiten Stufe der Behandlung adjuvante Chemotherapie verabreicht, um das Rezidivrisiko zu verringern und die Überlebenszeit zu verlängern.

Die Gesamtwirkung lässt sich in zwei Dimensionen - eine physikalisch-chemische und eine psychosoziale - abbilden. Zum therapeutischen Gesamterfolg trägt gewissermaßen ein Faktorenbündel bei - vereinfacht auf die genannten zwei Dimensionen reduziert. Es wäre nicht schlüssig, eine substitutionale Beziehung zwischen diesen zu unterstellen. Vielmehr könnte man davon ausgehen, dass es eine Art optimales Mischungsverhältnis beider gibt. Der pharmakologische Bestandteil nimmt möglicherweise ein größeres Gewicht ein, kann sich jedoch ohne eine Ergänzung über psychosoziale Maßnahmen nicht in wünschenswerter Weise entfalten. Eine sinnvolle klinische und ökonomische Bewertung von Gesundheitsleistungen setzt die Kenntnis der Bedeutung der einzelnen pharmakologischen und psychosozialen Faktoren voraus.

Abbildung 6 zeigt die in kontrollierten klinischen Studien entfaltbare Therapie beim kolorektalen Karzinom als zweistufigen Prozess einer kurativen Operation mit anschließender adjuvanter Folgetherapie, von dessen konkreter Ausgestaltung abhängig verschiedene Transmissionsmechanismen ineinander greifen.

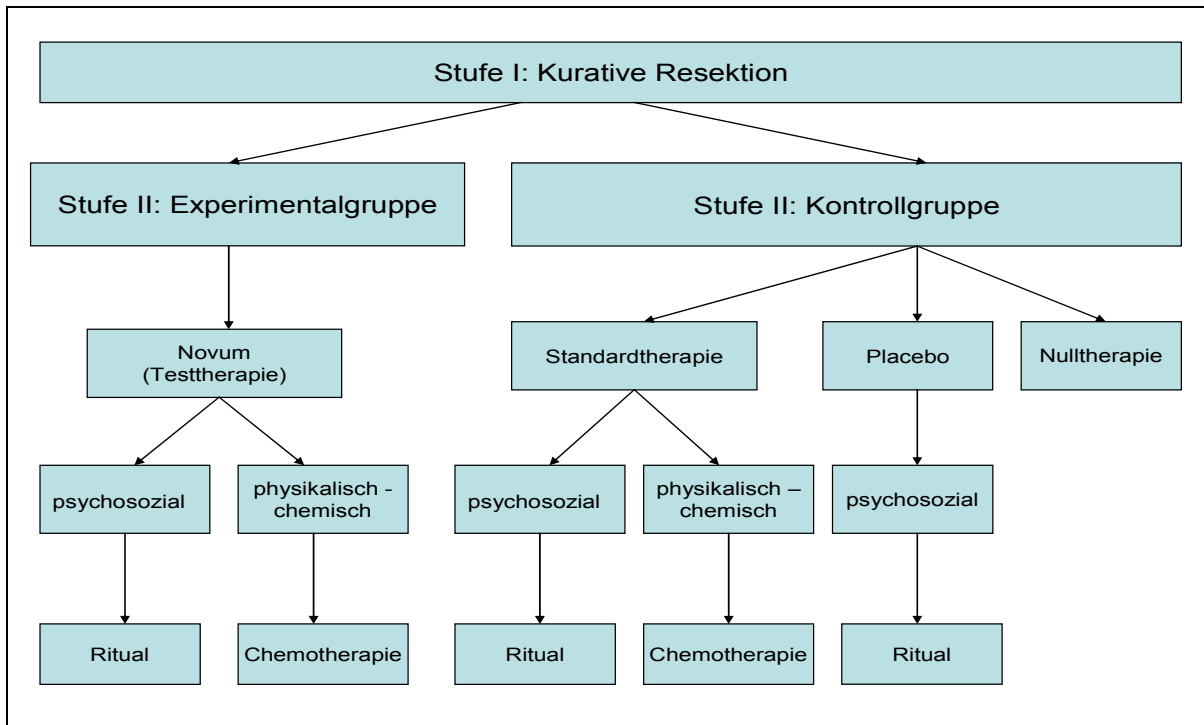


Abbildung 6: Behandlungsebenen und Studiendesign, eigene Darstellung

Bei einer Nulltherapie-Patientengruppe, welche ausschließlich einer operativen Maßnahme unterzogen wird, löst sich das pharmakologische und psychosoziale Wirkungsspektrum auf. Diese Situation ist anhand des rechten Behandlungspfads im Diagramm aufgeführt (Arm 3: Nulltherapie). Eine placeboinduzierte Therapie nährt ihren Erfolg gänzlich über psychosoziale Faktoren (Arm 2: Placebo). Die aktive Therapie (Novum, Testtherapie oder Standardtherapie (Arm 1 und 2) schließlich ist sowohl über physikalisch-chemische als auch psychosoziale Faktoren wirksam.

Obigen Ausführungen folgend kann ein in kontrollierten klinischen Studien beobachteter Gesamteffekt einer Therapie vereinfacht in zwei Wirkungsdimensionen aufgeschlüsselt werden, in eine pharmakologische Komponente (physikalisch-chemische Therapiewirkung) und eine psychosoziale. Die Argumentation und entsprechenden Schlussfolgerungen in klinischen Therapiestudien basieren jedoch auf der Annahme einer (nahezu) ausschließlich pharmakologischen Wirkungsweise.

5.4 Ganzheitliche Wirkungsweise einer medizinischen Intervention

Die im Rahmen der Durchführung von kontrollierten klinischen Studien gemessenen Wirksamkeitseffekte sind entweder auf die physikalisch-chemische Intervention, auf psychosoziale Mechanismen oder auf eine Überlagerung beider Einflüsse zurückzuführen. Zur Abgrenzung des wahren Behandlungserfolges erscheint es zweckmäßig, einen Kontexteffekt zu berücksichtigen. Dieser könnte alle psychologischen und weiteren nicht-pharmakologischen Einflüsse vor, während und nach der eigentlichen Therapiemaßnahme beinhalten. Die Gesamtwirkung einer therapeutischen Intervention ist der „Zugewinn an Besserung“ im Vergleich zu dem Ergebnis, was sich im natürlichen Krankheitsverlauf ohne Therapie ergeben hätte.

Die folgende Abbildung 7 zeigt in schematischer Darstellung den Gesamteffekt als Summe einer therapiespezifischen Wirkung, unspezifischer Reaktionen und des Ergebnisses eines natürlichen Krankheitsverlaufs. Die Gesamtheit unspezifischer Reaktionen umfasst dabei affektive Faktoren wie Suggestion oder Auswirkung der Erwartungshaltung des Patienten.

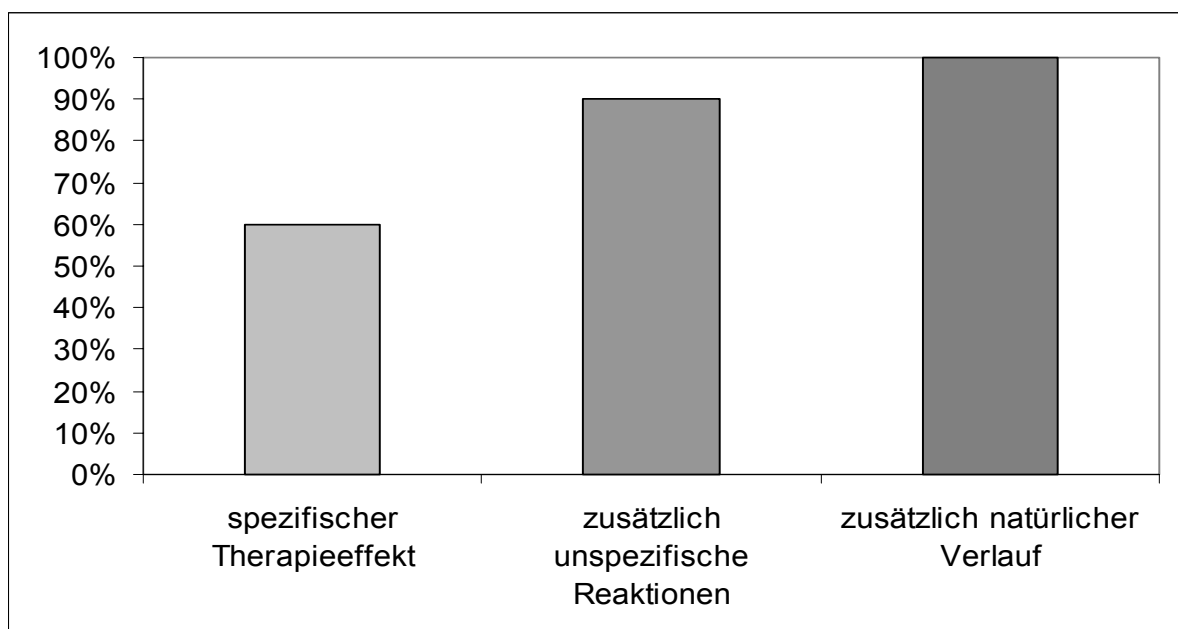


Abbildung 7: Globaler Therapieeffekt und Einzelkomponente in Anlehnung an [4]

Möglicherweise sind unspezifische und spezifische Effekte nicht substitutional zueinander. Sie könnten sich gegenseitig verstärken, so dass das Gesamtergebnis die Summe der Einzelwirkungen übertreffen kann. Vielleicht existiert ein „optimales Verhältnis“ beider maßgebend für eine Maximierung des Gesamterfolgs.

Schwierig stellt sich eine konkrete Quantifizierung des Ausmaßes der unspezifischen Reaktionen dar. Dennoch ist eine detaillierte Zuordnung der Wirkungskomponenten zur Erklärung des globalen Therapieeffektes von wissenschaftlichem Interesse nicht zuletzt auch vor dem Hintergrund gesundheitsökonomischer Aspekte. Selbst wenn, und dies liegt klinischen Studien nach erfolgreichen präklinischen Untersuchungen zugrunde, die physikalisch-chemische Intervention mit dem Therapieerfolg assoziiert ist, stellt sie möglicherweise nicht die einzige direkte Ursache dar. Eine Abschätzung der Einwirkung unspezifischer, nicht näher differenzierter Therapiewirkungen könnte ansatzweise als Residualrechnung gelöst werden. Der Placeboeffekt kann vom übrigen Wirksamkeitseffekt dadurch abgegrenzt werden, dass eine Placebo-Behandlung einerseits einer aktiven Therapie, andererseits aber auch einer Nulltherapiegruppe gegenübergestellt wird. Die Erforschung von Placebo-Effekten ist wichtig, da jede wirksame Behandlung auch eine Placebokomponente hat.

Ein messbarer Behandlungserfolg kann beispielsweise schon alleine deshalb eintreten, dass die Patienten im Rahmen von randomisierten klinischen Studien ein wesentlich höheres Ausmaß an Zuwendung durch häufige Untersuchungen erhalten als in der üblichen Versorgungspraxis. Darüber hinaus kann über selektive Ein- und Ausschlusskriterien ein spezielles Kollektiv erzeugt werden, bei welchem eine überdurchschnittliche Wirksamkeit gemessen wird. Theoretisch wäre vorstellbar, dass eine Therapie auf rein pharmakologische Weise wirkt (erster Balken in Abbildung 8 auf der folgenden Seite). Möglicherweise entfaltet sich eine Therapie aber schwerpunktmäßig auf nicht-pharmakologische Weise (letzter Balken). Es ist zu vermuten, dass mindestens ein Teil der beobachteten Wirkung über nicht-pharmakologische Faktoren zu erklären ist (mittlere Darstellung).

Es lässt sich argumentieren, dass es für den Patienten im Hinblick auf das Hauptziel „Gesamtüberlebenszeit“ unerheblich ist, wie sich eine Zielerfüllung realisieren lässt. In ökonomischer Hinsicht dagegen ist es im Zusammenhang mit Fragen bezüglich einer solidarischen Finanzierung von Gesundheitsleistungen und einer effizienten Allokation der aufgrund von Kosten limitierten Ressourcen entscheidend, den Beitrag einzelner Komponenten zum Zustandekommen des Gesamtergebnisses zu identifizieren.

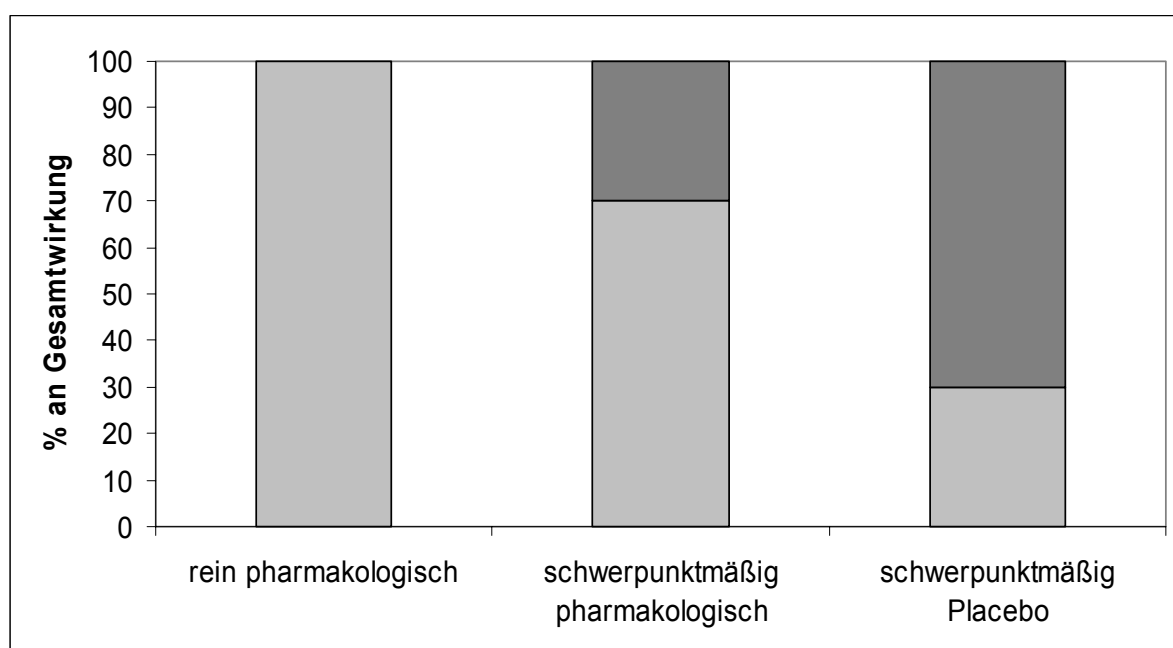


Abbildung 8: Therapeutische Gesamtwirkung in Anlehnung an [4]

5.5 Interaktion zwischen Arzt und Patient: Kommunikationstheoretische Aspekte

Unspezifische Reaktionen können durch die kommunikative Beziehung zwischen Arzt und Patient konditioniert werden. Diese Beziehung schließt jede verbale und nonverbale Kommunikation zwischen Arzt und Patient ein sowie die Bedingungen, unter denen diese stattfindet. In diesem Zusammenhang ist der Begriff des Knowledge Framing geprägt worden [45].

Eine zwischen Arzt und Patient bestehende emotionale Interaktion, eingeleitet über vorangegangene persönliche Erfahrungen und/oder Empfehlungen, ist die Grundlage für eine erfolgreiche Kommunikation. Einen zentralen Stellenwert nimmt dabei das Aufklärungsgespräch ein. Zunächst hat jedes Aufklärungsgespräch einen rechtlichen Aspekt [7]. Artikel 1 und 2 des Grundgesetzes garantieren Selbstbestimmungsrecht und körperliche Integrität. Daneben kommt diesem jedoch auch eine anxiolytische Funktion zu. Den Informationsgehalt selbst bilden inhaltliche Schwerpunkte prozeduraler und sensorischer Art: zeitlicher und organisatorischer Ablauf des Eingriffs, spezifische Details wie Behandlungsraum, Behandlungsteam, Art der verwendeten Instrumente oder Positionierung während des Eingriffs und Aspekte der vom Patienten erwarteten Mitarbeit.

Vor diesem Hintergrund zeigt sich die Bedeutung des Knowledge Framing in der Rolle des Prüfarztes bei klinischen Studien insbesondere dann, wenn ausschließlich der Patient im Hinblick auf die Art der Therapie verblindet ist. Sinnvollerweise sollte hierbei die Art der Wahrnehmung durch den Patienten nicht nur bezüglich Aussehen, Geschmack und Form vergleichbar sein. Die Therapiemaßnahmen sollten auch in der gleichen Art und Weise dargeboten werden, so dass die mit einem bestimmten Medikament verbundenen, beim Patienten hervorgerufenen kognitiven Assoziationen in den zu vergleichenden Gruppen ähnlich erzeugt werden. Asymmetrien in der Patienteninformation aufgrund von Präferenzen und Überzeugungen des behandelnden Arztes bewirken möglicherweise eine Verzerrung des gemessenen Effekts.

5.6 Art der Verblindung

Man kann davon ausgehen, dass über die physikalisch-chemische Wirkung hinaus ein beobachteter Gesamteffekt durch psychosoziale Aspekte wie Arzt-Patienten-Beziehung oder Erfahrungsaustausch mit anderen Patienten konditioniert wird. Dies hat methodische Implikationen in Bezug auf die Durchführung kontrollierter klinischer Studien, denn die Art der Verblindung kann die Ergebnismessung beeinflussen. Sofern ein bestimmtes Behandlungsschema durch den Prüfarzt bevorzugt wird, kann dieser dem Patienten entsprechend seiner Gruppenzugehörigkeit unterschiedliche Prognosewerte und

therapeutische Vorteile beziehungsweise Nachteile kommunizieren. Die dem Patienten zur Verfügung gestellte ärztliche Information kann dadurch mit positiven oder negativen Assoziationen verbunden sein und auf diese Weise eine Auswirkung auf dessen emotionale Haltung und bedingt auch auf den Therapieerfolg haben.

Speziell eine Nulltherapie-Kontrollgruppe als reine Beobachtung kann dadurch eine ausgeprägte kognitive Dissonanz beim behandelnden Arzt und/oder Patienten hervorrufen [45]. Hierbei wird das Prinzip der Behandlungsgleichheit eventuell stark verletzt, geschätzte Therapieeffekte möglicherweise nicht realitätskonform abgebildet. Ein doppelblindes Design hingegen in Einheit mit standardisierter Patientenaufklärung würde derartigen informationellen Asymmetrien entgegenwirken.

In den betrachteten Studien kann nicht von einem doppelblinden Design ausgegangen werden. Bei einem Vergleich mit einer reinen Beobachtungsgruppe sind die gemessenen Effekte möglicherweise überzeichnet. Es ist nicht davon auszugehen, dass ein in derartigen Studien beobachteter Überlebensvorteil ausschließlich mit einer pharmakologischen Spezifität, sondern auch mit der Information des Patienten über die durchgeführte Therapie korreliert ist. Die beobachtete Wirkung könnte rechnerisch korrigiert werden, um eine Abschätzung des Beitrags der stofflichen Komponente durchzuführen. Ein solches Vorgehen würde doppelblinde Studien mit hoher Fallzahl erfordern. Eine doppelblinde Studie kommt ohnehin nur bei einem placebokontrollierten Experiment beziehungsweise bei einem Vergleich mit einer alternativen Therapie in Frage. In diesen Fällen stellt die Doppelblindheit jedoch ein unerlässliches Instrument für valide Ergebnisse dar.

Aber selbst bei einem doppelblinden Vergleich zweier alternativer Behandlungsschemata kann vermutet werden, dass mindestens die Prüfarzte, wenn nicht auch ein Teil der Patienten aufgrund beispielsweise leicht zu unterscheidender Applikationswege die Art der verabreichten Therapie erraten können. Eine Therapie offenbart sich möglicherweise schon durch die Unterschiede im Applikationsmuster. Damit sind aber die Grundprinzipien klinischer Studien zum Wirksamkeitsnachweis einer therapeutischen Maßnahme verletzt.

Im Hinblick auf die praktische Durchführung einer klinischen Studie stellt die Doppelblindheit insbesondere im Zusammenhang mit einem Placebo-Vergleich ein ethisch kontrovers diskutiertes Thema und aufgrund persönlicher Überzeugungen in der Forschergemeinschaft möglicherweise schwer zu implementierendes Design dar.

5.7 Randomisierung

Eine Randomisierung gewährleistet im allgemeinen eine unverzerrte Zuteilung von Patienten zu den Behandlungsgruppen. Im Falle der einfachen Randomisierung basierend auf einer einzigen Folge von Zufallszahlen, können Unausgewogenheiten hinsichtlich einzelner demographischer und prognostischer Faktoren entstehen. Solche Effekte lassen sich durch eine stratifizierte, blockweise Randomisierung minimieren. Hierbei werden innerhalb der Vergleichsgruppen für bestimmte (nach Zentren sowie wichtigen demografischen und anamnestischen Faktoren getrennte) Subpopulationen separate Zufallszahlen erzeugt. Ein blockweises Randomisierungsverfahren garantiert darüber hinaus, dass im Rahmen einer festgelegten Länge die Vergleichsgruppen zahlenmäßig jeweils ausgeglichen sind [16]. Grundsätzlich sollte die Randomisierung von einem unabhängigen Institut durchgeführt werden [8].

Die Sichtung des verfügbaren Studienmaterials zeigte jedoch gerade hinsichtlich dieser Punkte teilweise Informationslücken. So wird bei einigen der in dieser Arbeit betrachteten Publikationen in den Ausführungen zu Material und Methodik auf diese Aspekte nicht detailliert eingegangen. In den meisten Fällen wird eine Randomisierung – abgesehen von Hinweisen auf eine Stratifizierung – nur erwähnt. Eine genauere Spezifikation der zur Generierung von Zufallszahlen verwendeten Methode und des Zuteilungsmechanismus wird nicht angegeben. Dies wäre jedoch im Hinblick auf die Validität der Randomisierung wichtig. Derartige Schlussfolgerungen ergeben sich im Zusammenhang mit einer Übersichtsarbeit aus dem Jahre 1990. Hier wurden 80 Berichte über randomisierte klinische Parallelgruppenstudien aus vier führenden Fachzeitschriften überprüft [21].

In 60% der Veröffentlichungen existierten keine Angaben zum Randomisierungsverfahren, in 30% der Studien war nicht eindeutig belegt, dass eine randomisierte Zuteilung zu den Behandlungsgruppen erfolgt war. Als weitere Problemschwerpunkte wurde unzureichender Vergleich der Ausgangscharakteristika (41%) sowie fehlende statistische Power-Berechnung (61%) identifiziert.

5.8 Multiple Kontrollgruppen

Bei einem Vergleich mit einer Beobachtungskontrollgruppe ist nicht eindeutig, auf welche pharmakologischen oder psychologischen Teilkomponenten der Behandlung die beobachtete Wirksamkeit zurückzuführen ist. Insbesondere kann die gemessene Wirksamkeit dadurch überzeichnet sein, dass die Patienten in der Experimentalgruppe eine hohe Adhärenz bezüglich der festgelegten Therapieziele aufzeigen. In einer neueren Meta-Analyse [53] wurde dieser Sachverhalt unter Verwendung von 21 Einzelstudien, in denen auch Placebo-Vergleiche eingeschlossen waren, untersucht.

Die Autoren deuten die Daten dahingehend, dass ein substanzieller positiver Zusammenhang zwischen Therapiecompliance und beobachteter Wirksamkeit besteht. Möglicherweise auch dadurch bedingt zeigen sich in Überlebenszeitstudien im Vergleich zur Nulltherapie über einen genuinen pharmakologischen Mechanismus hinaus teilweise erhebliche Therapievorteile. Es wäre deshalb zweckmäßig, einen neuen Therapieansatz simultan mit jeweils einer Standardtherapie- und Beobachtungskontrollgruppe zu vergleichen. Die spezifische Wirksamkeit einer definierten Behandlungsstrategie könnte sich dann aufzeigen lassen, indem anhand (für den Patienten) relevanter Zielgrößen ein quantifizierbarer zusätzlicher therapeutischer Nutzen im Vergleich zu beiden Kontrollgruppen beobachtet wird. Zusätzlich könnte die Kontrollgruppe auch als Placebo-Behandlung entworfen werden. Damit könnte der beobachtete Gesamteffekt als Residualrechnung näher aufgeschlüsselt werden.

6. Folgerungen und Empfehlungen

6.1 Assay Sensitivity und Kontrollgruppenwahl

Die Überprüfung neuentwickelter Therapieansätze erfordert doppelblinde randomisierte konfirmatorische klinische Studien. Schwerpunkt von Überlegenheitsstudien ist die Untersuchung der spezifischen Wirksamkeit für eine definierte medizinische Intervention. Es wird empfohlen, Studien so zu strukturieren, dass ein Höchstmaß an „Assay Sensitivity“ gegeben ist. Als „Assay Sensitivity“ wird die Fähigkeit einer Studie bezeichnet, eine wirksame Therapie gegen eine weniger wirksame und unwirksame Behandlungsvariante abzugrenzen. Im Rahmen von Überlegenheitsstudien wird in einschlägigen Richtlinien hierzu empfohlen, eine Studie mit multiplen Kontrollgruppen anzusetzen [8]. Die einfachste Variante stellt eine Studie mit drei Behandlungsarmen dar: Novum versus Standardtherapie versus Placebo. Hierbei wird die Effektgröße im Vergleich der Experimentaltherapie gegen das Placebo bei gleichzeitiger Gegenüberstellung zu der aktiven Kontrollgruppe abgeschätzt. Assay Sensitivity wird durch den Vergleich der aktiven Kontrollgruppe zum Placebo gewährleistet.

6.2 Add-on Studien

Insbesondere in onkologischen Studien kann eine Verwendung von Placebos unter ethischen Gesichtspunkten Bedenken hervorrufen, was Implikationen für Akzeptanz und Durchführbarkeit hat. Vor diesem Hintergrund sind Add-On-Studien interessant: Alle Patienten erhalten eine Standardtherapie. Die Behandlungsgruppen unterscheiden sich dann, indem der Standardtherapie jeweils das Novum beziehungsweise das Placebo modular hinzugefügt wird. Add-On-Studien bieten sich an, wenn für die Standardtherapie bereits ein ausreichender Wirksamkeitsnachweis vorliegt.

6.3 Behandlungsgleichheit, Randomisierung und Verblindung

Das für einen Parallelgruppenvergleich erforderliche Kriterium der Behandlungsgleichheit könnte durch eine Standardisierung der Patienten-Einwilligungserklärung besser implementiert werden. Weitere zwingende Kriterien im Hinblick auf einen validen Gruppenvergleich sind Randomisierung und Verblindung. Bei fehlender Information bezüglich eines dieser Aspekte ist es ohne Rücksprache mit den Autoren für den Leser nicht direkt erkennbar, ob dieser Punkt gar nicht berücksichtigt wurde oder ob er nur nicht erwähnt wurde. Eine identische Fallzahl in den zu vergleichenden Behandlungsgruppen deutet auf eine effiziente Randomisierung hin, könnte jedoch im ungünstigsten Fall auch über korrigierende Zuteilung zusätzlicher Patienten zu den Gruppen erzeugt werden. Ein stärkerer Schwerpunkt bei der Durchführung zukünftiger klinischer Studien kann auf der Verankerung dieser Prinzipien und deren Überprüfbarkeit gesetzt werden.

6.4 Berichtswesen

Das Berichtswesen sollte möglichst transparent und vollständig sein. Nur auf diese Weise können den verschiedenen Zielgruppen (Forschergemeinschaft, Behörden, Industrie, Patienten) Ergebnisse glaubhaft kommuniziert werden. Es könnte eine Reihe inhaltlicher Kriterien aufgestellt werden, welche auch im Abstrakt zwingend beschrieben werden müssen. Dazu wird vorgeschlagen, Autoren bei Einreichung einer Veröffentlichung zum Ausfüllen einer Checkliste zu verpflichten. Für jeden Punkt insbesondere im Hinblick auf Studiendesign und statistische Methodik sollen Seite und Absatz, wo sich die entsprechende Information findet, angegeben werden [21]. Die Diskussion der Forschungsergebnisse könnte mit den Bedingungen in der Praxis abgestimmt werden, um effektive Handlungsempfehlungen zu formulieren. In diesem Zusammenhang könnte in jeder Veröffentlichung auch ein Ausblick auf die in der täglichen Anwendung zu erwartenden Wirkungen und Schwierigkeiten enthalten sein, um die punktuellen Ergebnisse der konkreten Studie praxisorientiert zu gewichten.

6.5 Leitlinien

Es erscheint empfehlenswert, die in einschlägigen Leitlinien beschriebenen Verfahrensstandards für jede klinische Studie, gleich ob in Industrieunternehmen, universitären Einrichtungen oder unabhängigen Instituten durchgeführt, als verbindlich zu implementieren. Besonders hervorhebenswert sind in diesem Zusammenhang die ICH Guidelines, Sektionen E3: Berichtswesen klinischer Studien, E9: Statistische Prinzipien klinischer Studien und E10: Kontrollgruppen in klinischen Studien [8].

7. Zusammenfassung und Schlussbemerkungen

Basierend auf den Substanzen 5-Fluororacil und Leucovorin wird ein Therapieansatz geschaffen, welcher im Rahmen der adjuvanten Behandlung beim kolorektalen Karzinom standardmäßig eingesetzt wird. Es existieren verschiedene Ausgestaltungen je nach Applikationsart, Dosisstärke und Anwendungsdauer.

In der vorliegenden Arbeit interessiert die Spezifität dieser Behandlung, exemplarisch untersucht anhand eines in Leitlinien empfohlenen Schemas. Möglicherweise ist ein in klinischen Studien beobachteter Wirkungseffekt nicht vollständig durch einen pharmakologischen Zusammenhang zu erklären. Weitere Wirkungsmechanismen – über psychosoziale Faktoren gesteuert - könnten das Gesamtergebnis beeinflussen.

Zu diesem Zweck wurde eine Medline©-Recherche über die relevanten Publikationen zu diesem Thema durchgeführt. Anhand der Ergebnisse randomisierter klinischer Studien lässt sich für das definierte Chemotherapieprotokoll eine spezifische Wirksamkeit nicht aufzeigen. Bei einem Vergleich mit einer Kontrollgruppe ohne Behandlung werden signifikante Effekte berichtet, die sich jedoch im Vergleich mit einer alternativen aktiven Therapie nicht bestätigen.

Es wird gefolgert, dass die Effektivität einer therapeutischen Maßnahme nicht ausschließlich über den Vergleich mit einer Nulltherapie Kontrollgruppe evaluiert werden sollte, sondern simultan auch über den direkten Vergleich mit alternativen Behandlungsschemata. Eine stärkere Verankerung und Überprüfbarkeit der grundlegenden Prinzipien von Randomisierung und Doppelblindheit könnte die Validität von Forschungsergebnissen erhöhen.

Die Erfolgsintensität einer medizinischen Behandlung wird möglicherweise über die reine pharmakologische Wirkung hinaus insbesondere auch über die Erwartungshaltung seitens des Patienten gesteuert.

In diesem Zusammenhang erscheinen weitere placebokontrollierte Studien notwendig, um den physikalisch-chemischen Wirkungsmechanismus adjuvanter Chemotherapie präziser abzugrenzen. Möglicherweise kann eine Quantifizierung der Bedeutung von Kontextfaktoren vorgenommen werden. Nicht zuletzt muss vor dem Hintergrund der Diskussion um die Finanzierung von Gesundheitsleistungen eine postulierte pharmakologische Wirksamkeit auf gesicherten klinisch-statistischen Erkenntnissen gestützt werden. Daran kann sich das Leistungsentgelt definierter medizinischer Maßnahmen ökonomisch effizient orientieren.

Zur statistischen Absicherung von Wirksamkeitseffekten muss der Begriff der Kausalität in klinischen Studien stärker implementiert werden. Gegeben einer biologischen Plausibilität, welche man im Rahmen der betrachteten Behandlung aufgrund mikrobiologischer Untersuchungen voraussetzen darf, ist zunächst die Festlegung einer klinisch relevanten Höhe zur Wirksamkeit notwendig. Dies wird im Rahmen von Fallzahlschätzungen teilweise auch durchgeführt, jedoch fehlt es hier noch an einem breiten Konsens sowie einer verpflichtenden methodischen Integration. Die gefundenen Ergebnisse zeigen eine nur unzureichende Spezifität im Hinblick auf ein Therapieprotokoll, denn über jede Form einer aktiven Therapie wird eine höhere Wirksamkeit beobachtet als durch eine Nulltherapie. Eine profunde Dosis-Wirkungsbeziehung kann zumindest für die betrachtete 5-FU / LV Behandlung aufgrund von Studienergebnissen ebenfalls nicht abgeleitet werden.

Für den medizinischen Alltag könnte basierend auf den Erkenntnissen zur spezifischen Wirkungsweise medizinisch-physikalischer Interventionen ein ganzheitliches Therapiekonzept entwickelt werden, welches den Patienten in seiner bio-psychozialen Gesamtheit berücksichtigt, dementsprechend alle Möglichkeiten ausschöpft und zielorientiert eine integrierte Behandlung ansetzt.

Wissensgenerierung kann optimiert werden, wenn ein interdisziplinärer Ansatz aus der Sicht aller an der Forschung beteiligten Akteure angesetzt wird. Auf diese Weise könnte eine interdisziplinäre Zusammenführung klinischer, statistisch-mathematischer und ökonomischer Forschungsmethoden in einer konsensfähigen übergreifenden

Expertenmeinung münden. Politische Entscheidungsträger und private Investoren könnten diesen Prozess über die Einrichtung eines Innovationsfonds fördern. Für die Solidargemeinschaft würde daraus über eine effiziente Budgetverwendung ein Kostenvorteil, für die Patienten ein höheres Maß an Sicherheit bei der Entwicklung standardisierter Therapieverfahren erwachsen.

Vor über 50 Jahren begann die klinische Forschung, spezielle therapeutische Maßnahmen zur (neo-) adjuvanten Therapie des kolorektalen Karzinoms zu entwickeln. Spätestens mit der Veröffentlichung erster signifikanter Ergebnisse in größeren randomisierten Studien Mitte der 80er Jahre, in welchen die adjuvante Chemotherapie jeweils einer Beobachtungskontrollgruppe gegenübergestellt wurde, hegen Ärzte und Patienten die Überzeugung auf eine nachhaltige Heilbarkeit des kolorektalen Karzinoms durch Operation und Chemotherapie. Nach fast drei Jahrzehnten Anwendung chemotherapeutischer Standards bleibt nach neueren Untersuchungen zu diesem Thema die klinische Relevanz fraglich. Die Eckpfeiler der Therapie haben sich seit 30 Jahren nicht verändert. Vielmehr wurde über modulare Erweiterung durch zusätzliche pharmakologische Komponenten kontinuierlich nach einer Verbesserung des Gesamtergebnisses geforscht. Dabei ist es gelungen, verschiedene Detailspekte insbesondere im Hinblick auf die Nebenwirkungen teilweise deutlich zu verbessern. Heute lässt sich die Chemotherapie wesentlich günstiger steuern als in den Anfängen. Eine sprunghafte und relevante Entwicklung zur entscheidenden Lebenszeitverlängerung ist jedoch ausgeblieben.

Mit den Ergebnissen der vorliegenden Literaturrecherche lassen sich die Standards, welche aktuell in den Leitlinien festgelegt sind, nicht vorbehaltlos stützen. Folglich verfährt die medizinische Praxis möglicherweise nicht vollständig kongruent zur epidemiologischen Evidenz. Patienten schöpfen ihre Erwartungen und Hoffnungen aus den vom behandelnden Arzt vermittelten Perspektiven zu empfohlenen Therapiekonzepten. Hierbei existiert möglicherweise ein Mangel an für Laien verständlich aufbereiteter Information bezüglich der epidemiologisch nachweisbaren Wirksamkeit und Toxizität. Es wäre denkbar, die Meinung informierter Patienten als Laien in Zukunft stärker bei der Gestaltung von Behandlungsstandards einzubinden.

Aktuell besteht seitens verschiedener am Gesundheitswesen beteiligter Akteure (Patienten, Ärzte, pharmazeutische Industrie, Gesundheitspolitiker) explizit der Wunsch nach adjuvanten Therapiekonzepten, wenn dadurch ein (für die Erzeugerindustrie) monetär messbarer oder auch ein (insbesondere für Ärzte und Patienten) subjektiv empfundener Mehrwert erzeugt wird. Beim behandelnden Arzt und Patienten ist dieser Mehrwert mit erheblichen positiven psychologischen Assoziationen verknüpft.

Vom ordnungspolitischen Standpunkt betrachtet kann ein auf dem Solidarprinzip basierendes Gesundheitssystem langfristig jedoch nur diejenigen Leistungen finanzieren, die epidemiologisch nachweislich klinisch sinnvoll und nützlich sind. Rationale Entscheidungsträger werden Art und Umfang solidarischer Gesundheitsleistungen auf epidemiologischen Fakten definieren müssen. Dadurch ergibt sich ein Dilemma, in welchem aufgrund der Überlagerung verschiedener subjektiver Interessen und Wunschvorstellungen ein System perpetuiert wird, welches sich nicht vollständig an empirischen Fakten orientiert, sondern an der Implementierung einer gefühlten Sicherheit. Diese Essenz lässt sich treffend formulieren mit den Worten eines evangelischen Pfarrers: *„Du sollst niemandem seine Überzeugung nehmen, es sei denn Du kannst etwas Besseres bieten“*.

8. Quellenverzeichnis

8.1 Verzeichnis der im Text verwendeten Lehrbücher, Leitlinien und Datenbanken

(1) AWMF Online – Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Interdisziplinäre Leitlinie der Deutschen Krebsgesellschaft und ihrer Arbeitsgemeinschaften, der Deutschen Gesellschaft für Chirurgie und der Deutschen Gesellschaft für Verdauungs- und Stoffwechselkrankheiten.

Leitlinien-Register 032/011: Kolonkarzinom

(2) AWMF Online – Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Interdisziplinäre Leitlinie der Deutschen Krebsgesellschaft und ihrer Arbeitsgemeinschaften, der Deutschen Gesellschaft für Chirurgie und der Deutschen Gesellschaft für Verdauungs- und Stoffwechselkrankheiten.

Leitlinien-Register 032/012: Rektumkarzinom

(3) AWMF Online – Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Interdisziplinäre Leitlinie der Deutschen Krebsgesellschaft und ihrer Arbeitsgemeinschaften, der Deutschen Gesellschaft für Chirurgie und der Deutschen Gesellschaft für Verdauungs- und Stoffwechselkrankheiten.

Leitlinien-Register 021/007: Kolorektales Karzinom

(4) Fletcher RH, Fletcher SW, Wagner EH: Klinische Epidemiologie: Grundlagen und Anwendungen, deutschsprach. Ausgabe: Haerting J, Rink Ch, Ullstein Medical: Wiesbaden, Kapitel 7, S. 189-230 (1999)

(5) Glenn AW: Common Statistical Methods for Clinical Research with SAS Examples, second edition, SAS Institute Inc. Cary, NC, USA, Kapitel 9, S. 77-156, Kapitel 22, S.365-388 (2002)

(6) Globocan 2002 database project <http://www-dep.iarc.fr/>

- (7) Grundgesetz der Bundesrepublik Deutschland, Artikel 1 und 2
- (8) ICH: International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use. ICH Steering Committee. 1998. E9: Statistical Principles For Clinical Trials. E10: Choice of Control Group and Related Issues In Clinical Trials.
- (9) Michl M: Hämatologie, 1. Auflage, Urban & Fischer: München, S. 114-117 (2005)
- (10) NIH – National Institute of Health: Adjuvant Therapy for Patients with Colon and Rectum Cancer. NIH – Consensus Statement 1990 Apr 16-18; 8(4): 1-25
- (11) Porzsolt F, Kumpf J, Pöppel E: Evidence-based Medicine und “Knowledge-Framing” in Porzsolt F, Williams AR, Kaplan RM: Klinische Ökonomik. Effektivität & Effizienz von Gesundheitsleistungen. Ecomed, S. 185-189 (2003)
- (12) PubMed, National Library of Medicine and National Center for Biotechnology Information, USA: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi> (2005).
- (13) Robert-Koch Institut: Dachdokumentation Krebs.
http://www.rki.de/cln_006/nn_227180/DE/Content/GBE/DachdokKrebs/Datenbankabfragen/datenbankabfragen__node.html__nnn=true
- (14) Sachs L, Hedderich J: Angewandte Statistik, Methodensammlung mit R, 12. Auflage, Springer: Berlin, Heidelberg, New York, Kapitel 7, S. 305-414, Kapitel 8, S. 560-597 (2006)
- (15) SAS/STAT User’s Guide – SAS Institute Inc. 1999, Cary, NC, USA
- (16) Schumacher M, Schulgen G: Methodik klinischer Studien: methodische Grundlagen der Planung, Durchführung und Auswertung, Springer: Berlin, Heidelberg, New York, Kapitel 5, S. 71-87, Kapitel 6, S. 89-105 (2002)

(17) Spiegel: Giftkur ohne Nutzen 41/2004 – 04.Oktober 2004.
http://www.pilhar.com/News/Presse/2004/20041004_Spiegel_Giftkur_Chemo.htm

(18) Tumorzentrum Freiburg: Empfehlungen zur standardisierten Diagnostik, Therapie und Nachsorge: Kolorektales Karzinom. Stand: Februar 2000. http://www.tumorzentrum-freiburg.de/medizin_info/kolorektal_karzinom.htm

(19) Tumorregister München: Jahresbericht 1998 / Schwerpunkt kolorektale Karzinome.
<http://trm.web.med.uni-muenchen.de/annals.html>

(20) Weiß Ch.: Basiswissen Medizinische Statistik, 3. Auflage, Springer: Berlin, Heidelberg, New York, Teil III, Kapitel 9-12, S. 171-243 (2005)

8.2 Verzeichnis der im Text verwendeten Publikationen

(21) Altman DG, Doré CJ: Randomisierung und Vergleich der Ausgangswerte in klinischen Studien. *Zeitung für ärztliche Fortbildung und Qualität im Gesundheitswesen* 99: 629-634 (2005)

(22) André T, Boni C, Mounedji-Boudiaf L, Navarro M, Taberero J, Hickish T, Topham C, Zaninelli M, Clingan P, Bridgewater J, Tabah-Fisch I, de Gramont A: Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *New England Journal of Medicine* 350: 2343-51 (2004)

(23) Benson III Al B, Schrag D, Somerfield MR, Cohen AM, Figueredo A T, Flynn PJ, Krzyzanowska MK, Maroun J, McAllister P, Van Cutsem E, Brouwers M, Charette M, Haller DG: American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. *Journal of Clinical Oncology* 22: 3408-19 (2004)

(24) Bradford Hill A: The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* 58: 295-300 (1965)

- (25) Carmichael J, Popiela T, Radstone D, Falk S, Borner M, Oza A, Skovsgaard T, Munier S, Martin C: Randomized comparative study of Tegafur/Uracil and oral Leucovorin versus parenteral fluorouracil and Leucovorin in patients with previously untreated metastatic colorectal cancer. *Journal of Clinical Oncology* 20: 3617-3627 (2002)
- (26) Chau I, Norman A R, Cunningham D, Tait D, Ross PJ, Iveson T, Hill M, Hickish T, Lofts F, Jodrell D, Webb A, Oates JR: A randomized comparison between 6 months of bolus fluorouracil/leucovorin and 12 weeks of protracted venous infusion fluorouracil as adjuvant treatment in colorectal cancer. *Annals of Oncology* 16: 549-57 (2005)
- (27) Dencausse Y, Hartung G, Sturm J, Kopp-Schneider A, Hagmuller E, Wojatschek C, Lindemann H, Fritze D, Queisser W: Adjuvant chemotherapy in stage III colon cancer with 5-fluorouracil and levamisole versus 5-fluorouracil and leucovorin. *Onkologie*. 25: 426-430 (2002)
- (28) Focan C., Bury J., Beauduin M. et al.: Adjuvant intraportal chemotherapy for Dukes' B2 and C colorectal cancer also receiving systemic treatment: results of a multicenter randomized trial. *Anticancer Drugs*. 11: 549-554 (2000)
- (29) Gennatas C, Mouratidou D, Androulakis G, Georgoulas V, Tsavaris N, Phillippakis M, Michailakis E, Kalofonos C, Mpsempeas S, Katsos J, Tsitoura M, Retalis G, Mallas E, Vosrso D, Andreadis C, Hatzistylianos G, Pisiotis C, Kamilarios D, Kakoliris S, Komporozos V, Kannas D, Legakis N, Mpatakis T, Ntamtsios J, Papaevangelou E, Perios G, Photopoulos A, Pouli A, Prigouris S, Samanidis L, Sakellariou V, Smyrniotis V, Polymeneas G, Vasiliou J, Athanasiou A, Papadimitriou J: Adjuvant systematic therapy protocol for Dukes' B2 and C resectable colon carcinoma. *Tumori* 88: 32-36 (2002)
- (30) Gill S, Sargent D: End points for adjuvant therapy trials: Has the time come to accept disease-free survival as a surrogate end point for overall survival? *The Oncologist* 11: 624-629 (2006)

- (31) Gill S, Loprinzi CL, Sargent DJ, Thomé SD, Alberts SR, Haller DG, Benedetti J, Francini G, Shepherd LE, Seitz F, Labianca R, Chen W, Cha SS, Heldebrant MP, Goldberg RM: Pooled analysis of fluorouracil-based adjuvant chemotherapy for stage II and III colon cancer: Who benefits and by how much? *Journal of Clinical Oncology* 22: 1797-1906 (2004)
- (32) Grothey A, Kleeberg UR, Stauch M, Hieke K: Behandlungskosten Fluoropyrimidin-basierter Therapien des kolorektalen Karzinoms in verschiedenen Behandlungssituationen. *Gesundheitsökonomie und Qualitätsmanagement* 9: 229-235 (2004)
- (33) Haydon A: Adjuvant chemotherapy in colon cancer: What is the evidence? *Internal Medicine Journal* 33: 119-124 (2003)
- (34) IMPACT Investigators: Efficacy of adjuvant fluorouracil and folinic acid in colon cancer. International Multicenter Pooled Analysis of Colon Cancer Trials (IMPACT) investigators. *Lancet* 345: 939-944 (1995)
- (35) Kajnar H., Awa A., Fässler M., Porzsolt F.: Structural Analysis of Health Technology Assessments (HTA) Demonstrates the Need for Standardized Assessment of Transparency, Applicability and Consistency (TAC Criteria), Abstract (2004)
- (36) Laurie JA, Moertel CG, Fleming TR, Wieand HS, Leigh JE, Rubin J, Mc Cormack GW, Gerstner JB, Krook JE, Malliard J, Twito DI, Morton RF, Tschetter LK, Barlow JF: Surgical adjuvant therapy of large-bowel carcinoma: an evaluation of levamisole and fluorouracil. The North Central Cancer Treatment Group and the Mayo Clinic. *Journal of Clinical Oncology* 7: 1447-56 (1989)
- (37) Laurie JA, Moertel CG, Fleming TR, McDonald JS, Haller DG, Laurie JA, Tangen CM, Ungerleider JS, Merson WA, Tormey DC, Glick JH: Fluorouracil plus levamisole as effective adjuvant therapy after resection of stage III colon carcinoma: a final report. *Annals of Internal Medicine* 122: 321-326 (1995)

- (38) Macdonald J.S., Astrow A.B.: Adjuvant therapy of colon cancer. *Seminars in Oncology*. 28: 30-40 (2001)
- (39) Marsoni S.: Efficacy of adjuvant fluorouracil and leucovorin in stage B2 and C colon cancer. *IMPACT Seminars in Oncology* 28: 14-19 (2001)
- (40) Min JS, Kim NK, Park JK, Yun SH, Noh JK: A prospective randomized trial comparing intravenous 5- fluorouracil and oral doxifluridine as postoperative adjuvant treatment for advanced rectal cancer. *Annals of Surgical Oncology* 7: 674-9 (2000)
- (41) Moertel CG., Fleming TR., Macdonald JS, Haller DG, Laurie JA, Goodman PJ, Ungerleider JS, Emerson WA, Tormev DC, Glick JH: Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. *The New England Journal of Medicine* 320: 352-358 (1990)
- (42) O'Connell MJ, Laurie JA, Kahn M, Fitzgibbons RJ Jr, Erlichman C, Shepherd L, Moertel CG, Kocha WI, Pazdur R, Wieand HS, Rubin J, Vukov AM, Donohue JH, Krook JE, Figueredo A: Prospectively randomized trial of postoperative adjuvant chemotherapy in patients with high-risk colon cancer. *Journal of Clinical Oncology* 16: 295-300 (1998)
- (43) O'Connell MJ, Laurie JA, Kahn M, Macdonald JS, Haller DG, Mayer RJ, Wieand HS: Controlled trial of fluorouracil and low-dose leucovorin given for 6 months as postoperative adjuvant therapy for colon cancer. *Journal of Clinical Oncology* 15: 246-25 (1997)
- (44) Poplin EA, Benedetti JK, Estes NC, Haller DG, Mayer RJ, Goldberg M., Weiss GR, Rivkin SE, Macdonald JS: Phase III Southwest Oncology Group 9415/Intergroup 0153 randomized trial of fluorouracil, leucovorin, and levamisole versus fluorouracil continuous infusion and levamisole for adjuvant treatment of stage III and high-risk stage II colon cancer. *Journal of Clinical Oncology* 23: 1819-25 (2005)

- (45) Porzsolt F., Schlotz-Gorton N., Biller-Andorno N. et al.: Applying Evidence to Support Ethical Decisions: Is the Placebo Really Powerless?, *Science and Engineering Ethics* 10: 119-132 (2004)
- (46) QUASAR collaborative group: Comparison of fluorouracil with additional levamisole, higher-dose folinic acid, or both, as adjuvant chemotherapy for colorectal cancer: a randomised trial. *Lancet* 355: 2588-96 (2000)
- (47) Saini A, Norman AR, Cunningham D, Chau I, Hill M, Tait D, Hickish T, Iveson T, Lofts F, Jodrell D, Ross PJ, Oates J: Twelve weeks of protracted venous infusion of fluorouracil (5-FU) is as effective as 6 months of bolus 5-FU and folinic acid as adjuvant treatment in colorectal cancer. *British Journal of Cancer* 88: 1859-1865 (2003)
- (48) Sakamoto J, Ohashi Y, Hamada C, Buyse M, Burzykowski T, Piedbois P: Efficacy of oral adjuvant therapy after resection of colorectal cancer: 5-year results from three randomized trials. *Journal of Clinical Oncology* 22: 484-492 (2004)
- (49) Saltz LB, Cox JV, Blanke C, Rosen LS, Fehrenbacher L, Moore MJ, Maroun JA, Ackland SP, Locker PK, Pirotta N, Elfring GL, Miller LL: Irinotecan plus fluorouracil and leucovorin for metastatic colorectal cancer. Irinotecan Study Group. *New England Journal of Medicine* 343: 905-914 (2000)
- (50) Sargent DJ, Goldberg RM, Stacy DJ, Jacobson MD, Macdonald JS, Labianca R, Haller DG, Shepherd LE, Seitz JF, Francini G: A pooled analysis of adjuvant chemotherapy for resected colon cancer in elderly patients. *The New England Journal of Medicine* 345: 1091-1097 (2001)

- (51) Sargent DJ, Wieand HS, Haller DG, Gray R, Benedetti JK, Buyse M, Labianca R, Seitz JF, O' Callaghan CJ, Francini G, Grothey A, O' Connell M, Catalano PJ, Blanke CD, Kerr D, Green E, Wolmark N, Andre T, Goldberg RM, De Gramont A: Disease-free survival versus overall survival as a primary end point for adjuvant colon cancer studies: individual patient data from 20,898 patients on 18 randomized trials. *Journal of Clinical Oncology* 23: 8664-8670 (2005)
- (52) Siewczynski R, Arnold D: Neoadjuvante und adjuvante Therapie des Rektumkarzinoms. *Onkologie-heute* 03: 16-24 (2007)
- (53) Simpson S H, Eurich D T, Majumdar S R, Padwal R S, Tsuyuki R T, Varney J, Johnson J A: A meta-analysis of the association between adherence to drug therapy and mortality. *British Medical Journal* 333: 15-19 (2006)
- (54) Schlesinger-Raab A, Eckel R, Engel J, Sauer HJ, Löhrs U, Molls M, Hölzl D: Metastasiertes Mammakarzinom: Keine Lebensverlängerung seit 20 Jahren. *Deutsches Ärzteblatt* (40): 2706-2714 (2005)
- (55) Smith RE, Colangelo L, Wieand HS, Begovic M, Wolmark N: Randomized trial of adjuvant therapy in colon carcinoma: 10-year results of NSABP protocol C-01. *Journal of the National Cancer Institute* 96: 1128-32 (2004)
- (56) Staib L, Link KH, Beger HG.: Toxicity and effects of adjuvant therapy in colon cancer: results of the German prospective, controlled randomized multicenter trial FOGT1. *Journal of Gastrointestinal Surgery* 5: 275-281 (2001)
- (57) Wils J, O'Dwyer P, Labianca R: Adjuvant treatment of colorectal cancer at the turn of the century: European and US perspectives. *Annals of Oncology* 12: 12-22 (2001)

9. Danksagung

Zunächst gilt mein herzlicher Dank Herrn Prof. Dr. med. Franz Porzsolt, mir das vorliegende Thema zur Bearbeitung im Rahmen einer Dissertation zur Verfügung gestellt zu haben. Herr Prof. Dr. Franz Porzsolt leitet seit Jahren erfolgreich den Lehrstuhl für Klinische Ökonomik an der Universität Ulm mit interdisziplinärer Ausrichtung. Infolgedessen fand ich in seiner Obhut als Doktorvater ein geeignetes Umfeld für die Bearbeitung dieser Fragestellung. Besonders hervorzuheben ist seine engagierte Betreuung auch bei persönlichen Gesprächen, im Verlauf derer sich stets neue Aspekte und Denkansätze ergaben. Seine Art des kollegialen Umgangs mit Menschen erzeugte ein positives Arbeitsklima, das entscheidend zum erfolgreichen Gelingen des Projekts beigetragen hat.

Insbesondere möchte ich auch Frau Dr. sc. hum. Christel Weiß (Leiterin der statistischen Abteilung an der Medizinischen Fakultät Mannheim der Universität Heidelberg) für ihre Unterstützung und Anregungen im Zusammenhang mit dieser Arbeit danken. Die statistisch-inhaltlichen Erörterungen waren mir sehr hilfreich und haben die Arbeit sehr bereichert. Schließlich kam erst durch ihre Initiative aufgrund der guten wissenschaftlichen Beziehungen zum Lehrstuhl für Klinische Ökonomik in Ulm das vorliegende Dissertationsprojekt zustande.

Auch allen anderen, die mir mit Anregungen behilflich waren, ein herzliches Dankeschön an dieser Stelle. Dabei möchte ich insbesondere auch Frau Sylvia Büttner (Medizinische Dokumentarin) und Frau Rosemarie Černý (Medizinische Dokumentarin) nennen, die mir in Layoutfragen sowie in der grafischen Aufbereitung der Ergebnisse mit Rat und Tat zur Seite standen.