

Aus dem Institut für Medizinische Informationsverarbeitung, Statistik und Biomathematik, München

Versuchsplanung in Phase I

Von K. Überla

Zusammenfassung: Klinische Pharmakologen arbeiten vor allem in der Phase I. Hier liegt ihr Hauptinteresse, gewissermaßen ihr Heiligtum, in dem der Methodiker wenig zu tun hat. Im folgenden wird auf einige Probleme und Möglichkeiten in der Phase I eingegangen.

Die Formalisierung der Phase I ist nicht weit fortgeschritten. Häufig hört man die Meinung, daß eine Versuchsplanung im Sinne der Statistik in dieser Phase nicht nötig sei. Gerade das Gegenteil ist der Fall. Wenn man Menschen zum erstenmal mit einem neuen Stoff in Berührung bringt, ist es besonders nötig, sich mit der Versuchsplanung im wissenschaftlichen Sinne zu beschäftigen.

Zunächst sollen einige allgemeine Bemerkungen gemacht und der methodische Hintergrund des Systems formal beschrieben werden. Dann werden einige Gesichtspunkte bei der Versuchsplanung an einem Menschen — dem ersten, der mit einer neuen Substanz in Berührung kommt — dargestellt. Im dritten Teil werden lateinische Quadrate behandelt, die bei 2–7 Probanden in der Phase I das wichtigste Instrument der Versuchsplanung sind.

1. Allgemeine Beschreibung der Phase I als System

Die Ziele der Beteiligten sind bei den Phase-I-Untersuchungen unterschiedlich: Der Verbraucher (Patient) erwartet bessere und wirksamere Medikamente für die häufigsten und wichtigsten Krankheiten; neue Präparate sollen schnell wirksam sein und keine Nebeneffekte aufweisen.

Die Versuchspersonen erwarten, daß sie zunächst durch den Versuch selbst nicht zu Schaden kommen (Schmerzen, mögliche Nebenwirkungen usw.). Die Grundrechte sollen nicht verletzt werden. Ferner wird ein gewisser Vorteil durch die Teilnahme an solchen Versuchen (gewissenhafte Untersuchung und Behandlung, Honorar usw.) gesehen.

Die Ziele des Pharma-Unternehmens sind wieder andere: Toleranz und Effektivität nachzuweisen, die gesetzlichen und wissenschaftlichen Voraussetzungen für die Phase II zu schaffen und kostengünstig zu bleiben.

Die Ziele des klinischen Versuchsleiters sind, eine neue und wirksame Substanz zu finden, sowie als Wissenschaftler erfolgreich zu sein.

Vom allgemein wissenschaftlichen Standpunkt aus sind wieder andere Ziele gegeben: Die Toleranz nachzuweisen, die Effektivität zu zeigen (Pharmakodynamik), die Dosis zu etablieren, alle toxischen und Nebeneffekte zu überwachen, die pharmakokinetischen Effekte zu studieren und die Bioverfügbarkeit zu studieren.

Die Ziele der Beteiligten sind widersprüchlich und kaum auf einen Nenner zu bringen. Bei jedem Versuch der Phase I wird bewußt oder unbewußt eine Gewichtung solcher Ziele vorgenommen. Die Schwierigkeiten in der Phase I liegen häufig darin, daß der jeweilige Zielmix für eine bestimmte Studie nicht genau und explizit, sondern nur implizit definiert ist.

Das Verhältnis Proband — Untersucher ist der Ausgangspunkt jeder Systembeschreibung (Abb. 1). Eine neue Substanz wird das erstmal einem Menschen gegeben. Der Un-

Summary: Experimental Design in Phase I

Clinical pharmacology has its domain during phase I. The formalisation of phase I is not very advanced. Sometimes it is argued that the design of experiments during that phase is unnecessary. Just the opposite is true. Trying a new substance for the first time in man poses a special necessity for the design of experiments in a scientific context.

Phase I is described in a formal way. The design of experiments with a single subject — the first one to be exposed to a new substance — is described. Finally Latin squares are dealt with as a useful tool for phase I.

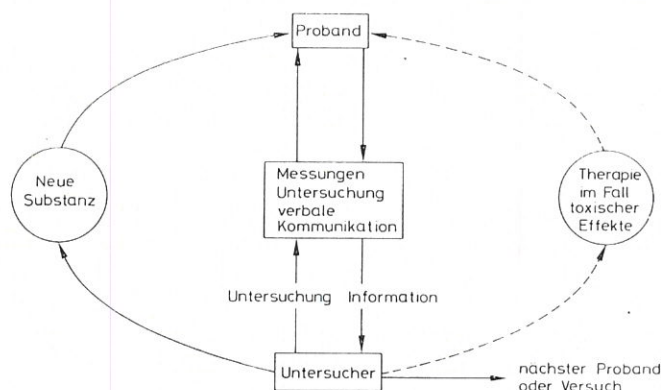


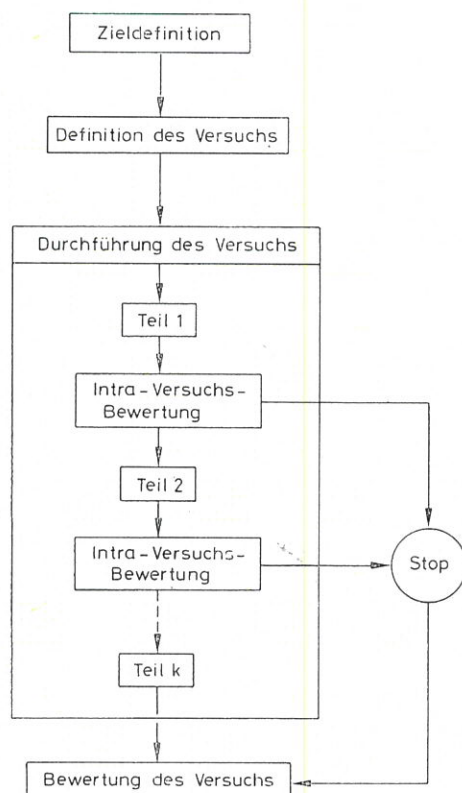
Abb. 1: Beziehung zwischen Proband und Untersucher.

tersuchte steht in intensiver Beziehung zum Arzt. Messungen verschiedenster Art werden durchgeführt, es findet eine verbale Kommunikation statt, und zwar in beiden Richtungen. Tritt bei dieser Informationssammlung oder Untersuchung ein gefährlicher Zustand auf, so ist im Hintergrund eine Therapie verfügbar im Fall toxischer Effekte. Die Abb. beschreibt die Beziehung zwischen Untersucher und Probanden mit ihren Möglichkeiten sehr global.

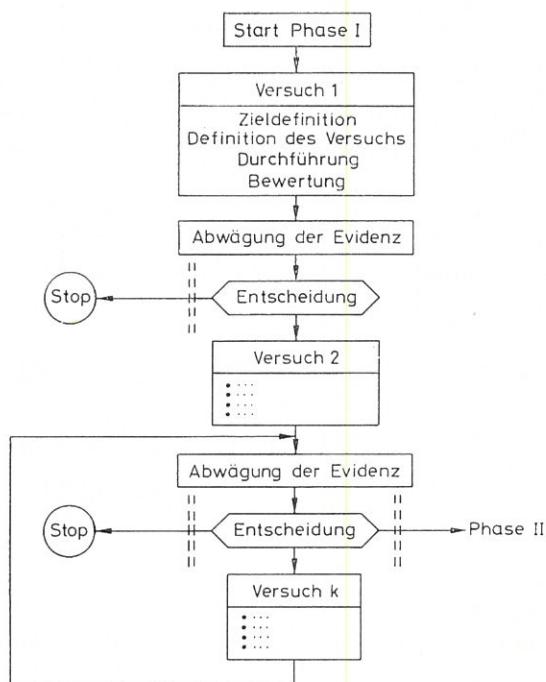
Jede Phase I läßt sich in mehrere einzelne Versuche aufteilen, z. B. zur Bestimmung der Toleranz, der Wirksamkeit, der therapeutisch zweckmäßigen Dosis, der Pharmakokinetik oder der Bioverfügbarkeit. Jeder einzelne Versuch der Phase I läßt sich formal beschreiben, wie dies im Schema 1 dargestellt ist.

Zunächst werden die Ziele definiert, dann der Versuch. Schließlich wird der Versuch durchgeführt, und zwar in einzelnen Schritten, z. B. die Aufnahme jeweils eines Probanden oder einer kleinen Gruppe von Probanden. Bei jedem dieser Teile muß es möglich sein, den Versuch abubrechen. Am Ende ist der gesamte Versuch auszuwerten und zu bewerten.

Die gesamte Phase I läßt sich als System aufeinanderfolgender Einzelversuche beschreiben; ein dynamisches System, wie Schema 2 zeigt. Zunächst wird die Phase I gestartet, wer immer diese Entscheidung trifft. Dann läuft der erste Versuch ab, der aufgeteilt ist in einzelne Schritte: Definition der Ziele, Definition des Versuchs, Durchführung und Be-



Schema 1: Grundschrte eines Versuches der Phase I.



Schema 2: Systembeschreibung von Phase I-Versuchen.

wertung. Nach jedem Versuch schließt sich eine Abwägung aller nunmehr vorliegenden Evidenzen an, z. B. was von anderen Gruppen gefunden wurde, der allgemeine Fortschritt oder die Veränderung der Situation, die es z. B. unzweckmäßig erscheinen lassen kann, die Entwicklung weiter zu verfolgen. Daraus folgt der Schluß, ob die Phase I weitergeführt werden soll.

So folgt Versuch nach Versuch in der Phase I. Das Systemverhalten und besonders der Ausgang der Phase I wird im wesentlichen bestimmt durch die Bedingungen, die erfüllt sein müssen, um zur Phase II überzugehen, sowie durch die Bedingungen, die zum Stopp der Phase I führen. Eine Sy-

stemanalyse zeigt also, daß die Phase I ein dynamisches System aufeinanderfolgender Einzelversuche ist, das in seinem Systemverhalten im wesentlichen abhängt von solchen „stop and go-Bedingungen“. Diese formale Systembeschreibung der Phase I ist der nötige Hintergrund für die Versuchsplanung.

2. Versuchsplanung bei einem Patienten

Es gibt keine geschlossene Theorie der Versuchsplanung in der Phase I. Die bekannten Instrumente der Phasen II und III können mehr oder weniger sensibel angewandt werden. Die Situation der ersten Anwendung einer Substanz am Menschen ist so einzigartig, daß es nötig ist, irgendeine Versuchsplanung anzuwenden, auch wenn die Versuchsplanung allgemein verlangt, daß eine Zufallszuteilung erfolgt — was hier nicht möglich ist —, daß eine Wiederholung erfolgt, was hier ebenfalls nur eingeschränkt möglich ist.

Wenn wir nur eine Versuchsperson haben, können wir nicht auf die Grundgesamtheit aller Versuchspersonen zurückschließen. Unsere Schlüsse sind begrenzt auf diese eine Versuchsperson. Dafür können wir aber alle bekannten statistischen Techniken einsetzen. Lediglich die Übertragung auf andere Personen geht nicht mehr mit Hilfe statistischer Methoden, sondern nur auf dem Weg des Analogieschlusses. Dies ist jedoch ausreichend, um weitere Versuche der Phase I anzusetzen oder abubrechen.

Die wichtigste Kontrolle in dieser Situation ist der Zustand des Probanden vor Anwendung der Substanz. Wenn man ein Merkmal vor und nach der Gabe einer einzigen Dosis bei einem Probanden mißt, hat man den Unterschied zwischen diesen beiden Situationen, der möglicherweise auf die Substanz zurückgeführt werden muß. Natürlich kann dieser Unterschied auch Zufallsschwankungen oder Tagesrhythmen enthalten. Das Problem einer einzigen Differenz zwischen vorher und nachher ist jedoch, daß keine Varianz zur Verfügung steht, an der wir die Größe dieser Differenz messen können. Im allgemeinen verwendet man in der Versuchsplanung die Varianz zwischen den Personen, um die Größe des Effekts daran zu messen. Dies ist hier nicht möglich, da wir nur eine Person haben.

Der einzige Ausweg, der sich anbietet, ist die wiederholte Messung desselben Merkmals vor und nach Gabe der Substanz. Diese Wiederholung ist ein wirksames Instrument bei einer Person. Im einfachsten Fall haben wir eine Situation wie in Abb. 2.

Es werden k-Messungen eines Merkmals an einer Person vor der Gabe und k-Messungen nach der Gabe durch-

Messungen vor Medikament	Messungen nach Medikament
1	1
2	2
•	•
•	•
k	k

Varianzanalyse

	df	ssq	msq	F
Behandlung	1			
Wiederholung	(k - 1) 2			
Gesamt	2k - 1			

$$x_{ij} = \mu + \eta_i + \epsilon_{ij}$$

Abb. 2: Einfacher Versuchsplan für eine Person.

geführt. Wendet man auf diese Situation eine Varianzanalyse oder einen T-Test an, so ist dies erlaubt. Man erhält die klassische Varianzanalysentafel, auch die Modellgleichung ist in Abb. 2 angegeben.

Der entscheidende Punkt ist die Bedeutung des Signifikanztests. Für den Fall, daß die Irrtumswahrscheinlichkeit klein genug ist, z. B. kleiner als 1%, kann man behaupten, daß man bei einer Person einen meßbaren Effekt gefunden hat. Wir haben die Genauigkeit der Messung verglichen mit der Größe der Veränderung. Man kann also sagen, daß die Genauigkeit der Meßmethode hinreichend groß war, um einen Unterschied zwischen den Zeitpunkten zu zeigen, und dies mit einer bestimmten Irrtumswahrscheinlichkeit. Ob dieser so „signifikante“ Unterschied auf die Substanz zurückgeht, oder auf andere Einflüsse, kann man mit dieser Versuchsanordnung nicht sagen. Aber man kann wesentlich mehr sagen, als dies ohne eine solche Aussage geschehen kann.

Nicht alle Statistiker würden mit dieser Darlegung eingehen. Doch ist der statistische Test bei einer Person unter bestimmten Bedingungen — wie sie hier beschrieben worden sind — korrekt, wenn man die Bedeutung des Wortes „signifikant“ entsprechend erläutert. Nach derselben Philosophie lassen sich alle anderen Tests für analoge Situationen auf eine einzige Person übertragen. Der entscheidende Punkt ist freilich, durch genügende Wiederholungszahl vor und nach der Messung eine Varianzschätzung für den Fehler der Meßmethodik zu finden. Da dieser Fehler der Meßmethodik im allgemeinen klein sind, erhält man durch derartig modifizierte Signifikanzaussagen relativ gute Chancen für Sachaussagen. Dies ist besser als nichts: Die Aussage nämlich, der beobachtete Unterschied ist größer als unser Meßfehler. Behält man die Nullhypothese bei, kann man nur sagen, daß der Meßfehler vermutlich in der gleichen Größenordnung ist wie der Effekt — daß ein Effekt bei einer Person empirisch also nicht gezeigt werden konnte.

3. Lateinische Quadrate

Bei mehr als einem Patienten, z. B. 2—7 Probanden, sind lateinische Quadrate das geeignete Instrument der Versuchsplanung für Phase I. Das Prinzip ist in der Abb. 3 erläutert.

Man hat z. B. 3 Probanden, die an 3 Tagen mit 3 Dosen untersucht werden. Im oberen Teil der Abb. ist dargestellt, wie sich die neun Elemente (3 Probanden an 3 Tagen) im ungeordneten Zustand darstellen. In der Mitte ist die Situation wiedergegeben, wenn man die Elemente nach Tagen

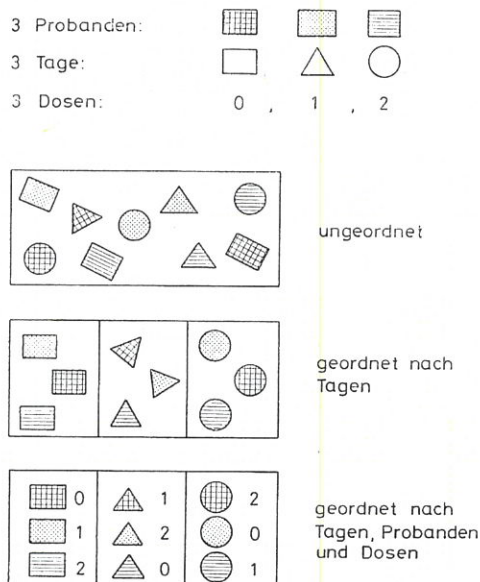


Abb. 3: Prinzip der lateinischen Quadrate in Phase I.

ordnet. Im unteren Teil schließlich ist dargestellt, wenn man sowohl nach Tagen als auch nach Probanden ordnet. Außerdem ist dann jedem der neuen Elemente eine und nur eine Dosis zugeordnet, und zwar so, daß in jeder Zeile und jeder Spalte jede Dosis einmal und nur einmal vorkommt.

Diese Anordnung ist entscheidend. Das lateinische Quadrat schließt aus, daß man Wechselwirkungen zwischen Probanden, Dosen oder Tagen erfassen kann. Im allgemeinen kann ein Fehlen derartiger Wechselwirkungen zunächst unterstellt werden.

„Die Blockung“ nach Zeilen und Spalten ermöglicht es, den Versuchsfehler klein zu halten. Die Zufallszuteilung ist bei dieser Versuchsanordnung eingeschränkt, die Chancen aber, ein signifikantes Ergebnis für die Dosen zu bekommen, sind besonders hoch.

Diesem Vorteil stehen zwei Nachteile gegenüber:

1. Die Anzahl der Probanden, der Tage und Dosen muß gleich sein, d. h. die Versuchsanordnung ist weniger flexibel.
2. Wechselwirkungen werden als nicht existent angenommen, die drei Faktoren sind orthogonal.

Das Schema für ein solches lateinisches Quadrat in der Phase I ist in Abb. 4 wiedergegeben. Es läßt sich standardmäßig mit einer Varianzanalyse auswerten, wie dies unten angedeutet wurde. Die Varianz der Dosen wird durch die Fehlervarianz geteilt, und man erhält den kritischen F-Wert, wie der Pfeil dies andeutet.

	A	B	C	
I	0	1	2	A, B, C: Personen I, II, III: Tage 0, 1, 2: Dosen
II	1	2	0	
III	2	0	1	

Varianzanalyse

	df	ssq	msq	F
Personen	2			
Tage	2			
Dosen	2			
Fehler	2			
Gesamt	8			

Abb. 4: Lateinisches Quadrat.

Für lateinische Quadrate verschiedener Dimension gibt es zahlreiche Realisierungsmöglichkeiten. Tab. 1 zeigt, wieviele lateinische Quadrate es bei den Dimensionen 3, 4, 5 und 6 gibt. Von einer Standardform spricht man, wenn ein lateinisches Quadrat so angeordnet ist, daß es in der ersten Zeile und in der ersten Spalte die Behandlungen in aufsteigender Reihe enthält. Bei 3×3 Quadraten gibt es eine solche Standardform, aus der sich durch Vertauschen der Zeilen und Spalten 12 verschiedene lateinische Quadrate ableiten lassen. Bei einem konkreten Experiment ist eine dieser 12 Möglichkeiten zufällig zu wählen. Wie Tab. 1 zeigt, gibt es bei 4×4 Quadraten 4 Standardformen mit je 144 Permutationen, d. h. insgesamt 576 verschiedene lateinische Quadrate. Bei 5 Dimensionen sind es schon über 160 000, bei 6 mehr als 812 000 000. Es ist nötig, aus diesen verschie-

Tab. 1: Anzahl verschiedener lateinischer Quadrate.

Dimension	Standardformen	$r! (r-1)!$	Gesamtzahl
3×3	1	12	12
4×4	4	144	576
5×5	56	2880	161 280
6×6	9408	86400	812 851 200

Tab. 2: Rechenschema lateinisches Quadrat (Ausgangsdaten).

	(3)	(1)	(4)	(6)	(2)	(5)	Zeilensumme
	21	14	11	28	18	10	102
	(2)	(6)	(3)	(5)	(1)	(4)	
	28	33	27	18	24	14	144
	(5)	(3)	(6)	(2)	(4)	(1)	
	12	31	28	30	17	20	138
	(4)	(2)	(5)	(1)	(3)	(6)	
	13	24	8	16	24	23	108
	(1)	(5)	(2)	(4)	(6)	(3)	
	8	11	20	22	24	17	102
	(6)	(4)	(1)	(3)	(5)	(2)	
	26	13	26	(24)	19	18	126
Spaltensumme	108	126	120	138	126	102	720
	(1)	(2)	(3)	(4)	(5)	(6)	
Beh.-Summe	108	138	144	90	78	162	720

Tab. 3: Varianzanalyse lateinisches Quadrat (Rechenschema).

$$A = 21^2 + \dots + 18^2 = 16\,026$$

$$C = \frac{1}{36} (720)^2 = 14\,400$$

$$R = 1/6 (102^2 + 144^2 + \dots) = 14\,688$$

$$K = 1/6 (108^2 + 126^2 + \dots) = 14\,544$$

$$T = 1/6 (108^2 + 138^2 + \dots) = 15\,312$$

$$r = 6$$

	FG	SAQ	MAQ	F
Zeilen	$r-1 = 5$	$R-C = 283$	57,6	
Spalten	$r-1 = 5$	$K-C = 144$	28,8	
Behandl.	$r-1 = 5$	$T-C = 912$	182,4	
Fehler	$(r-1)(r-2) = 20$	$(\text{subtr.}) = 284$	14,2	12,8
Gesamt	$r^2-1 = 35$	$A-C = 1628$		

denen lateinischen Quadraten jeweils eines zufällig auszuwählen. Dazu gibt es Tabellen, die die Standardformen enthalten, von denen man ausgehen kann. Das Rechenschema

	A	B	C	
I	00	01	02	
	≡≡	≡≡	≡≡	A, B, C : Personen
II	01	02	00	
	≡≡	≡≡	≡≡	I, II, III : Tage
III	02	00	01	
	≡≡	≡≡	≡≡	0, 1, 2 : Dosen
				≡≡ : Wiederholungen

Varianzanalyse für eine Dosis

	df	ssq	msq	F
Personen	2			
Tage	2			
Dosen	1			
Fehler	18			
Gesamt	23			

Abb. 5: Lateinisches Quadrat, kombiniert mit Wiederholungen.

für lateinische Quadrate ist an einem Beispiel in Tab. 2 (Ausgangsdaten) und in Tab. 3 (Rechenschema) dargestellt. Anhand dieser Zahlen kann man das Vorgehen nachvollziehen.

Es soll auf eine mögliche Weiterentwicklung der lateinischen Quadrate für die Phase I hingewiesen werden: Es ist möglich, die lateinischen Quadrate zu verbinden mit der Wiederholung der Merkmalsmessung vor und nach Gabe der Dosis. Daraus würde sich die Versuchsanordnung in Abb. 5 ableiten.

Wir haben 3 Personen an 3 Tagen mit 3 Dosen. An jedem der neuen Felder messen wir vor und nach der Gabe des Medikaments, und zwar je 4mal. Wir haben dann $3 \times 3 \times 3 \times 2 \times 4 = 216$ Beobachtungen.

Der einfachste Weg der Auswertung ist, jede Dosis für sich zu betrachten mit einer Varianzanalyse, deren Schema im unteren Teil der Abbildung angegeben ist. Bei der Dosis Null dürfte kein Effekt auftreten. Es gibt andere Ansätze zur Auswertung eines solchen Materials. Die Kombination der lateinischen Quadrate mit der Idee der wiederholten Messung vor und nach einer Behandlung bei jeder Gelegenheit öffnet einen neuen methodischen Weg für die Versuchsplanung und statistische Auswertung der Phase I. Dabei handelt es sich jedoch noch um Entwicklungsarbeit.

Die statistische Auswertung von Experimenten der Phase I darf nicht auf die Testung weniger Hypothesen begrenzt werden. Die Hypothesentestung ist nötig, sie ist jedoch abzugrenzen von der Hypothesengenerierung. Bei der Wichtigkeit eines gut dokumentierten Materials der Phase I ist eine Hypothesengenerierung, die als solche deutlich abgesetzt ist vom geplanten Test, immer nötig. Für die Hypothesengenerierung kann die Deskription aller beobachteten Merkmale in Untergruppen und kompliziertere statistische Verfahren angesetzt werden. Das statistische Auspressen eines Materials der Phase I mit allen verfügbaren Mitteln ist dann gerechtfertigt, wenn es als solches gekennzeichnet wird.

Die Anzahl der Phase I-Studien wird immer begrenzt bleiben aus den verschiedensten Gründen. Es ist entscheidend wichtig, menschliche und finanzielle Ressourcen nicht zu vergeuden in dieser Phase. Auf lange Sicht gesehen sind Intuition und individuelle Erfahrung nicht die einzigen Mittel für eine effektive Planung der Phase I. Formale Instrumente müssen hinzutreten, von denen einige aufgezeigt wurden.

Die Natur arbeitet über Redundanz und Zufall. Dies trifft auch zu für den faszinierenden Prozeß der Entwicklung neuer Medikamente — ein Prozeß, der sich erst in einem frühen formalen Entwicklungsstand befindet. Wir haben noch keine Wissenschaft in einem engeren Sinn, diesen Prozeß zu unterstützen. Wir sammeln in der Phase I nur Daten, durch trial and error, mit individueller Erfahrung und „guidelines“ wie bei der Ausübung einer Kunst.

In Zukunft wird es sehr viel mehr Wissenschaft in dieser Phase geben. Das Ziel, zumindest eine genaue Systembeschreibung der Phase I zu geben und daraus die Versuchsplanung abzuleiten, ist für den Methodiker vorgegeben. Es ist zu hoffen, daß in der Zusammenarbeit der verschiedenen Wissenschaften, zu der wir heute alle beitragen, sich die Kunst und das Glück bei der Entwicklung neuer Medikamente in Strategien überführen läßt mit höherer Erfolgsaussicht.

Anschr. d. Verf.: Prof. Dr. med. K. Überla, Institut für Medizinische Informationsverarbeitung, Statistik und Biomathematik, Marchioninistr. 15, 8000 München 70

Verantwortlich für die Redaktion: Prof. Dr. Hans Georg Classen, Stellvertreter: V. Schramm. Redaktions-Sekretariat: W. Frey. Verlag: Editio Cantor, D-7960 Aulendorf/Württ., Ruf: Aulendorf (0 75 25) 4 31—4 33. Druck: Vereinigte Buchdruckereien, August Sandmaier & Sohn, D-7952 Bad Buchau. Alle Rechte vorbehalten.

Bezugsbedingungen: Die Zeitschrift erscheint monatlich. Pro Jahr 2 Bände (I u. II). 6 Hefte ergeben einen Band (keine Berechnung zusätzlicher Sonderhefte). Die Zeitschrift kann vom Verlag oder durch eine Buchhandlung bezogen werden. Der Bezugspreis beträgt im Abonnement pro Band 130,— DM inklusive MwSt. zuzügl. Versandkosten. Auslandsabonnement: 140,— DM zuzügl. Versandkosten. Das Einzelheft kostet 26,— DM inklusive MwSt. zuzügl. Versandkosten. Kostenlose Probehefte liefert der Verlag auf Anforderung.

Printed in W. Germany — ISSN 0004—4172