

## Übersichten

### Prospektive, kontrollierte Studien in der Chirurgie. Kontroverse Standpunkte zur Motivierung und Durchführung

W. Lorenz und H. Rohde

Zentrum für Chirurgie der Philipps-Universität, Marburg (Lahn),  
Abteilung für experimentelle Chirurgie und pathologische Biochemie und Chirurgische Klinik

#### Prospective Controlled Clinical Trials in Surgery. Controversial Issues in Motivation and Performance

**Summary.** Controversial aspects in prospective controlled trials in surgery can be detected at various stages in the course of this research work, such as in motivation, planning and performance as well as in assessment and evaluation of the results. Especially the introduction of control groups, randomization and "blindness" of patients, operators and doctors running the follow-up are subjects of a controversial discussion. Subthreshold arguments for defenders and deniers of the trial are taken from different theories of science, such as empirism and rationalism which are more effective in the daily routine work than expected.

Provided that convincingly only therapeutic and diagnostic procedures are compared which were considered as equi-effective and reliable to the same extent according to the best and most complete knowledge available it is ethically justified to perform controlled randomized trials. It should, however, not be expected that they (always) produce differences between treatments of high degree of difference. The aim of the controlled trials is not the introduction of sensational novelties into medicine, but the abolition and prevention of unnecessary and unworthy modes.

**Key words:** Controlled clinical trials – Operation – Historical controls – Randomisation – Ethical aspects

**Zusammenfassung.** Kontroverse Standpunkte zu prospektiven, kontrollierten Studien in der Chirurgie gibt es bei der Motivierung, Planung und Durchführung

*Sonderdruckanfragen an:* Prof. Dr. W. Lorenz (Adresse s. nach Literatur)

sowie bei der Auswertung dieser klinischen Forschungsarbeiten. Im besonderen werden die Bildung von Vergleichsgruppen, die Randomisierung und die Blindheit von Patient, Operateur und Untersucher diskutiert. Unterschwellig spielen bei Befürwortern und Gegnern erkenntnistheoretische Gesichtspunkte eine Rolle, die mehr die ärztliche Praxis beeinflussen, als im Alltag offenkundig wird.

Unter der Voraussetzung, daß in überzeugender Weise nur therapeutische oder diagnostische Verfahren miteinander verglichen werden, die nach bestem Wissen und Gewissen als gleich wirksam oder gleich zuverlässig angesehen werden, sind prospektive kontrollierte, randomisierte Studien als ethisch voll gerechtfertigt anzusehen. Man darf aber auch dann von ihnen nicht erwarten, daß sie hochgradige Unterschiede zwischen den Verfahren zutage fördern. Ihr Ziel ist es nicht, sensationelle Neuheiten in die Medizin einzuführen, sondern in erster Linie unnötige und wertlose Moden auszumerzen und zu verhindern.

**Schlüsselwörter:** Kontrollierte klinische Studien – Operativer Eingriff – Historische Kontrolle – Randomisierung – Ethische Aspekte

---

Prospektive, kontrollierte Studien zur Entscheidungshilfe in Therapie und Diagnostik wurden im ersten Jahrzehnt nach dem 2. Weltkrieg in einige klinische Teilgebiete eingeführt, darunter auch in die Chirurgie [2, 18]. Aufgrund der Fragestellungen – ausschließlich Vergleich von klinisch als gleichwertig eingeschätzten Verfahren [1] – waren sie zunächst weniger in der praktischen Krankenversorgung als in der Veränderung des ärztlichen Denkens wirksam. In dieser Hinsicht war ihr Einfluß allerdings so nachhaltig, daß ihre Ergebnisse heute weder irgendeine medizinische Disziplin einschließlich der theoretischen Fächer [35] noch der Gesetzgeber [21] unbeachtet lassen darf.

**Tabelle 1.** Inhalt der Begriffe „Prospektiv“ und „Retrospektiv“. Die Tabelle wurde von Barth [3] nach Feinstein [14] zusammengestellt. „Cohort“ wurde von der römischen Cohorte abgeleitet, bei der die einzelnen Soldaten in Reih und Glied von dem geplanten Standort ihrer Aufstellung an *vorwärts* marschieren. „Trohoc“ ist die Umkehrung von „cohort“, d.h. hier marschieren die Soldaten von einem Halteplatz ihres Marsches gleichsam nach rückwärts. Daß es dabei viel Verwirrung gibt, kann man sich vorstellen. Es war die Absicht von Feinstein, mit diesem Bild die Problematik einer in die Vergangenheit gerichteten Auswahl der Patienten zu verdeutlichen

Zeitliche Richtung	Nicht differenziert	Differenziert nach	
		Sammlung von Daten	Betrachtung Population
Zukunft	Prospektiv	Prolektiv	Cohort
Vergangenheit	Retrospektiv	Retrolektiv	Trohoc

Gerade weil das Prinzip der prospektiven, kontrollierten Studien mit seinem hohen Anspruch an die medizinische Methodik erhebliche, oft geradezu erdrückende Konsequenzen in der Praxis hat, konnten kontroverse Standpunkte zu ihrer Notwendigkeit und zu den einzelnen Teilaspekten ihrer Durchführung nicht ausbleiben.

## 1. Begriffsbestimmungen

Die Begriffe „prospektiv“ und „retrospektiv“ stammen aus der Epidemiologie, wobei „prospektiv“ ursprünglich einfach „geplant“ bedeutete. Aber inzwischen gibt es den Begriff „geplante retrospektive Studien“, z.B. beim Zusammenhang Rauchen – Krebs [5]. Auch wurde die Unschärfe der beiden Begriffe immer wieder kritisiert. Deshalb schlug Feinstein [14] die in Tabelle 1 gegenübergestellten Bezeichnungen vor, die sich, wie eine Anfrage bei verschiedenen Statistikern ergab, beachtlicher Beliebtheit erfreuen und deshalb auch in dieser Arbeit empfohlen werden sollen. „Prolektiv“ bezieht sich dabei auf die *Datensammlung* nach einem vorher festgelegten Plan (Definitionen, Meßmethoden, Patientenauswahl, Untersucher etc.), „retrolektiv“ auf das Datenherausdestillieren aus bereits vorliegenden Krankengeschichten, auf deren Qualität (Fehler, Auslassungen, undefinierter Beobachter) wir nachträglich keinen Einfluß mehr haben. „Cohort“ bezieht sich dagegen auf unsere *Patienten*, wenn wir z.B. wissen wollen, wieviele Rezidive sie nach Beginn unserer Studie über die Vagotomie entwickeln *werden*, „trohoc“ (Umkehrung von „cohort“), wenn wir wissen wollen, wieviele von unseren Ulkus-Patienten vor ihrem Eintritt in die Studie wegen chronischem Ulcus duodeni streßvolle Lebenser-

eignisse hatten, z.B. Flüchtlinge aus dem letzten Krieg.

Eine „kontrollierte“ Studie bedeutet nicht eine Studie, deren Durchführung anhand bestimmter Kriterien überwacht wurde (z.B. Qualitätskontrolle), sondern eine Studie mit Kontroll- bzw. Vergleichsgruppen [27]. Da die Zuteilung von Patienten zur Testgruppe und Vergleichsgruppe mit verschiedenen Methoden erfolgen kann (s. unten bei 3.2.), muß diese in der Bezeichnung der Studie angegeben werden, z.B. eine „kontrollierte, randomisierte Studie“.

Bei Signifikanzprüfungen werden immer wieder die Begriffe „Nullhypothese ( $H_0$ )“ und „Prüfgröße ( $P$ )“ verwendet [27]. Unter der Nullhypothese versteht man die Annahme, daß die in der Test- und Vergleichsgruppe erhaltenen Daten alle zur selben Grundgesamtheit (Population) gehören. Ihr Unterschied, den man durch eine bestimmte Untersuchungs- bzw. Meßmethode in einer Stichprobe ermittelt hat, wäre daher rein zufällig zustande gekommen und damit kein „signifikanter“ Unterschied. Mit der Prüfgröße ( $P$ ) wird dann berechnet, ob die Behauptung der Nullhypothese zutrifft.

$P=1$  bedeutet Zutreffen in 100% der Fälle (kein Unterschied),  $P=0.05$  bedeutet Zutreffen in nur 5% der Fälle, was in der Regel und vereinbarungsgemäß genügt, um die Nullhypothese abzulehnen und die Alternativhypothese zu akzeptieren (signifikanter Unterschied).

## 2. Kontroverse Standpunkte bei der Motivierung zu prospektiven, kontrollierten Studien

Die Motivierung zu prospektiven, kontrollierten Studien hängt von der Anerkennung ihrer prinzipiellen Notwendigkeit und der Bejahung ihrer Ergebnisse im allgemeinen ab (Tabelle 1).

Nun haben sich aber in Europa über die letzten drei Jahrhunderte hin im wesentlichen zwei erkenntnistheoretische Schulen entwickelt, die als Empirismus (Bacon, Locke, Hume usw.) und Rationalismus (Descartes, Spinoza, Leibniz usw.) charakterisiert werden [45]. Für die erstere gilt nur die Beobachtung als Quelle der Erkenntnis, für letztere nur die intellektuelle Anschauung von klaren und deutlichen Vorstellungen [45]. Dieser von Popper so scharf herausgestellte Gegensatz erscheint aber überspitzt: Man findet Elemente des Empirismus auch bei den Rationalisten und umgekehrt. Ohne daß wir es im Alltag unseres Berufes merken, beeinflussen diese beiden Schulen in entscheidender Weise unsere Einstellung.

(1) Der empiristische Grundsatz „in 10, 20, 30 ... Fällen wurde beobachtet, daß ... also gilt, daß ...“ führt uns unmittelbar zur prospektiven, kontrol-

lierten Studie mit all ihren Forderungen nach Objektivität (Kontrollgruppen, Wiederholung, Randomisierung, Blindheit usw.).

(2) Der rationalistische Grundsatz „mein Nachdenken, meine Überzeugung, meine Ansicht, mein Gefühl sagen mir in diesem Fall, daß ...“ benötigt grundsätzlich keine prospektiven kontrollierten Studien und räumt ihnen deshalb auch keinen wesentlichen Einfluß auf die klinische Praxis ein. Die Betonung des Einzelnen, des Subjekts in diesem Denksystem (Descartes „*Ich* denke, also bin *ich* [50]) führt in der ärztlichen Praxis zur Autorität und findet ihre logische Formulierung in der Redewendung „*meine* Erfahrung sagt mir, daß ...“. Autorität mag im Einzelfall notwendig oder hemmend, positiv oder negativ sein, sie ist und bleibt ein anderes Prinzip als die 20–30 Fälle.

Damit beruht der gegensätzliche Standpunkt bei der Motivierung zu prospektiven kontrollierten Studien zwischen den Angloamerikanern und Kontinentaleuropäern [39, 48, 49] nur scheinbar und oberflächlich auf „modernen“ bzw. „antiquierten“ Ansichten, in Wirklichkeit aber auf deutlichen, jahrhundertalten Unterschieden in der Denkweise. Gegenwärtig feiert der Empirismus Triumphe. Nach Karl Popper [46] irren aber beide erkenntnistheoretischen Schulen: Zuverlässiges Wissen ist weder durch Anschauung allein noch durch die Beobachtung von noch so vielen Fällen zu erlangen, sondern nur durch kritische Anpassung an den von der jüngsten Widerlegung noch nicht betroffenen Kenntnisstand. Diesen Schluß findet man *nach* Beendigung kontrollierter Studien immer wieder bestätigt, wenn eine „perfekt“ geplante und durchgeführte Studie eben doch Mängel aufweist, weil Faktoren vergessen oder in der Laufzeit der Studie dazugekommen sind [10, 59].

### 3. Kontroverse Standpunkte zur Planung und Durchführung von prospektiven kontrollierten Studien

Das oben geschilderte erkenntnistheoretische Dilemma durchzieht wie ein roter Faden auch die Kontroversen zu den einzelnen Schritten der Planung und Durchführung von kontrollierten Studien (Tabelle 1).

#### 3.1. Bildung von Vergleichsgruppen

Sie ist ein Hauptschritt in der Planung der prospektiven kontrollierten Studien. In Tierexperimenten wird sie ohne prinzipielle Einwände und seelische Skrupel bejaht, in klinischen Studien dagegen aus ethischen und juristischen Gründen von verschiedenen Autoren

entweder völlig [15] oder unter bestimmten Bedingungen [20] abgelehnt. Im Einzelfall entzündet sich der Streit vor allem an der Abwägung von Notwendigkeit gegen Risiko bei Vergleichsgruppen (Tabelle 1). Beide werden aus Vorinformationen, aber auch aus den persönlichen Vorstellungen und – nicht zu vergessen – Vorurteilen über die Größe des erwarteten Fortschritts bzw. Schadens abgeschätzt. Die Kontroverse wird dann besonders heftig, wenn das Risiko den Verlust von Menschenleben miteinschließt. Dies ist bei chirurgischen Eingriffen immer der Fall.

Die *Befürworter* von prospektiven kontrollierten Studien bejahen die Bildung von Vergleichsgruppen. Sie halten sie für unbedingt notwendig (eine *conditio sine qua non*) und auch für ethisch vertretbar. Denn sie nehmen für das Zeitintervall zwischen Planung und Ende der Studie die gleiche Effektivität der in der Prüfung stehenden Verfahren hinsichtlich erwünschter *Wirkungen* und unerwünschter *Nebenwirkungen* für gültig an (Nullhypothese) [1, 27, 30]. Diese statistische Aussage, in die chirurgische Praxis übertragen, stellt aber an den prüfenden Chirurgen erhebliche Anforderungen. Er muß nicht nur subjektiv, sondern auch aufgrund alles nur irgendwie erreichbaren Wissens (das heißt heute, nach äußerst aufwendigem Studium) überzeugt sein und glaubhaft machen können, daß die zu testenden Diagnoseverfahren *gleich* zuverlässig und die zu untersuchenden Therapieverfahren *gleich* wirksam sind. Dies hat z.B. D. Johnston [29] bis heute daran gehindert, beim chronischen unkomplizierten Ulcus duodeni eine kontrollierte Studie über die selektiv-proximale Vagotomie ohne Drainage gegen ein anderes dabei verwendbares Operationsverfahren durchzuführen: Nach Abwägung aller Pro- und Contraargumente ist er von den Vorteilen seiner Methode bis zum jetzigen Zeitpunkt absolut überzeugt.

Kontrollierte Studien in der Chirurgie eignen sich deshalb nicht zur Bestätigung einer positiven oder negativen Anschauung ohne glaubhafte persönliche Unsicherheit (kein subjektiver Zweifel!). Sie sind nicht gerechtfertigt bei überzeugenden Erfolgen eines neuen Therapieverfahrens, wo bisher alle anderen Mittel versagten (z.B. in einer Situation wie bei der Einführung der Sulfonamide oder des Insulins) (kein objektiver Zweifel!). Sie sind nicht als sportliches Vergnügen im Zeitalter der prospektiven Studien geeignet (x-te Bestätigung bereits mehrfach geprüfter und eindeutig entschiedener Fragestellungen, z.B. zum Zwecke der Erlangung einer Habilitation) oder „weil man halt so eine Studie gemacht haben muß, sonst ist man nicht modern“.

Die *Gegner* von prospektiven kontrollierten Studien sind vor allem aus ethischen Gründen gegen Vergleichsgruppen. Fincke [15] fragt zum Beispiel, was

mit einem Prüfer geschehen soll, der ein im Tierversuch wirksames Antikrebsmittel in einer kontrollierten klinischen Studie mit Erfolg eingesetzt hat und nun von der Witwe eines Patienten in der Vergleichsgruppe angezeigt wird, weil er ihm das lebensrettende Mittel versagt habe. Diese Frage erscheint unsinnig, wenn von der oben erläuterten Ausgangslage vor und während der Studie (Annahme der Nullhypothese) ausgegangen wird. Fincke erscheint gegenüber kontrollierten Studien voreingenommen [30] und seine Auslegung der Trapnellstudie hinsichtlich Trasyolol und Pankreatitis [61] hat letzterer mit einer weiteren kontrollierten Multicentrestudie unabsichtlich, aber überzeugend widerlegt [42].

Aber so einfach sind die Entscheidungen für oder gegen die kontrollierten Studien in der Chirurgie nicht: Soll man nach Tierversuchen sofort mit einer kontrollierten Studie beginnen oder zuerst in einer Pilotstudie dem neuen Verfahren die Chance geben, ein überragender Erfolg zu sein? Ausgehend von der Chirurgie der portalen Hypertension spricht sich Warren [68] für letzteres aus und wird generell von Burdette und Gehan [5] unterstützt, Chalmers [9] plädiert dagegen für die kontrollierte randomisierte Studie bereits beim ersten Patienten. Weil jeder Prüfer die gleiche Effektivität der in der Studie angewendeten Verfahren für sich und andere *überzeugend* vertreten muß, wozu ein dem Problem angemessenes Wissen notwendig ist, empfehlen wir, nicht nach Tierversuchen sofort mit einer kontrollierten Studie am Patienten zu starten.

Der Statistiker mahnt uns: „Vergleichsgruppen müssen vergleichbar sein“ (Immich [27]).

(1) Freiwillige Versuchspersonen als Kontrollgruppe und Patienten als Testgruppe sind nicht vergleichbar [11, 13, 26]. Zu viele Unterschiede wurden zwischen ihnen entdeckt [12]. Studien über chirurgische Fragestellungen weisen seltener diesen Fehler auf: Beim Vergleich zweier Operationsverfahren läßt sich wohl niemand freiwillig ohne Indikation operieren. Aber Studien über pathophysiologische Fragen oder Arzneimittel, die mit diesem Fehler behaftet sind, gibt es zu Hunderten.

(2) In der Praxis klinischer Studien ist es oft schwierig, geeignete Vergleichsgruppen zu bilden. Möglichst sollen nicht nur solche mit sog. „gesunden“ Patienten zusammengestellt werden (z.B. Unfallgeschädigte mit Frakturen beim Vergleich mit Ulcus-duodeni-Kranken), sondern auch Vergleichsgruppen mit Patienten, die eine Erkrankung desselben Organbereichs (z.B. Magenulkus, Magenkarzinom) oder ähnliche Beschwerden aufweisen (z.B. „chronische Galle“ oder chronische Pankreatitis). Gründe hierfür ergeben sich aus den Erfahrungen mit der systematischen Kontrolluntersuchung [67], aber auch in der

klinischen Chemie [19]. Diesen Forderungen haben wir mit erheblichem Aufwand versucht, in einer kontrollierten Studie beim chronischen Ulcus duodeni weitgehend gerecht zu werden [64, 65].

### 3.2. Zuteilung zu den Vergleichsgruppen

Der zweite Hauptschritt in der Planung von prospektiven kontrollierten Studien – sieht man vom Prinzip der Wiederholung hier ab – ist die Zuteilung von Patienten zu den einzelnen Test- und Vergleichsgruppen. Hierfür gibt es mehrere, aus der Medizinhistorie übernommene und systematisch entwickelte Verfahren, die wiederum Gegenstand heftiger Kontroversen sind.

Zunächst einmal können Patienten überhaupt nur der Testgruppe zugeordnet werden, während die Vergleichsgruppe erdacht, aus der Literatur herausdestilliert oder einem früheren (eigenen oder fremden) Krankengut entnommen wird. Werden Testgruppe und Vergleichsgruppe tatsächlich gebildet, so kann dies hintereinander (konsekutiv) oder gleichzeitig geschehe. Zur Kritik der konsekutiven Studien s. S. 308. Bei der gleichzeitigen Zuordnung kann man dies nach verschiedenen Ordnungsprinzipien („systematisch“, z.B. alternierend, je nach Einweisung in das Krankenhaus oder dem Geburtstag an einem Datum mit gerader oder ungerader Ziffer, nach Adaptation an einen Vorgänger [60, 69]) oder nach der Losentscheidung (Randomisierung) durchführen. Heftige Kontroversen entzündeten sich heute fast ausschließlich an der Frage, ob und ob immer randomisiert werden soll.

Die *Befürworter* der Randomisierung halten dieses Verfahren für absolut notwendig, allen anderen Zuteilungsverfahren überlegen und für ethisch vertretbar (Tabelle 1). Zu ihnen gehören wenigstens 90 Prozent aller Versuchsplaner, Informationswissenschaftler, Biostatistiker und auch klinischer Biostatistiker [14]. Argumente für die Randomisierung finden sich bei Yates [70], Fisher [17], Greenberg [24], Hill [26], Pflanz [44], Immich [27, 28], Colton [11], Ederer [13], Byar [7], Peto [43] und Feinstein [14]. Auch in der Chirurgie verfiert eine beträchtliche Mehrzahl von Ärzten, die klinische Studien durchgeführt *haben*, das Prinzip der Randomisierung, nachdem sie sich „am eigenen Leibe“ von den Vorteilen dieses Zuteilungsverfahrens überzeugen konnten [1, 23, 47]. Auch in unserer Klinik haben wir es uns angewöhnt, bei jeder nur möglichen klinischen und sogar tierexperimentellen Fragestellung eine Randomtabelle zur Hand zu nehmen (z.B. eine Kopie von Tabelle A1 aus Snedecor und Cochran [58]). Damit wird in wenigen Minuten die Zuordnung von Patienten, Testpersonen oder Ver-

suchstieren – statt „irgendwie“ – streng zufällig zu den Test- und Vergleichsgruppen erreicht [16, 31, 32, 36].

Die Randomisierung ist notwendig, da nur so eine bewußt und unbewußt gewollte wie auch bei Blindstudien eine völlig unbeabsichtigte Einseitigkeit bei der Patientenzuteilung am ehesten verhindert werden kann. Weiterhin hat nur so jeder Patient die optimal gleiche Chance, einem jeden der im Vergleich stehenden Verfahren zugeordnet zu werden, wodurch die Anwendung statistischer Prüfverfahren überhaupt erst zulässig wird. Nur *so* verteilen sich schließlich die Effekte von unbekanntem Störfaktoren gleichmäßig auf alle Gruppen und werden damit am ehesten neutralisiert [27]. Bei anderen Zuteilungsverfahren wurden Vorkommen und Möglichkeit systematischer Fehler nachgewiesen, so bei alternierender Reihe [24, 26, 11, 13] und bei der Zuteilung anhand der Aufnahme in das Krankenhaus oder der Geburtsdaten [11]. Dies ist besonders für chirurgische Fragestellungen von Bedeutung. Beim Vergleich von zwei Operationsmethoden mit alternierender Reihe weiß der Operateur immer, welches Verfahren beim nächsten Patienten „drankommt“. Je nach persönlicher Überzeugung wird er dazu neigen, den nächsten Patienten entweder in die Studie zu nehmen oder aus ethischen Gründen auszuschließen, wodurch die Begünstigung eines Verfahrens (ein systematischer Fehler) zustandekommt.

Das Verfahren der Randomisierung sollte kritisch propagiert, aber nicht idealisiert werden. Es gibt Studien, bei denen die durch Randomisierung gewonnenen Test- und Vergleichsgruppen geradezu wie eineiige Zwillinge übereinstimmen (Beispiel s. [13]) oder nur geringfügig voneinander abweichen (Beispiele s. [11, 36]). Häufig aber findet man doch beachtliche Unterschiede, je mehr Merkmale der Patienten, wie z.B. Sozialstruktur, Begleitkrankheiten etc., man zum Vergleich heranzieht. Man sollte wissen, daß Randomisierung nur über die sehr große Zahl zu einer optimalen Ähnlichkeit von Prüf- und Vergleichsgruppe führt – eine Zahl, die in der Praxis oft unerreichbar ist.

Die ethische Rechtfertigung erfährt die Randomisierung ebenso wie die Einrichtung einer Vergleichsgruppe (s. oben) aus der begründeten Annahme der gleichen Effektivität der in der Prüfung befindlichen Verfahren. Man sollte sich deshalb aber in der Chirurgie dann auch nicht wundern, wenn bei vielen kontrollierten klinischen Studien am Ende eben auch kein Vorteil für eines der verglichenen Operationsverfahren herauskommt (Zusammenfassung für die Studien „Vagotomie gegen Resektion“ s. Troidl et al. [63]).

Die *Gegner* der Randomisierung finden sich heute überwiegend in zwei Lagern. Die einen halten sie vor

allem bei Krebsstudien für ethisch fragwürdig und den „historischen Kontrollen“ gegenüber nicht unbedingt von Vorteil [20]. Die anderen finden, daß die einfache (= nicht stratifizierte) Randomisierung zu schlechter Balancierung und relativ großer Heterogenität des Krankengutes führt (Tabelle 1). Sie schlagen deshalb andere Wege der Zuordnung von Patienten zu Test- und Vergleichsgruppen vor, die als „adaptive“ Verfahren [69] bzw. als „mimization“ [33, 60] bezeichnet werden.

Verschiedene Autoren in unserem Land, die bei der Neuschaffung des Arzneimittelgesetzes (Zusammenfassung s. Fincke [15]) bzw. als Gegner des modischen Trends der 70iger Jahre, kontrollierte klinische Studien zu starten, sich gegen randomisierte Studien ausgesprochen hatten, benutzten die Argumente der oben genannten amerikanischen Autoren zu einer Attacke gegen die Randomisierung überhaupt. Hier liegen verschiedene Mißverständnisse vor, die gelöst werden müssen: 1. Keiner der genannten amerikanischen Wissenschaftler war prinzipiell gegen die Randomisierung bei klinischen Studien, Gehan beschrieb in seinem Buch sehr ausführlich die Vorzüge dieser Patientenzuteilung [5]. 2. Die neuen Verfahren zur Allokation von Patienten zu Test- und Vergleichsgruppe sind als Alternativen zur Randomisierung vorgeschlagen und diskutiert worden, ihre Bewährung steht noch aus. Für die Beseitigung des Nachteils der sog. einfachen Randomisierung (Heterogenität der Gruppen) wird seit langer Zeit die Bildung von Blöcken [27] bzw. die stratifizierte Randomisierung [14] vorgeschlagen. 3. Die Nachteile der sog. „historischen Kontrolle“ (Verwendung des früheren Krankengutes der Klinik als Vergleichsgruppe für die Testgruppe) werden auch von Gehan und Freireich [20] gesehen. Am Beispiel des Prostatakarzinoms wiesen aber Byar et al. [7] die enormen Fehlschlüsse bei historischen Kontrollen nach, Pocock gelang dies bei 10 von 19 Studien über eine maligne Erkrankung [43]. Die ausgewogene Stellungnahme der MRC-Gruppe [43] zur Durchführung randomisierter klinischer Studien auf dem Karzinomgebiet erscheint uns dabei besonders empfehlenswert und für den Kliniker und experimentellen Chirurgen als „Anleitung“ besonders wertvoll.

### 3.3. Blindheit von Patient, Operateur und Untersucher

Verfahren zur Kanalisierung und Einschränkung von Entscheidungen durch Nichtwissen („Blindheit“) in einer kontrollierten klinischen Studie dienen wie die Randomisierung der Vermeidung von Voreingenommenheit im speziellen und von systematischen Fehlern (Trendbildung) im allgemeinen. Wie mit der Randomisierung erreichen wir damit nicht das Idealziel einer

perfekten Planung. Die vielen Doppel-Blindstudien mit dem Ulkustherapeutikum Cimetidin sind hierfür ein auffälliges Beispiel [6]: Je nach Einstellung des „blinden“ Therapeuten zum Patienten und zum neuen Heilverfahren werden dem Kranken so viele unterschiedliche Einflüsse (Trends, Voreingenommenheiten) vermittelt, daß die Placeboheilungsrate sich zwischen 30 und 80 Prozent bewegt und damit im oberen Extremfall dem Cimetidin überhaupt keine Chance mehr läßt, „signifikant“ wirksam zu sein [4].

In der Chirurgie ist Blindheit von Patient, Operateur und Untersucher ein so vielschichtiges Problem, daß es mit Patentrezepten nicht gelöst werden kann. Für den Operateur ist die Aufklärungspflicht gegenüber dem Patienten komplizierter als für den Tabletten verschreibenden Arzt, da eine Operation für den Kranken in der Regel ein akutes, bedrohliches Ereignis mit irreversiblen Folgen darstellt. Bemerkenswert erscheint in diesem Zusammenhang die Auffassung von Atkins [1], die er als Präsident der englischen Chirurgenvereinigung vertrat: Wenn man wirklich nicht weiß, welches von zwei Operationsverfahren besser in der Behandlung eines schwerkranken Patienten ist, kann es *nicht* ärztlich vertretbar sein, ihm klarzumachen, daß man um seine Entscheidung lost. Wir kamen allerdings zu der Meinung, daß es meist von der Kunst der Gesprächsführung abhängt, dem Patienten zwei gleichwertige Möglichkeiten der Therapie nahezubringen, wie wir in unserer Kontrolluntersuchungsklinik für Magenkranke immer wieder erlebten [51]. Deshalb plädieren wir in der kontrollierten Studie für das Ausmaß an Aufklärung, das wir auch dem nicht in die Studie einbezogenen Kranken vermitteln.

Kein Operateur ist während des chirurgischen Eingriffs „blind“, welches Verfahren er durchführt. Hierin liegt ein beträchtliches Risiko für kontrollierte Studien. Trotz bester Absicht, fair zu sein, operiert der Chirurg „lieber“ nach seiner Technik, die er selbst erfunden hat, oder er überkompensiert aus Angst, nicht fair zu sein. Gerade die letzte Möglichkeit wurde bisher zu selten bedacht [26], die „Redlichkeit des Forschers“ [40] kann auch sein größter systematischer Fehler werden. Daß eine nicht gleichwertige Anwendung von zwei Techniken zum Scheitern der Studie führt, wurde von Liedberg und Oskarson [34] bei der selektiv-proximalen Vagotomie anschaulich demonstriert. Entscheidend wichtig ist aber, daß der Chirurg bei der Operationsindikation hinsichtlich der Technik „blind“ ist. Er muß erst entscheiden, ob bei einem Patienten *beide* zu prüfenden Techniken oder Verfahren aus ärztlichen Gründen möglich sind (auch „nicht operieren“ ist eine Therapie!), andernfalls muß er den Kranken in die Fluchtklausel nehmen und nach seinem Gewissen behandeln. Nachdem er sich entschie-

den hat, erfährt er z.B. nach öffnen eines Briefumschlages oder bei einer multizentrischen Studie z.B. nach einem Telefonat, welche der beiden Techniken er aufgrund der Randomisierung verwenden muß. Um dem Operateur für den individuellen Patienten die freie ärztliche Entscheidung zu lassen, muß mit dem Bekanntgeben der Losentscheidung so lange wie möglich gewartet werden. Oft wird erst nach gründlicher Exploration des Abdomens entschieden, ob der Patient „in die Studie kommt oder nicht“ (vgl. Goligher et al. [23]).

Ob der Vor- und Nachuntersucher chirurgischer Patienten „blind“ sein soll, ist ebenfalls Gegenstand noch nicht entschiedener Kontroversen und kann im Einzelfall unterschiedlich beantwortet werden. Die gegensätzlichen Entscheidungskriterien sind hierfür „Vermeidung von Voreingenommenheit bei Kenntnis des Verfahrens“ und „Beurteilungsfähigkeit durch Verständnis spezieller Probleme des chirurgischen Patienten“. Hierzu haben wir bereits ausführlich Stellung genommen [37, 51] und uns für eine komplexe Untersuchung durch Stellungnahme des Patienten, des unabhängigen Internisten und eines Panels aus Chirurgen, theoretischen Chirurgen [35, 51, 67] und Internisten entschlossen (Ergebnisse s. Seidel et al. [57], Troidl et al. [64, 65, 66]). Die Beurteilung des Operationserfolges unterliegt einer enormen Beobachtervariation [22, 25] und wird vom Patienten erheblich anders gesehen als vom untersuchenden Arzt [8].

### 3.4. Aufwand (Stichproben- und Meßgrößenumfang)

*Befürworter* von kontrollierten klinischen Studien sehen den Aufwand für gerechtfertigt an, weil die Erkennung von objektiv besseren Behandlungsmethoden letztlich Kosten sparen hilft [37]. *Gegner* dieser Methode sehen sie dagegen als zu teuer hinsichtlich ihres ungewissen Ausgangs oder gar Scheiterns, als zu risikoreich und in der Anwendung ihrer Ergebnisse auf den Patienten als zu fragwürdig an, als daß sie der Gesellschaft wirklich in ausreichendem Maße nützt (Tabelle 1).

Bei kontrollierten klinischen Studien denken viele Ärzte zunächst an Stichprobenumfänge von Hunderten von Patienten, an jahrelange Laufzeit, Notwendigkeit des Computereinsatzes oder an eine Serie von Universitätskliniken (multizentrische Studie). All dies ist keinesfalls notwendig. Stichprobenumfänge von zweimal 16 oder zweimal 30 Patienten oder Testpersonen [27] reichen für viele Fragestellungen bereits aus. Beispiele hierfür finden sich in unseren Arbeiten über Blutersatz [36, 38], über Magensekretionsteste bei Vagotomie [62, 52] sowie in den Studien von Fischer und Hamelmann [16], Largiadèr und Säuberli [32]

**Tabelle 2.** Kontroverse Standpunkte zu prospektiven kontrollierten Studien (Übersicht)

Phase der kontrollierten Studie	Formulierung des Standpunkts	
	Zustimmend	Ablehnend
<i>Motivierung</i>		
Prinzipielle Notwendigkeit	Quelle der Erkenntnis ist die Beobachtung	Q. der E. ist die intellektuelle Anschauung von klaren und deutlichen Vorstellungen
Ergebnis im allgemeinen	Zuverlässig und nützlich	Fragwürdig und ohne Einfluß auf die klinische Praxis
<i>Planung und Durchführung</i>		
Bildung von Vergleichsgruppen	Notwendig und ethisch vertretbar, weil $H_0$ angenommen	Nicht notwendig und ethisch verwerflich
Zuteilung zu Vergleichsgruppen	Randomisierung ist notwendig, allen anderen Verfahren überlegen und ethisch vertretbar	Randomisierung ist nicht nötig, führt oft auch zu schlechter Homogenität, ist ethisch fragwürdig
Blindheit von Patient und Untersucher	Äußerst empfehlenswert zur Ausschaltung von Voreingenommenheit	Widerspricht der Aufklärungspflicht und der Redlichkeit des Forschers
Aufwand (Stichproben- und Meßgrößenumfang)	Gerechtfertigt durch zuverlässiges Ergebnis	Zu teuer, zu risikoreich und zu fragwürdig
<i>Auswertung</i>		
Wert der induktiven Statistik	Testverfahren sind nötig für zuverlässige Entscheidung	Mit Statistik kann man alles beweisen
Signifikanz für die Klinik (determinacy)	Statistische Signifikanz ist prinzipiell wertvoll, unabhängig von der klinischen Signifikanz	Klinische Signifikanz ist wichtiger als statistische Signifikanz

und in der ausgezeichneten Plasmaersatzmittelstudie von Schöning und Koch [55], die als einzelne Kliniker innerhalb eines Jahres erfolgreich ihre Arbeit abschließen konnten.

#### 4. Kontroverse Standpunkte in der Auswertung von prospektiven, kontrollierten Studien

Man kann bei der Auswertung von kontrollierten klinischen Studien zwei Arten von Statistik unterscheiden, die deskriptive und induktive [14]. Erstere dient der Verminderung des Datenberges [17] durch Zusammenfassung der Ergebnisse in Mittelwerten, Bereichen, Klassen etc. Letztere hilft uns die Frage beantworten, ob Beobachtungen etwas miteinander zu tun haben oder sich voneinander mit einer bestimmten Wahrscheinlichkeit unterscheiden.

Für das praktische Vorgehen gibt es eine Serie von Testen, mit denen wir eine bestimmte „Signifikanz“ errechnen können. Diese mathematischen Prozeduren werden – pars pro toto – im täglichen Umgang als Statistik bezeichnet und „kräftig beschimpft“. „Mit Statistik kann man alles beweisen“

lautet das ablehnende Urteil der Gegner von prospektiven, kontrollierten Studien, während die Befürworter Testverfahren für eine zuverlässige Entscheidung für notwendig halten (Tabelle 1).

Leider haben viele Theoretiker in der Medizin infolge fehlender Ausbildung ein gestörtes Verhältnis zur Biostatistik. Aus dem eingangs geschilderten erkenntnistheoretischen Dilemma wurde diese angewandte Mathematik abgelehnt, weil sie in der Tat beim Versuchstier häufig nicht benötigt wurde. Die experimentellen Bedingungen wurden so lange „hingetrimmt“, bis die massiven Unterschiede zwischen den Ergebnissen von Test- und Kontrollversuch keine statistische Analyse mehr verlangten. Bei der Argumentierung gegen kontrollierte Studien wurde übersehen, daß prospektive klinische Studien nicht unter den Extrembedingungen von Experimenten ablaufen dürfen und deshalb nicht extrem unterschiedliche Resultate in Test- und Kontrollgruppe erwarten lassen. Die Definition der Statistik als Entscheidungshilfe, die unter anderem auf Abraham Wald (1902–1950) zurückgeht, ist eine Zusammenfassung von Methoden, die uns erlauben, vernünftige optimale Entscheidungen im Falle von Ungewißheit zu treffen [54],

bei der die Kliniker vom Theoretiker in der Regel allein gelassen werden. Ein Beispiel hierfür erlebten zwei Chirurgen letztes Jahr in den USA. Sie besuchten zunächst M. Mc Knealey in Albany (New York), der erfolgreich eine kontrollierte klinische Studie über die chirurgische Therapie des Bronchialkarzinoms (Stadium I) in Kombination mit BCG Immunstimulierung durchgeführt hatte [41]. Dann fragten sie den Immunpathologen Borsos im National Cancer Institute, Bethesda, was er aufgrund seiner Tierexperimente von der Mc Knealey-Studie halte. Er lehnte sie ab, sagte er, da seine Meerschweinchen anders reagiert hätten. Auf die Frage, was dann die Kliniker tun sollten, meinte er nur: „Das ist deren Entscheidung“.

In kontrollierten klinischen Studien ist oft der statistisch gesicherte (signifikante) Unterschied sehr klein. Es kann nicht oft genug betont werden, daß dies meist die Folge der Ausgangssituation für die Durchführung der Studie ist, nämlich der Vergleich von augenscheinlich gleichwertigen Verfahren. Befürworter der Studien sehen einen statistisch signifikanten Unterschied prinzipiell für wertvoll an, weil er eine Entscheidungshilfe darstellt, während die Gegner der Studien einen „beträchtlichen“, klinisch signifikanten Unterschied verlangen (Tabelle 1).

Die klinische Signifikanz (determinacy nach Atkins [1]) kann aber in Folgestudien oft erreicht werden, wenn die statistische Signifikanz analysiert wird und die Erkenntnisse daraus zu mehr eingegrenzten Studienbedingungen führen (geeigneter Blockbildung, im Falle der Behandlung von Brustkrebs s. Atkins [1]).

## 5. Schlußfolgerungen

Gegenwärtig ist in Deutschland die Stellungnahme zu kontrollierten klinischen Studien in der Chirurgie geprägt und geleitet durch kritischen Optimismus und prinzipielle Bejahung, aber auch durch Unsicherheit, Zweifel und verständliche Emotionen gegen Zwänge, Belastungen und Bedrohungen [53]. Nach unserer Überzeugung, entwickelt aus Umgang und Auseinandersetzung mit ihr, ist die geplante, kontrollierte klinische Studie ein wertvolles Werkzeug. Wir müssen lernen, wann generell und wann im Einzelfall wir es einsetzen müssen und dürfen. Wir müssen lernen, wie wir es handhaben können und sollen und mit welchen vermeidbaren und unvermeidlichen Fehlern wir zu rechnen haben.

Nach unserer Überzeugung ist es gegenwärtig aber auch ethisch nicht zu vertreten, die kontrollierte klinische Studie grundsätzlich abzulehnen. Nicht nur der begeht einen Fehler, der beim einzelnen Patienten

oder bei einer Serie von Patienten vorsätzlich oder fahrlässig falsche Hilfe leistet, sondern auch der, der – nicht nur im Einzelfall, sondern auch in der Serie – Hilfe versagt. Die Diskussion in Presse, Rundfunk und Fernsehen zur kontrollierten Studie in der Arzneimittelforschung wird überwiegend emotional, weltanschaulich und mit Schlagworten geführt, statt daß die umfangreiche Literatur zu diesem Thema gelesen und durchdacht sowie eigene Erfahrungen erworben werden. Angesichts der Wichtigkeit und Grundsätzlichkeit des Problems ist diese Haltung wenig verantwortungsbewußt.

In „chirurgische Perspektiven“ hat H.W. Schreiber [56] die breitere Entwicklung kontrollierter klinischer Studien für die Zukunft gefordert. Sie werden nicht geniale Entdeckungen in der Chirurgie verhindern, wohl aber so viele unnötige und wertlose Moden, wie sie in der ersten Hälfte dieses Jahrhunderts produziert wurden [1]. Der durch die Krankenversorgung überlastete klinische Chirurg kann nicht allein auch noch die Last der Planung, Organisation und Auswertung von prospektiven, kontrollierten Studien tragen. Hier ist die experimentelle Chirurgie aufgefordert, an der praktischen Krankenversorgung teilzunehmen [35, 51]. Es gibt kein Gebiet, auf dem sie ihre Daseinsberechtigung mehr unter Beweis stellen könnte.

*Danksagung.* Die Autoren danken vielmals H. Immich für wertvolle Anregungen und die kritische Prüfung des Manuskriptes. Sie danken weiterhin H.W. Schreiber, H. Jahrmärker, B. Schöning und H. Barth für ihre Mühe bei der Durchsicht der Arbeit und für ihre Diskussionsbeiträge sowie der Deutschen Forschungsgemeinschaft (Schwerpunkt Klinische Pharmakologie, Lo 199/8) für ihre finanzielle Unterstützung.

## Literatur

1. Atkins, H.J.B.: Conduct of a controlled clinical trial. *Brit. Med. J.* **1966****II**, 377
2. Atkins, H.J.B., Falconer, M.A., Hayward, J.L., Mac Lean, K.S.: Adrenalectomy and hypophysectomy for advanced cancer of the breast. A comparative study, *Lancet* **1957****I**, 489
3. Barth, H.: Planung und Datenanalyse bei prospektiven kontrollierten klinischen Studien. Kongreßband der 95. Tagung der Deutschen Gesellschaft für Chirurgie 1978, *Langenbecks Arch. klin. Chir.* (in Druck)
4. Blum, A.L.: Vortrag auf dem Göttinger Symposium über Cimetidin, Nov. 1977
5. Burdette, W.J., Gehan, E.A.: Planning and analysis of clinical studies. Springfield, Illinois: C.C. Thomas 1970
6. Burland, W.L., Simkins, M.A.: Cimetidine. *Proc. 2<sup>nd</sup> Intern. Symp. on histamine H<sub>2</sub>-receptor antagonists.* Amsterdam: Excerpta Medica 1977
7. Byar, D.P., Simon, R.M., Friedewald, W.T., Schlesselman, J.J., DeMets, D.L., Ellenberg, J.H., Gail, M.H., Ware, J.H.: Randomized clinical trials. Perspectives on some recent ideas. *New Engl. J. Med.* **295**, 74 (1976)
8. Cay, E.L., Philip, A.E., Small, W.P., Neilson, J., Henderson, M.A.: Patient's assessment of the result of surgery for peptic ulcer. *Lancet* **1975****I**, 29

9. Chalmers, T.C.: Randomization of the first patient. *Med. Clin. North America* **59**, 1035 (1975)
10. Christensen, E., Juhl, E., Tygstrup, N.: Treatment of duodenal ulcer. Randomized clinical trials of a decade (1964 to 1974). *Gastroenterology* **73**, 1170 (1977)
11. Colton, Th.: *Statistics in medicine*. p. 255. Boston: Little, Brown and Comp., 1974
12. Crocetti, A.: Volunteering in medical research. Dissertation, John Hopkins Univ., Baltimore 1970
13. Ederer, F.: Why do we need controls? Why do we need to randomize? *Amer. J. Ophthalmol.* **79**, 759 (1975)
14. Feinstein, A.R.: *Clinical biostatistics*. St. Louis: C.V. Mosby Company 1977
15. Fincke, M.: *Arzneimittelprüfung. Strafbare Versuchsmethoden*. Heidelberg, Karlsruhe: C.F. Müller Juristischer Verlag, 1977
16. Fischer, M., Hamelmann, H.: Dehnung oder Sphincterotomie als Behandlung der primär-chronischen Analfissur. *Chirurg* **49**, 215 (1978)
17. Fisher, R.A.: *Statistical methods for research workers*. Edinburgh: Oliver and Boyd 1950
18. Forrest, A.P.M.: The treatment of duodenal ulcer by gastroenterostomy, gastroenterostomy and vagotomy and partial gastrectomy. *Gastroenterologia (Basel)* **89**, 307 (1958)
19. Galen, R.S., Gambino, S.R.: *Beyond normality: The predictive value and efficiency of medical diagnoses*. p. 5, 6. New York: J. Wiley & Sons 1975
20. Gehan, E.A., Freireich, E.J.: Non-randomized controls in cancer clinical trials. *New Engl. J. Med.* **290**, 198 (1974)
21. Gesetz zur Neuordnung des Arzneimittelrechts, BGBl. I 2445 (1976)
22. Gill, P.W., Leaper, D.J., Stainland, J.R., Horrocks, J.C., De Dombal, F.T.: Observer variation in clinical diagnosis. A computer-aided assessment of its abdominal pain. *Methods of Information in Medicine* **12**, 108 (1973)
23. Goligher, J.C., Pulvertaft, C.N., Watkinson, G.: Controlled trial of vagotomy and gastroenterostomy, vagotomy and antrectomy and subtotal gastrectomy in elective treatment of duodenal ulcer: interim report. *Brit. med. J.* **1964I**, 455
24. Greenberg, B.G.: Why randomize? *Biometrics* **7**, 309 (1951)
25. Hall, R., Horrocks, J.C., Clamp, S.E., DeDombal, F.T.: Observer variation in assessment of results of surgery for peptic ulceration. *Brit. med. J.* **1976I**, 814
26. Hill, A.B.: *Principles of medical statistics*. p. 34. London: The Lancet LTD 1966
27. Immich, H.: *Medizinische Statistik. Eine Einführungsvorlesung*. S. 38. Stuttgart, New York: F.K. Schattauer 1974
28. Immich, H.: Grundsätzliches zur Planung und Durchführung klinischer Versuche. *Med. Welt* **27**, 1517 (1969)
29. Johnston, D.: Persönliche Mitteilung
30. Koller, S.: Angriff auf den Fortschritt der Medizin. Behauptung der Strafbarkeit kontrollierter klinischer Therapieversuche. *Fortschr. Med.* **95**, 2570 (1977)
31. Kusche, J., Stahlknecht, C.D., Lorenz, W., Reichert, G., Richter, H.: Diamine oxidase activity and histamine release in dogs following acute mesenteric artery occlusion. *Agents Actions* **7**, 81 (1977)
32. Largaier, F., Säuberli, H.: The role of gastrin in duodenal ulcer surgery. *Klin. Wschr.* **54**, 957 (1976)
33. Lasagna, L.: Randomized clinical trials. *New Engl. J. Med.* **295**, 1086 (1976)
34. Liedberg, G., Oscarson, J.: Selective proximal vagotomy – short-term follow-up of 80 patients. *Scand. J. Gastroent.* **8**, Suppl. 20, 12 (1973)
35. Lorenz, W., Hamelmann, H., Troidl, H.: Marburg experiment on surgical research: A five-year's experience on the cooperation between clinical and theoretical surgeons. *Klin. Wschr.* **54**, 927 (1976)
36. Lorenz, W., Doenicke, A., Dittmann, I., Hug, P., Schwarz, B.: Anaphylactoide Reaktionen von Blutersatzmitteln beim Menschen. *Anaesthesist* **26**, 644 (1977)
37. Lorenz, W., Rohde, H.: Probleme bei der Durchführung einer prospektiven kontrollierten Studie über chirurgische Behandlungsmethoden. *Klin. Wschr.* **51**, 457 (1973)
38. Lorenz, W., Doenicke, A.: Histamine release in clinical conditions. *Mount Sinai J. Med.* **45**, 297 (1978)
39. Mainland, D.: Some statistical problems in the planning and performance of clinical trials. In: *Pharmacologic techniques in drug evaluation*. Modine, J.H. and Siegler, P.E. (eds.), p. 27. Chicago 1964
40. Martini, P., Oberhofer, G., Welte, E.: *Methodenlehre der therapeutisch klinischen Forschung*. Berlin, Heidelberg, New York: Springer 1968
41. Mc Kneally, M.F., Maver, C., Kausel, H.W.: Regional immunotherapy of lung cancer with intrapleural B.C.G., *Lancet* **1976I**, 377
42. M.R.C. Multicentre trial of glucagon and aprotinin. *Lancet* **1977II**, 632
43. Peto, R., Pike, M.C., Armitage, P., Breslow, N.E., Cox, D.R., Howard, S.V., Mantel, N., Mc Pherson, K., Peto, J., Smith, P.G.: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. *Brit. J. Cancer* **34**, 585 (1976)
44. Pflanz, M.: Versuchspersonen und Kontrollgruppen – psychologische und epidemiologische Gesichtspunkte. *Med. Welt* **11**, 682 (1968)
45. Popper, K.R.: Von den Quellen unseres Wissens und unserer Unwissenheit. In: *Mannheimer Forum 1975/76*. H.v. Ditfurth (Hrsg.), S. 9. Mannheim: Boehringer
46. Popper, K.R.: *Objektive Erkenntnis. Ein evolutionärer Entwurf*. S. 44. Hamburg: Hoffmann und Campe: 1973
47. Price, W.E., Guzzle, J.E., Postlethwait, R.W., Johnson, W.D., Grabichi, P.: Results of operation for duodenal ulcer. *Surg. Gynec. Obstet.* **131**, 233 (1970)
48. Rahlfs, V.R.: Über die Beurteilung klinisch-therapeutischer Untersuchungen. *Med. Welt* **21**, 2135 (1970)
49. Rahlfs, V.R.: *Biometrie und klinisch-pharmakologische Forschung*. Intern. J. clin. Pharmakol. **5**, 1 (1971)
50. Reichenbach, H.: *Der Aufstieg der wissenschaftlichen Philosophie*. S. 45. Berlin: F.A. Herbig 1955
51. Rohde, H., Troidl, H., Lorenz, W.: Systematic follow-up: A concept for evaluation of operative results in duodenal ulcer patients. *Klin. Wschr.* **55**, 925 (1977)
52. Rohde, H., Troidl, H., Lorenz, W.: Reproduzierbarkeit und Spezifität des Insulininjektions- und Insulininfusionstestes zur Prüfung einer kompletten Vagotomie. *Langenbecks Arch. Chir. Suppl.*, 192 (1976)
53. Rundtischgespräch (Leiter: W. Lorenz): Die prospektive Studie, Methode zur Ermittlung des Therapieerfolges. Kongreßband der 95. Tagung der Deutschen Gesellschaft für Chirurgie. *Langenbecks Arch. klin. Chir.* (in Druck)
54. Sachs, L.: *Angewandte Statistik. Planung und Auswertung – Methoden und Modelle*. S. 26. Berlin, Heidelberg, New York: Springer 1974
55. Schöning, B., Koch, H.: Pathergiequote verschiedener Plasma-substitute an Haut und Respirationstrakt orthopädischer Patienten. *Anaesthesist* **24**, 507 (1975)
56. Schreiber, H.W.: *Chirurgische Perspektiven*. *Langenbecks Arch. Chir.* **345**, 477 (1977)
57. Seidel, W., Troidl, H., Lorenz, W., Rohde, H., Richter, H., Drews, H., Hamelmann, H.: Eine prospektive, kontrollierte Studie zur selektiven Vagotomie beim chronischen Duodenalulkus: Frühergebnisse mit einer standardisierten Operationsauswahl und Operationstechnik. *Klin. Wschr.* **51**, 477 (1973)
58. Snedecor, G.W., Cochran, W.G.: *Statistical methods*. 6. Ausgabe. The Iowa State University Press, Ames, Iowa 1967

59. Sturdevant, R.A.L.: How should results of controlled trials affect clinical practice? *Gastroenterology* **73**, 1179 (1977)
60. Taves, D.R.: Minimization: A new method of assigning patients to treatment and control groups. *Clin. Pharmacol. Ther.* **15**, 433 (1974)
61. Trapnell, J.E., Rigby, C.C., Talbot, C.H., Duncan, E.H.L.: A controlled trial of Trasylol in the treatment of acute pancreatitis. *Brit. J. Surg.* **61**, 177 (1974)
62. Troidl, H., Lorenz, W., Hegewald, C., Rohde, H., Acker, G., Seidel, W.: Präzision von Magensekretionstesten bei der chirurgischen Therapie des Ulkusleidens. *Langenbecks Arch. Chir., Suppl. Chir. Forum* **235** (1973)
63. Troidl, H., Lorenz, W., Rohde, H., Fischer, M., Hamelmann, H.: Was ist gesichert in der Behandlung der Ulkuserkrankheit durch Vagotomie? *Internist* **16**, 575 (1975)
64. Troidl, H., Lorenz, W., Rohde, H., Häfner, G., Ronzheimer, M.: Histamine and peptic ulcer: a prospective study of mucosal histamine concentration in duodenal ulcer patients and in control subjects suffering from various gastrointestinal diseases. *Klin. Wschr.* **54**, 947 (1976)
65. Troidl, H., Rohde, H., Lorenz, W., Häfner, G., Hamelmann, H.: Effect of selective gastric vagotomy on histamine concentration in gastric mucosa of patients with duodenal ulcer. *Brit. J. Surg.* **65**, 10 (1978)
66. Troidl, H., Lorenz, W., Rohde, H., Fischer, M., Vestweber, K.H., Hamelmann, H.: Pathophysiologie, Diagnostik und Operationsvorbereitung bei der benignen Magenausgangsstenose: Eine prospektive Studie an 209 Patienten mit peptischem Ulkus. In: *Komplikationen beim chronischen peptischen Ulkus* (Ed. Häring). Stuttgart: Georg Thieme (in Druck)
67. Troidl, H.: Einrichtung einer systematischen Kontrolluntersuchungsklinik. Kongreßband der 95. Tagung der Deutschen Gesellschaft für Chirurgie 1978, *Langenbecks Arch. klin. Chir.* (in Druck)
68. Warren, W.D.: Controlled clinical research. Opportunities and problems for the surgeon. *Amer. J. Surg.* **127**, 3 (1974)
69. Weinstein, M.C.: Allocation of subjects in medical experiments. *New Engl. J. Med.* **291**, 1278 (1974)
70. Yates, F.: Some examples of biased sampling. *Annals of Eugenics* **6**, 202 (1935)

Eingegangen am 17. Juli 1978

Angenommen am 22. Januar 1979

Prof. Dr. W. Lorenz  
Abteilung für experimentelle Chirurgie  
und pathologische Biochemie  
an der Chirurgischen Univ.-Klinik  
Robert-Koch-Straße 8  
D-3550 Marburg (Lahn)  
Bundesrepublik Deutschland