

Der lange Weg zur validierten Ersatzmethode

Thomas Hartung* und Horst Spielmann**

*Biochemische Pharmakologie der Universität Konstanz, D-Konstanz

**ZEBET (Zentralstelle zur Erfassung und Bewertung von Ersatz- und Ergänzungsmethoden zum Tierversuch), D-Berlin

Zusammenfassung

Es wird versucht, neuere Konzepte zu einem besseren Vorgehen bei Validierungsstudien zusammenzufassen. Im Licht der bisherigen Versuche zur Validierung von Ersatz- und Ergänzungsmethoden zum Tierversuch drängt sich ein Überdenken des aufwendigen Prozesses auf. Auf Grund der verschiedenen Anwendungen erfordern toxikologische Tests (Verbraucherschutz) und pharmakologische Methoden (Identifizierung von Wirksubstanzen) unterschiedliche Validierungsstrategien. Es werden deshalb Vorschläge zur Verbesserung der Entwicklung, Evaluierung, Prävalidierung und Validierung eines Tests sowie der endgültigen Bewertung der Ergebnisse präsentiert.

Summary: The sophisticated process of validation.

Current concepts on the efficient performance of validation studies are summarized. From the experience of recent attempts to validate alternative methods to animal experiments, the difficult procedure of validation should be refined. The applications of toxicological tests (safety testing) and pharmacological methods (identification of putative drugs) require different validation strategies. Suggestions as to an improved design of test development, evaluation, prevalidation, definitive validation as well as final assessment of the result are presented.

Keywords: alternative methods, validation strategies, safety testing, evaluation, prevalidation, definitive validation, final assessment

Die Validierung einer Ersatzmethode im Vergleich zum Tierversuch soll ihre Reproduzierbarkeit und Relevanz ermitteln, um eine abschließende Beurteilung zu ermöglichen. Sie ist entscheidend für die Akzeptanz von möglichen Anwendern und unabdingbar für Behörden. Im Spannungsfeld von Politik und Wissenschaft darf nicht vergessen werden, daß es sich um einen primär wissenschaftlich rationalen Prozess handelt. Der Begriff der Validierung wird für sehr unterschiedliche Ansätze verwendet. Er umfaßt sowohl den Vergleich verschiedener Methoden in einer Einrichtung oder einem Unternehmen für interne Zwecke, als auch die kommerzielle Methodenentwicklung und das Bemühen um eine behördliche Anerkennung. In jedem Einzelfall werden sich daraus unterschiedliche Anforderungen an den Aufwand der Validierung ergeben. Zu klären bleibt jedesmal, ob überhaupt ein vollständiger Ersatz des jeweiligen Tierversuchs angestrebt werden soll, oder ob (zunächst) nur die Einbettung in ein hierarchisches Prüfverfahren versucht wird, d.h. als eine Art Vorfilter, durch den Tierversuche vermindert oder überflüssig werden. Es ist sogar vorstellbar, daß zunächst durch den Einsatz der Ersatzmethode parallel zum Tierversuch überhaupt erst eine Datenbasis erhoben werden kann, die das Potential der neuen tierversuchsfreien Methode erschließt.

Verschiedene Ansprüche an Ersatzmethoden in Pharmakologie und Toxikologie

Die Entwicklung und Sicherheitsprüfung von Arzneimitteln basiert heute wesentlich auf Tierversuchen. Bei der Etablierung von Ersatzmethoden werden Pharmakologie und Toxikologie gern in einem Atemzug genannt. Bei genauerer Analyse ergeben sich jedoch interessante Unterschiede, die wesentlichen Einfluß auf die Validierung der Ersatzmethode haben: Toxikologische Prüfungen

sind Sicherheitstests zum Schutz des Verbrauchers, d.h. es soll das Risiko schädigender Wirkungen selbst bei extremen Expositionen oder sogar bei unsachgemäßer Verwendung einer Substanz abgeschätzt werden. Dazu wird das gesunde Versuchstier so lange steigenden Mengen der Versuchssubstanz ausgesetzt, bis toxische Wirkungen auftreten. Wenn die Mechanismen von Aufnahme und Verteilung im Organismus und der Um- und Abbau der Substanz bekannt oder abschätzbar sind, scheint es auch möglich, die toxische Wirkung auf besonders empfindliche Gewebe direkt in vitro zu prüfen. Es kann z.B. die lebertoxische Wirkung einer Substanz an Leberzellen studiert werden, wenn diese Zellen einigermaßen in demselben Zustand sind wie im Organismus. Ziel muß es bei toxikologischen Prüfungen sein, alle potentiell gefährlichen Substanzen zu identifizieren, d.h. es dürfen vor allem keine "falsch negativen" Testergebnisse auftreten. Für die üblichen toxikologischen Sicherheitsprü-

98



fungen von Stoffen im Bereich der Kosmetika, Waschmittel, Verpakkungsmaterialien, Nahrungsmittelzusätzen und Pharmaka gibt es deshalb standardisierte Protokolle, deren Durchführung zum Teil gesetzlich vorgeschrieben ist.

Die Wirkstoffindung in der pharmakologischen Forschung hingegen basiert auf experimentellen Modellen zu Krankheitsbildern, in denen Wirksubstanzen mit dem Ziel getestet werden, sie zur Heilung von Krankheiten einzusetzen. Deshalb ist der Zugang in der Regel tierexperimentell, d.h. ausgehend von Beobachtungen des Arztes am Patienten wird versucht, eine entsprechende Erkrankung am Versuchstier auszulösen und anschließend zu heilen. Dieses Abbild der menschlichen Erkrankung wird zunächst an der Ähnlichkeit der hervorgerufenen Symptome bemessen. Der nächste Schritt ist dann die experimentelle Charakterisierung der "Modellerkrankung" des Versuchstiers. Insbesondere interessieren hierbei die Mechanismen der Schädigung. Daraus werden dann Hypothesen über die Entstehung der Erkrankung beim Menschen abgeleitet, die beim Patienten überprüft werden können. Ein gutes pharmakologisches Modell ahmt also nicht nur wesentliche Pathomechanismen der Krankheitsentstehung beim Menschen nach, sondern bewährt sich in der Praxis vor allem durch die Richtigkeit der Auswahl von Wirkstoffen für die weitere Prüfung bis hin zur Klinik. Im Gegensatz zur Toxikologie sind die Versuche deshalb weniger standardisiert. Ziel ist es vielmehr, durch eine möglichst ausgeklügelte Planung der Versuche die Wahrscheinlichkeit zu erhöhen, eine geeignete Wirksubstanz zu finden. Eine ständige Anpassung an den Stand der Erkenntnisse über Pathophysiologie und die technischen Möglichkeiten ist dazu wichtig. Dementsprechend gibt es nur wenige allgemein durchgeführte und praktisch keine gesetzlich vorgeschriebenen Tierversuche in diesem Bereich. Eine Ausnahme ist allerdings die Qualitätsprüfung von Chargen eines eingeführten Medikamentes, die z.B. das Arzneibuch vorschreibt. Das Problem des Pharmakologen bei der Wirkstoffindung sind aber somit weniger falsch positive Effekte, sondern es besteht eine Gefahr darin, daß eine wirkungsvolle Substanz zu früh ausgemustert wird.

Die Testentwicklung (Test Development)

Am Anfang aller Überlegungen zum Ersatz eines Tierversuchs sollte natürlich die Frage stehen, ob dieser überhaupt sinnvoll und notwendig ist (Hartung et al., 1995). Erst danach sollte die Entwicklung der Ersatzmethode anstehen. Es gilt dann zunächst zu zeigen, daß die jeweilige tierversuchsfreie Methode das adäquate Modell eines gültigen Tiermodells ist. Der Ausgangspunkt ist dabei die klinische und tierexperimentelle Erfahrung zum entsprechenden Krankheitsbild bzw. zur toxischen Wirkung. Orientierend an den daraus gewonnenen Vorstellungen findet als erstes die Nachahmung z.B. im perfundierten Organ oder der Zellkultur statt, d.h. bei der Arzneimittelentwicklung sollte eine experimentell erzeugte Schädigung in diesen Systemen durch eine Prüfsubstanz verhindert werden. Diese Aufbauarbeit orientiert sich zunächst nur an der erreichbaren offensichtlichen Ähnlichkeit der Ersatzmethode zur experimentellen Schädigung im Tierversuch. Als Maßstab muß hier die Plausibilität des Modellaufbaus dienen. Wichtig ist, daß die Testentwicklung während des gesamten nachfolgenden Validierungsprozesses nicht als abgeschlossen angesehen werden darf, um eine Überprüfung und gegebenenfalls Verbesserung möglichst früh zu ermöglichen (Vereinfachung und Standardisierung der Handhabung, Auswertung und ggf. Modifikation der Methode). Die Optimierung des Systems muß sowohl bezüglich des biologischen Systems (z.B. der Qualität der Zellkultur), des Testprotokolls (Handhabbarkeit, Festlegung von Vorinkubation, Exposition und Inkubation) als auch des verwendeten Meßparameters erfolgen.

Wichtig ist außerdem, daß jeder Test letztlich eine Hypothese prüft (z.B. "Die Substanz besitzt toxische Wirkungen auf eine gegebene Zielstruktur" oder "Die Substanz verhindert den Ablauf eines Pathomechanismus"). Diese Hypothese, also das, was die Methode leisten soll, muß für jeden Test möglichst früh festgelegt werden (hypothesis testing), um die Testentwicklung daran zu orientieren. Als Ersatzmethode zu einem Tierversuch muß sie geeignet sein, das Ergebnis eines konkreten Tierversuchs abzuschätzen (prediction model). Daraus ergibt sich auch die Anforderung, daß eine (toxikologische) Ersatzmethode auch dieselbe Sicherheit gewährleisten muß wie der zu ersetzende Tierversuch. Als vollständiger Ersatz kommt die tierversuchsfreie Methode damit nur in Frage, wenn dieselbe Qualität der Ergebnisse erreicht wird. Die Testentwicklung kann, solange dieses Ziel nicht erreicht wird, nicht als abgeschlossen angesehen werden. Natürlich ist es aber möglich, einen Tierversuch partiell zu ersetzen, wenn die Ersatzmethode eine Vorauswahl von Substanzen ermöglicht oder für bestimmte definierte Substanzgruppen gültige Ergebnisse liefert.

Die Evaluierung der Methode (Evaluation)

Das so erarbeitete Ersatzverfahren muß nun auf seine tatsächliche mechanistische Ähnlichkeit hin zum Geschehen in vivo untersucht werden. Auf dieser Basis wird dann die eigentliche Prüfung der Eignung für die Wirkstoffindung im Vergleich zum Tierversuch möglich. Bei toxikologischen Modellen muß ein entsprechender Vergleich der Mechanismen der Zell- und Gewebeschädigung zeigen, ob tatsächlich relevante Pathomechanismen im in vitro Modell studiert werden können. Tierversuchsfreie Methoden, die einen vollständigen Ersatz eines konkreten



Tierversuchs anstreben, werden in der Regel eine solche mechanistische Legitimierung benötigen. Der empirische Vergleich bezüglich einer begrenzten Auswahl von Testchemikalien kann da kaum befriedigen.

Folgende grundlegenden Probleme bestehen beim Vergleich pharmakologischer und toxikologischer Modelle *in vivo* und *in vitro*:

- Das Datenmaterial zu Tierversuchen ist uneinheitlich und begrenzt.
 - Die Durchführung desselben Tierversuches weicht in verschiedenen Laboratorien oft erheblich voneinander ab. Hinzu kommen verschiedenartige Dokumentation und der begrenzte Umfang der erhobenen Daten, da im Tierversuch angestrebt wird, mit einer minimalen Tierzahl zu arbeiten.
- In vivo wird häufig mit verschiedenen Tierarten gearbeitet.
 Speziesunterschiede (gelegentlich sogar Stammunterschiede) sind ein großes Problem der Validierung.
 Dies führt dazu, daß nur ein Teil des Datenmaterials aus dem Tier-

- versuch zum Vergleich mit der Ersatzmethode herangezogen werden kann.
- Es sind kaum Daten zu Substanzen publiziert, die im Tierversuch nicht wirksam sind.
- Der Vergleich von pharmakologischen Modellen erfolgt vor allem durch den Vergleich der schützend wirkenden Stoffe. Es ist tierexperimentell jedoch kaum auszuschließen, daß eine Substanz nicht protektiv wirken könnte. Beispielsweise könnten eine andere Dosierung, Verabreichung oder die Gabe zu einem anderen Zeitpunkt für eine Schutzwirkung notwendig sein. Aus diesem Grund werden Negativergebnisse kaum publiziert. Für einen Vergleich zwischen in vivo und in vitro Methoden ist es aber gerade wünschenswert, zu zeigen, daß auch negative Ergebnisse übertragbar sind.
- Zahlreiche Aspekte des Tiermodells werden im Zellmodell nicht erfasst.
- Einige wesentliche Aspekte des Tierversuchs können im Zellmo-

- dell nicht nachgeahmt werden (Rasmussen, 1993): Gelangt das Medikament überhaupt im Tier an seinen Wirkort? Welche Bedeutung hat die Elimination der Substanz z.B. über die Niere? Welche anderen Zellen und Organe wechselwirken mit den untersuchten Phänomenen z.B. durch Metabolisierung der Prüfsubstanz? Welche biologischen Barrieren bestehen (z.B. Blut/Hirn-Schranke)? Andert sich die Durchblutung des untersuchten Organs im Tierversuch? Welche Interaktionen bestehen mit dem Nervensystem? Welche Rolle spielen die nur teilweise nachzuahmende physiologische Zellumgebung oder Zell/Zell-Kontakte?
- Statt "lebend tot" erhält man in vitro eine kontinuierliche Reaktion. Letalität im Tierversuch und Toxizität in vitro sind zwei grundverschiedene Dinge: Im einen Fall stirbt der gesamte Organismus; im anderen Fall stirbt lediglich ein Teil einer sehr homogenen Zellpopulation. Es ist schwierig festzulegen, welches Ausmaß des Zellun-

Tabelle: Optimaler Ablauf eines Validierungsprozesses

PHASE	Gruppen	Substanzen	HAUPTZIELE
Testentwicklung	11 od 30 10 nog 20	direction clare?	angestrebte Analogie zum Tierversuch, Festlegung des Testzieles (getestete Hypothese)
Evaluierung shodtsild von generalist	1 (i.A. durch Entwickler)	?	mechanistische Charakterisierung des in-vitro-Modelles Erhebung der ersten Datenbasis empirischer Vergleich zum Tierversuch
Prävalidierung	2 2-3 3	ca. 5 ca. 15 ca. 20-30	Verbesserung des Protokolls Transferierbarkeit des Protokolls Probedurchführung des Protokolls
Validierung a) unabh. Testprüfung b) Vorbereitung c) vorläufige Phase d) definitive Phase	Leitung + Referenzlab. Experten, mögl. Lab. ca. 4/Test	ca. 20	Prüfung durch ECVAM, ZEBET, o.ä. Festlegung Studien-Design, -Plan uProtokoll Auswahl Laboratorien und Testsubstanzen Probelauf, Verbesserung des Studien-Plans, Identifizierung von Defiziten doppelt-blinder Laborvergleich
Auswertung a) Biometrie b) Bewertung	1 (unabh.) Leitung und unabhängige Experten	Instant Posts Interted and State of the Regulation section	Bestimmung von Varianz, Spezifität, Selektivität und prädiktivem Wert Gesamtbewertung Planung des weiteren Vorgehens
Report / Anerkennung	Leitung	i m allowed	Publikation, Dokumentation, unabhängiger Review, Schritte zur behördlichen Anerkennung



tergangs eine für den Organismus bedrohliche Schädigung repräsentiert. Würde eine Substanz, die die Hälfte der Zellen in einem Versuchsansatz schützt, bei vergleichbarer Konzentration im Zielgewebe des Versuchstiers das Überleben garantieren?

Die skizzierten Probleme erlauben lediglich, eine Ähnlichkeit des *in vitro* Modells mit der Pathophysiologie der Erkrankung bzw. des toxischen Mechanismus zu erarbeiten. Die entscheidende Überprüfung des Modells erfolgt dadurch, daß Voraussagen aus dem Studium des Zellmodells für den Tierversuch gemacht und mit entsprechenden tierexperimentellen Daten überprüft werden (prädiktiver Wert).

Die Prävalidierung (Prevalidation)

Eine Reihe von Überlegungen sollten vor dem Eintritt in den eigentlichen Validierungsprozeß stehen (Curren *et al.*, 1995; Flint, 1994):

• Besteht überhaupt eine Notwendigkeit für den Tierversuch?

- Werden im Tierversuch spezifische Parameter bestimmt, die auf spezifische Wirkungen einer Testsubstanz im Versuchstier rückschließen lassen?
- Besteht ein ausreichendes Verständnis der Pathomechanismen sowohl in vivo als auch in vitro, um einen aussagekräftigen Vergleich zu ermöglichen?
- Ist die Testmethode standardisiert?
- Gibt es akzeptable "positive Kontrollen" mit Stoffen, die im Modell wirksam sind und als Standard benutzt werden können?
- Besteht die Möglichkeit, die Methode in einer ausreichenden Zahl von Laboratorien (international?) zu evaluieren und zu validieren?
- Kann die Evaluierung mit ausreichend vielen Testsubstanzen unter ausreichend vielen Bedingungen erfolgen, um Variabilität, Eignung und Anwendbarkeit der Methode zu prüfen?
- Besteht eine ausreichende Basis für eine Kooperation zwischen Hochschule, Industrie und Behörden bei der Validierung, durch die eine internationale Akzeptierung gesichert werden kann?

Darüber hinaus ist es aus den bereits angeführten Gründen wichtig, den Einsatz des Testes möglichst genau zu definieren: Ein Screening-Test wird in der Regel weniger strengen Kriterien unterworfen werden müssen als eine Ersatzmethode für einen klassischen Toxizitätstest (Green, 1993).

Bevor in einem größeren Ansatz mehrere Laboratorien gemeinsam eine Validierung vornehmen, versucht die Prävalidierung das Testprotokoll zu prüfen, zu verbessern und die Robustheit einer Methode für die Transferierung in andere Labors zu prüfen. Ziel dieser experimentellen Vorprüfung ist es, zu einem optimierten Versuchsprotokoll zu gelangen, das es ermöglicht, den Test an einem beliebigen Ort zu etablieren. Dies betrifft sowohl die technischen und personellen Voraussetzungen als auch die Reproduzierbarkeit an verschiedenen Orten. Zu den vorbereitenden Maßnahmen sollte die Bestimmung der Intralabor-Variabilität gehören. Außerdem müssen mögliche Testsubstanzen für die spätere Validierung im Modell erprobt werden. Neben den Möglichkeiten der Methode sollten auch ihre Grenzen herausgearbeitet werden. Wichtig sind z.B. Substanzklassen, die nicht geprüft werden können. Den Abschluß der Prävalidierung sollte ggf. ein begrenzter (nicht notwendigerweise doppelt-blinder) Ringversuch mit 2–3 beteiligten Gruppen darstellen, der als erster Testlauf für die eigentliche Validierung angesehen werden kann. Einrichtungen wie ECVAM und ZEBET bieten sich hier an.

Auswertung unabhängig und Report Validierung Evaluierung Testentwicklung Aufwand

neu:

Ablauf

- · unabh. Leitung / Koordinierung
- Betonung der Testentwicklung

Aufwand

alt

- vorgeschaltete Prävalidierung mit unabhängiger Testprüfung
- bei Bedarf Wiedereintritt in Entw.
- Validierung mit weniger Gruppen
- Probelauf für Validierung
- unabhängige Auswertung

Abbildung: Ablauf und zeitlicher Aufwand von Validierungsstudien.

Anerkennung

Die Validierung (Validation)

Die Arbeiten der Prävalidierung sollen im positiven Fall die Basis für den Validierungsprozeß bilden (Balls et al., 1990). Entsprechend aufbereitet sollen sie geeignete Kooperationsgruppen/-unternehmen, Geldgeber und (inter-)nationale Einrichtungen für Validierungsstudien (ECVAM, ZEBET etc.) in die Lage



versetzen, zu beurteilen, ob ein Validierungsprozeß nötig und erfolgversprechend ist. Dazu muß zunächst eine Leitungsgruppe, die weitgehend unabhängig von der eigentlichen Versuchsdurchführung ist, eingesetzt werden. Sie koordiniert zwischen Geldgebern, beratenden Experten, Institutionen und beteiligten Arbeitsgruppen.

Der genaue Ablauf einer Validierungsstudie (s. auch Abb. und Tab.) muß zwischen allen Beteiligten vereinbart werden. Dazu sollte geprüft werden, ob mehrere Ersatzmethoden zum gleichen Tierversuch parallel einer Validierung unterzogen werden können. Dies ermöglicht sowohl die vergleichende Bewertung der Ersatzmethoden als auch die Beurteilung, ob ev. durch verschiedene Methoden komplementäre Ergebnisse erzielt werden, um gegebenenfalls den Tierversuch nicht durch einen Einzeltest, sondern durch eine Testbatterie zu ersetzen.

Zentral für die Erarbeitung eines Studien-Designs (Feltem *et al.*, 1995) müssen neben dem aus der Prävalidierung erarbeiteten Versuchsprotokoll insbesondere folgende Punkte geklärt werden:

- Welche Tests sollen gegebenenfalls vergleichend durchgeführt werden?
- Welche Laboratorien wollen/können an der Validierung teilnehmen?
- Welche gemeinsamen Quellen für Geräte und Verbrauchsmaterialien sollen genutzt werden?
- Welche Testsubstanzen sollen im Ringversuch eingesetzt werden (Stabilität, Kontrollen)?
- Definition der Datenbasis und Struktur (s.u.)
- Vereinheitlichtes Datenübergabe-Protokoll
- Wie wird das Ergebnis der Validierung ermittelt?
- Wann sollen von wem die daraus erwachsenden Konsequenzen bestimmt werden (nächste Schritte)?

Aus dem Studiendesign leitet sich der eigentliche Studienplan ab, d.h. die konkrete Durchführungsvereinbarung zwischen den Parteien. Im

Studienplan muß das Ziel der Studie eindeutig definiert sein, und er muß in jeder Phase der Studie eine aktuelle Übersicht über den experimentellen Fortgang ermöglichen. Außerdem muß er bei allen Änderungen im Projektablauf überarbeitet bzw. geändert werden, wie z.B. bei Änderung zeitlicher Vorgaben. Eine wichtige Voraussetzung zur Validierung einer Methode ist die Standardisierung der Komponenten in den verschiedenen Laboratorien. Wesentlich dafür sind die Verwendung von (a) Chemikalienbanken, (b) Zellbanken, (c) Datenbanken und (d) Referenz-Laboratorien (Goldberg et al., 1993).

Es empfiehlt sich, eine solche Validierung zunächst in einem begrenzten Probelauf (preliminary validation) durchzuführen, um Fehler im eigentlichen Ringversuch zu vermeiden (definitive validation). Die Daten des begrenzten Probelaufes können gegebenenfalls mit in die Validierung einfließen. Selbstverständlich sollte die Prüfung doppelt blind erfolgen (Kodierung von Substanzen und Laboratorien). Die unabhängige Auswertung durch Unbeteiligte ist anzustreben. Wegen der Unfallgefahr ist auf das Problem beim Umgang mit codierten Substanzen hinzuweisen: Es muß im Notfall jederzeit eine Identifizierung gefährlicher Substanzen zur Unfallbegrenzung gewährleistet sein.

Die Auswertung der Validierung

a) Die statistische Evaluierung der Ersatzmethode

Es muß nochmals betont werden, daß vor Durchführung der praktischen Validierung eine eindeutige Klärung der Art der zu erhebenden Daten und ihrer Dokumentation erfolgen muß. Ebenso müssen die Auswerteverfahren vereinheitlicht werden (gemeinsame Software). Der natürliche Wunsch nach einer einfachen Auswertungsmethode muß gegen die Tatsache abgewogen werden, daß jede Vereinfachung mit einem Informationsverlust verbunden ist.

Drei Arten von Informationen sollen zur Ersatzmethode selbst erzielt werden:

- 1 Die Validität des Verfahrens soll ermittelt werden, d.h. die Ähnlichkeit mit dem Tierversuch, also die Korrelation mit den biologischen Parametern.
- 2 Die Reproduzierbarkeit der Methode besagt, ob die Methode in verschiedenen Laboratorien zum selben Ergebnis führt.
- 3 Die **Präzision** der Methode soll die Streuung der Ergebnisse bei laborinternen Wiederholungen beschreiben. Da dies häufig direkt mit dem betriebenen Aufwand zusammenhängt, stellt sich hier die Effizienz-Frage.

Beim Vergleich von Tierversuch und Ersatzmethode geht es im einfachsten Fall (Ja/Nein-Antworten in Tierversuch und Ersatzmethode) au-Berdem um die Bestimmung von:

- 4 Sensitivität, d.h. die Wahrscheinlichkeit, daß das Ersatzverfahren ein positives Ergebnis erbringt, wenn eine Substanz im Tierversuch wirksam ist.
- 5 Spezifität, d.h. die Wahrscheinlichkeit, daß das Ersatzmodell ein negatives Ergebnis zeigt, wenn die Substanz im Tierversuch keine Wirkung hat.
- 6 Prädiktiver Wert, d.h. die Wahrscheinlichkeit für einen positiven Effekt sowohl im Tierversuch als auch im Ersatzverfahren (positiver prädiktiver Wert) bzw. umgekehrt die Wahrscheinlichkeit, daß die Substanz sowohl im Tierversuch als auch im *in vitro* System unwirksam ist (negativer prädiktiver Wert).

Falls in einem oder beiden Verfahren eine graduelle Bestimmung des Meßparameters erhoben wird, müssen Regressionsanalysen verwendet werden.

b) Bewertung

Entsprechend dem vorab vereinbarten "Studienplan" findet abschließend die Gesamtbewertung der Validierungsstudie statt. Dabei ist auf eine möglichst weitgehende Unabhängigkeit der Beurteiler von den



Interessen der Teilnehmer an der Validierungsstudie zu achten. Die Bewertung sollte vollständig und nachvollziehbar sein und entsprechend dokumentiert und publiziert werden. Im Rahmen der Bewertung sollten bereits die weitergehenden Schritte verabredet und mögliche Adressaten in die Beurteilung einbezogen werden, denn nach dem langen Weg bis zur validierten Ersatzmethode beginnt nun der Marsch durch die Institutionen.

Ausblick

Die Durchführung der hier beschriebenen Validierung eines Testverfahrens verlangt die enge Kooperation der Wissenschaftler verschiedener Laboratorien mit Biometrikern, Anwendern in der Industrie und den zuständigen nationalen und internationalen Behörden. Nur in einem derart konzertierten Vorgehen kann das Ziel einer allgemeinen Akzeptanz der Methode erreicht werden. Diese Aufgabe wird in der Regel den Einzelnen überfordern. Durch die Schaffung nationaler (ZEBET) und internationaler Einrichtungen (ECVAM) entsteht die nötige Infrastruktur für solche umfangreichen Validierungsprojekte. Angesichts des enormen Aufwandes des Verfahrens wird im Einzelfall zu prüfen sein, ob der jeweilige Tierversuch realistisch und repräsentativ für die erwarteten Aussagen ist. Allzu häufig bewirken historische Festlegungen, daß ein Tierversuchsprotokoll nicht auf der Höhe der heutigen wissenschaftlichen Erkenntnis ist. Keinesfalls darf es dazu kommen, mit viel Aufwand das Beinahe-Abbild einer schlechten Methode zu entwickeln. Es wird zu prüfen bleiben, ob die Validität einer neuen Methode nicht auch mit anderen Maßstäben (klinischen Erkenntnissen, mechanistischen Übereinstimmungen, Plausibilität) gemessen werden kann. Dies gilt umso mehr, als in vitro Methoden auch den Zugriff auf menschliches Material z.B. auch von dem entsprechenden

Krankheitsbild ermöglichen. Solches Material ist natürlich nur begrenzt verfügbar. Es wäre auch vorstellbar, daß die Validierung eines Modells mit Zellinien oder Zellen aus Labortieren im Vergleich zu solchen menschlichen Zellen erfolgt. Damit könnte der Begriff der Validierung eines *in vitro* Verfahrens eine ganz neue Bedeutung erlangen, die auch über den Tierversuch selbst hinaus geht.

Danksagung

Der Kommentar verwertet u.a. zwei "state-of-the-art-lectures" von Michael Balls (ECVAM, Ispra, Italien) und Berthold Schneider (MHH Hannover). Anlaß war ein Workshop zum Thema "Statistics and Validation" bei einer Tagung am Paul-Ehrlich-Institut (International Association of Biological Standardization: "Replacement, Reduction and Refinement of Animal Experiments in the Development and Control of Biological Products", 2.-4.11.94). Die Autoren danken A. Wendel (Universität Konstanz) für fruchtbare Diskussion und Durchsicht des Manuskripts.

Literatur

Balls, M., Blaauboer, B., Brusick, D., Frazier, J., Lamb, D., Pemberton, M., Reinhardt, C., Roberfroid, M., Rosenkranz, H., Schmid, B., Spielmann, H., Stammati, A-L. und Walum, E. (1990). Report and recommendations of the CAAT/ER-GATT workshop on the validation of toxicity test procedures. *ATLA – Alternatives to Laboratory Animals* 18, 313–337.

Curren, R. D., Southee, J. A., Spielmann, H., Liebsch, M., Fentem J. H. und Balls, M. (1995). The role of prevalidation in the test development and validation process. *ATLA*, im Druck.

Fentem, J. H., Prinsen, M. K., Spielmann, H., Walum, E. und Botham, P. (1995). Validation – lessons learned from practical experience. *Toxicol. in vitro*, zur Veröffentlichung

eingereicht.

Flint, O. (1994). Rational programmes for the assessment of drug toxicity. *Proceedings of the 6th EAVPT Congress* 1994, 9–12.

Goldberg, A. M., Frazier, J. M., Brusick, D., Dikens, M. S., Flint, O., Gettings, S. D., Hill, R. N., Lipnick, R. L., Renskers, K. J. und Bradlaw, J. A. (1993). Report on the validation and technology transfer committee on the Johns Hopkins Center for alternatives to animal testing. Framework for validation and implementation of in vitro toxicity tests. *Xenobiotica* 23, 563–572.

Green, S. (1993). Regulatory agency considerations and requirements for validation of toxicity. *Toxicol. Lett.* 68, 119–123.

Hartung, T., Odenthal, K. P., Sauer, A., Schwarz, T. und Wendel, A. (1994). Entwicklung eines Verfahrens zum pharmakologischen Screening von Wirksubstanzen gegen Septischen Schock im Zellkultursystem. In H. Schöffl, H. Spielmann, H. A. Tritthart, K. Cußler, U. Fuhrmann, A. F. Goetschel, F. P. Gruber, C. Heusser, H. Möller, H. Ronneberger, A. Vedani (Hrsg.), Ersatz- und Ergänzungsmethoden zu Tierversuchen. Wien, New York: Springer, im Druck.

Rasmussen, E. S. (1993). The role of in vitro experiments in animal welfare. *Hum. Exp. Toxicol.* 12, 522–527.

Korrespondenzadresse

Thomas Hartung Biochemische Pharmakologie der Universität Konstanz Postfach 5560 D-78434 Konstanz