

U. Schuhmacher-Wolz
M. Hassauer
F. Kalberlah

**Ableitung von Arbeitsplatzrichtwerten
für Stoffe mit reproduktionstoxischen
Eigenschaften**

Diese Veröffentlichung ist eine Zusammenfassung des Abschlussberichts zum Projekt „Erweiterung des ARW-Konzepts für Stoffe mit reproduktionstoxischen Eigenschaften (außer Störung der weiblichen Fertilität)“ - Projekt F 2008 - im Auftrag der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Autoren: Dr. Ulrike Schuhmacher-Wolz
Dr. Martin Hassauer
Dr. Fritz Kalberlah

Forschungs- und Beratungsinstitut Gefahrstoffe, FoBiG GmbH
Werderring 16, 79098 Freiburg i. Br.

Herausgeber: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin
Friedrich-Henkel-Weg 1-25, 44149 Dortmund
Telefon: 0231 9071-0
Telefax: 0231 9071-2454
E-Mail: poststelle@buaa.bund.de
Internet: www.buaa.de

Berlin:
Nöldnerstr. 40-42, 10317 Berlin
Telefon: 030 51548-0
Telefax: 030 51548-4170

Dresden:
Proschhübelstr. 8, 01099 Dresden
Telefon: 0351 5639-50
Telefax: 0351 5639-5210

Alle Rechte einschließlich der fotomechanischen Wiedergabe und des auszugsweisen Nachdrucks vorbehalten.

Ableitung von Arbeitsplatzrichtwerten für Stoffe mit reproduktionstoxischen Eigenschaften

Die Ableitung von Arbeitsplatzrichtwerten (ARW)¹ wird zurzeit nur für Stoffe durchgeführt, für die eine krebserzeugende oder reproduktionstoxische Wirkung nicht bekannt ist. Da für Reproduktionstoxizität die Existenz eines praktischen Schwellenwerts jedoch plausibel angenommen werden kann, existieren Überlegungen, auch für Stoffe mit diesem kritischen Endpunkt zukünftig Arbeitsplatzrichtwerte abzuleiten.

Bei der Aufstellung von ARW ist in der Regel von tierexperimentellen Befunden auszugehen und es hat ggf. eine Hochrechnung auf dauerhafte Exposition (gegenüber kürzerer experimenteller Exposition) zu erfolgen. Für das vorliegende Projekt ergibt sich die Fragestellung, ob die entsprechende allgemeine Methodik zur Ableitung von ARW auf die toxikologischen Endpunkte „männliche Fertilität“ und „Entwicklungstoxizität“ in Hinsicht auf Interspezies- und Zeitextrapolation qualitativ und quantitativ übertragbar ist?

Die Endpunkte „männliche Fertilität“ und „Entwicklungstoxizität“ untergliedern sich in mehrere einzelne spezifische Endpunkte, von denen einige möglicherweise besonders sensitiv Effekte auf den männlichen Reproduktionstrakt bzw. die sich entwickelnde Leibesfrucht anzeigen. Es war entsprechend auch zu prüfen, ob spezifische Endpunkte bereits bei niedrigerer Exposition indikativ für einen reproduktionstoxischen Effekt sind, welche das wären und welche regulatorischen Konsequenzen bei der Aufstellung eines ARW daraus folgen könnten.

Die Untersuchung basiert auf einer umfassenden Literaturlauswertung.

Männliche Fertilität

Qualitativer Speziesvergleich

Nachteilige Einwirkungen auf die männliche Fertilität werden hier definiert als „nachteilige Auswirkungen auf die Libido, das Sexualverhalten, alle Aspekte der Spermatogenese..., auf den Hormonhaushalt oder auf physiologische Reaktionen, die im Zusammenhang mit der Befruchtungsfähigkeit, der Befruchtung selbst oder der Entwicklung der befruchteten Eizelle bis zur Einnistung im Uterus stehen.“ Physiologisch betrachtet ist die Fortpflanzung beim Menschen und den Säugetieren prinzipiell vergleichbar. Bezogen auf den Endpunkt „männliche Fertilität“ existieren jedoch einige grundsätzliche Speziesunterschiede, die bei der Bewertung berücksichtigt werden sollten, um falsche Schlussfolgerungen zu vermeiden.

So weisen die gängigen Versuchstierspezies (Maus, Ratte Kaninchen) gemessen am Fertilitätsindex eine deutlich höhere Fertilität als der Mensch auf, bedingt durch die höhere Spermienreserve der Tiere. Eine Änderung im Fertilitätsindex tritt im Tierversuch erst nach massiver Schädigung der Spermienbildung ein (z.B. bei der Ratte erst nach > 90 %iger Reduktion der Spermienproduktion), während beim Menschen schon geringere Störungen der Spermienproduktion die Fertilität beeinträchtigen

¹ Wir wählen im vorliegenden Bericht den Begriff „Arbeitsplatzrichtwert“ (ARW) entsprechend dem Titel des Forschungsprojekts, wobei sich in jüngster Zeit entsprechend den Vorgaben der neuen Gefahrstoffverordnung die Terminologie geändert hat: ARW werden zu AGW (-gesundheitsbasierte- Arbeitsplatzgrenzwerte).

können. Tierexperimentelle Daten, bei denen nur die Fertilität als Reproduktionserfolg gemessen werden, werden deshalb allgemein als ungeeignet betrachtet, um daraus quantitative Abschätzungen bezüglich reproduktionstoxischer Risiken für den Menschen abzuleiten.

Da Informationen über die Spermienqualität (Zahl, Motilität, Morphologie) häufig die einzigen beim Menschen verfügbaren Daten sind, kommt diesem spezifischen Endpunkt ein besonderes Gewicht zu. Da die Dauer der Spermienreifung sich bei den einzelnen Spezies unterscheidet (Ratte ca. 50 d, Mensch ca. 74 d), sollte bei der Bewertung der einzelnen Studien darauf geachtet werden, dass nicht unterschiedliche Zeitpunkte der Spermienentwicklung als Speziesunterschiede interpretiert werden.

Auch die Bewertung von Substanzeffekten auf das endokrine System des Menschen anhand von Tierdaten ist schwierig, weil hier zum Teil Speziesunterschiede auftreten: Bei der Ratte ist z.B. alleine Testosteron ausreichend, um die Spermatogenese aufrecht zu erhalten, während beim Menschen Testosteron und FSH (Follikelstimulierendes Hormon) notwendig sind.

Weiterhin zeigt sich, dass die betroffenen spezifischen Endpunkte bei Tier und Mensch vergleichbar sein können (z.B. Dibromchlorpropan führt sowohl beim Menschen wie beim Versuchstier zu Testesatrophie, verminderter Spermienzahl und Fertilität), sich bei manchen Substanzen aber auch unterscheiden (z.B. Kepon beeinflusst beim Menschen die Morphologie, Zahl und Motilität der Spermien und führt zu verminderter Libido und Impotenz, während im Tierversuch morphologisch veränderte Spermien, Testesatrophie und verminderte Fertilität auftreten).

Weiterhin kann es sein, dass infolge unterschiedlicher Empfindlichkeiten nicht in allen Spezies Effekte auftreten.

Die qualitativen Betrachtungen führten insgesamt zu folgendem Ergebnis:

- Nicht bei allen Stoffen sind bei Mensch und Tier oder auch zwischen verschiedenen Tierspezies die Zielorgane (spezifische Endpunkte der Reproduktionstoxizität) identisch. Deshalb sollten auch bei unterschiedlicher Lokalisation tierexperimentelle Hinweise auf die schädigende Wirkung eines Stoffes auf den männlichen Reproduktionstrakt prinzipiell auch als Hinweis für eine solche Wirkung beim Menschen gewertet werden, sofern Humanbefunde diese Interpretation nicht widerlegen und keine speziesspezifischen Besonderheiten dafür verantwortlich gemacht werden können.
- Da nicht alle Spezies gleich empfindlich gegenüber bestimmten Stoffen reagieren, sollte ein negatives Resultat in einer Spezies immer erst durch Untersuchungen in mindestens einer weiteren Spezies bestätigt werden, bevor man auch für den Menschen keine schädigende Wirkung annimmt.
- Tierexperimentelle Untersuchungen zum Fertilitätsindex werden nur als eingeschränkt geeignet bewertet, um aussagekräftige Daten bezüglich einer fertilitätsmindernden Wirkung beim Menschen zu erhalten. Häufig erweisen sich die Tierversuche als zu unempfindlich, sofern nicht spezielle Versuchsanordnungen (z.B. kontinuierliche Verpaarungsstudie) verwendet wurden.
- Wegen der im Tierversuch meist höheren Untersuchungstiefe, die beim Menschen aus ethischen und praktischen Gründen meist nicht erreicht werden kann, liefern Tierversuche wichtige Hinweise auf Wirkmechanismen und Zielorgane.

- Einzelne tierexperimentelle Befunde, wie z.B. morphologische Veränderungen in der Testes einschließlich Störungen der Spermatogenese lassen sich wegen prinzipieller morphologischer Unterschiede nicht direkt mit der Situation beim Menschen vergleichen. Sie geben aber wichtige Hinweise auf Zielzellen/-organe.
- Toxikokinetische und –dynamische Unterschiede können dazu führen, dass bestimmte Befunde nicht direkt auf andere Spezies übertragen werden können (z.B. quantitative Übertragung der Befunde zu Phthalsäureestern nicht möglich). Sofern Kenntnisse über prinzipielle Speziesunterschiede vorliegen, sind diese in ihrer Relevanz zu diskutieren. Pharmakokinetische Modelle können ein geeignetes Hilfsmittel sein, um kinetische Unterschiede zu berücksichtigen.

Quantitativer Speziesvergleich

Eine Extrapolation der Tierdaten auf den Menschen dient dazu, Unterschiede in der Toxikokinetik und Unterschiede in der Toxikodynamik adäquat zu berücksichtigen. Durch die Interspeziesextrapolation wird

- unter Berücksichtigung der systematischen Unterschiede zwischen Spezies verschiedener Größe eine geeignete Dosismaßzahl für Effekte angeben und
- die Variabilität in der Sensitivität der Spezies

berücksichtigt. Letztere kann sowohl in Bezug auf toxikokinetische als auch toxikodynamische Faktoren auftreten.

Die Berücksichtigung systematischer toxikokinetischer Unterschiede dient dazu, eine mittlere Schätzung einer risikoanalogen Dosis aus dem Tierversuch für den Menschen abzuleiten. Da die früher praktizierte Vorgehensweise, die Dosis als Körpergewichtsäquivalente von einer Spezies auf die andere zu übertragen, sich als problematisch erwies (kleinere Organismen tolerieren höhere Dosen relativ zum Körpergewicht - Angabe in mg/kg Körpergewicht - als größere), ist man dazu übergegangen, Dosisangaben zwischen Organismen unterschiedlicher Größe unter Berücksichtigung allometrischer Beziehungen („Scaling“) zu übertragen. In jüngerer Zeit hat sich das Scaling nach Grundumsatz, bei dem die wirksame Dosis im Menschen äquivalent zur wirksamen Dosis im Tier $\times (\text{Körpergewicht}_{\text{Mensch}}/\text{Körpergewicht}_{\text{Tier}})^{3/4}$ ist, durchgesetzt. Im vorliegenden Projekt wurde geprüft, inwieweit auch bei dem Endpunkt männliche Fertilität bei der Interspeziesextrapolation ein Scaling nach Grundumsatz anwendbar ist und ob neben einem Scaling weitere Faktoren bei der Extrapolation notwendig sind.

In einer Literaturlauswertung wurden vergleichbare Daten unterschiedlicher Spezies zu Störungen der männlichen Reproduktion zusammengetragen. Dabei wurden Daten bevorzugt, wenn Befunde beim Menschen und mindestens einer Tierspezies ermittelt werden konnten. Da verlässliche epidemiologische Daten jedoch nur für wenige Stoffe vorliegen (z.B. Dibromchlorpropan, Ethylendibromid, Gossypol, Sulfasalazin), wurden auch Daten in zwei unterschiedlichen Tierspezies in die Auswertung mit einbezogen. Dann wurde darauf geachtet, dass die Untersuchungen übereinstimmten hinsichtlich

- dem Expositionspfad und
- der Expositionsdauer.

Zusätzlich wurde bei Gegenüberstellungen nach Möglichkeit auch auf eine Übereinstimmung in der Untersuchungstiefe geachtet (diese ist jedoch nicht immer gewährleistet).

Sofern Vergleiche zwischen Mensch und Tier vorgenommen wurden, wurden beim Menschen meist die spezifischen Endpunkte Spermienmenge/-motilität, -morphologie, ggf. Aussagen zum Hormonstatus herangezogen, während bei den Tierstudien die empfindlichsten getesteten Parameter verwendet wurden.

Insgesamt lag nur eine begrenzte Anzahl von geeigneten Datensätzen vor: 33 Speziesvergleiche wurden herangezogen. Häufig handelt es sich dabei um Studien in Ratten und Mäusen, bei denen ein Vergleich hinsichtlich der Speziesunterschiede schwierig ist, da Stammesunterschiede Speziesunterschiede überdecken können. Des Weiteren liegen nur wenige Studien vor, in denen ein NOAEL und LOAEL bestimmt wurde (10 Datensätze). In vielen Studien wurden selbst bei der höchsten getesteten Dosis keine Effekte beobachtet. Andererseits traten Effekte häufig schon in der niedrigsten getesteten Dosis auf, die dann als LOAEL in die Auswertung einging. Die Höhe der tatsächlichen NOAEL- bzw. LOAEL-Werte ist unbekannt, was in einzelnen Fällen zu Unter- oder Überschätzungen der daraus ermittelten Interspeziesfaktoren führen kann. Die folgende Tabelle fasst die Datenbasis und das Auswertungsergebnis zusammen:

Tab. 1.1 Interspeziesvergleich zu Effekten auf die männliche Fertilität; Gegenüberstellung von LOAEL- und/oder NOAEL-Werten

	Alle Daten und Pfade (oral*, inhalativ, intraperitoneal)
Substanzen n =	14
Speziesvergl. n =	33
NOAEL-/LOAEL-Quotienten n =/n =	17/26
AM	6,12
SD	19,34
GM	1,31
GSD	5,96
50-Perz.	1,00
90-Perz.	5,45

* nach Transformation über allometrisches Scaling nach kalorischem Grundumsatz

Daraus lässt sich schlussfolgern:

- Die Werte für das geometrische Mittelwert (GM) und das 50-Perzentil widersprechen nicht einem Scaling nach Grundumsatz.
- Die Daten weisen darauf hin, dass neben einem Scaling nach Grundumsatz bei der Interspeziesextrapolation noch ein weiterer Faktor (z.B. Faktor 5 beim 90 Perzentil) notwendig wäre, um einen größeren Anteil der beobachteten Unterschiede abzudecken.
- Diese Schätzwerte (Mittelwerte, 90-Perzentil) sind aufgrund ihrer schlechten statistischen Absicherung (kleine Anzahl von Datensätzen) in ihrer Qualität unbefriedigend.

- Da diesen Daten bereits eine umfangreiche Literaturrecherche zu Grunde liegt, ist zu vermuten, dass derzeit eine noch weitergehendere Literaturliteraturauswertung, als sie im Rahmen dieses Projekts möglich war, kaum so viele weitere geeignete Datensätze liefern dürfte, dass sich die statistische Absicherung entscheidend verbessern würde.

Zeitextrapolation

Störungen der Spermatogenese stellen einen wichtigen spezifischen Endpunkt bei der Beurteilung von Effekten auf die männliche Fertilität dar. Je nach Angriffspunkt einer Substanz und der betrachteten Spezies zeigen sich Effekte auf die Spermien im Ejakulat aber erst nach bis zu 90 Tagen (Mensch). Ziel ist es jedoch, bei der Risikobewertung auch auf Basis von Studien mit kürzerer Dauer möglichst definitive Aussagen treffen zu können. Insbesondere bei Tierversuchen können bereits nach früheren Zeiten mittels anderer Untersuchungsmethoden (z.B. histopathologische Untersuchungen) Effekte beobachtet werden, lange bevor die Spermien ausgereift sind. So ergab eine Analyse, dass reproduktionstoxische Effekte meist schon nach 4 Wochen erkennbar sind, wenn empfindliche Parameter wie Testeshistologie, Epididymishistologie oder Organgewichte der Reproduktionsorgane untersucht wurden, die möglichst durch Angaben zur Spermienzahl, -motilität und -morphologie sowie zur Körpergewichtsentwicklung ergänzt werden. Zur Untersuchung, ob mit zunehmender Expositionszeit eine Wirkungsverstärkung auftritt, wurden die NOAEL- bzw. LOAEL-Werte von Studien unterschiedlicher Dauer in einer Spezies mit möglichst vergleichbarer Untersuchungstiefe miteinander verglichen. Auch wenn auf Grund der begrenzten Anzahl an Daten kein statistisch gesicherter Zeitextrapolationsfaktor abgeleitet werden konnte, so zeigt sich, dass die ermittelten Zeitextrapolationsdaten für den Endpunkt männliche Fertilität alle etwas niedriger liegen, als die im Rahmen der ARW-Methodik verwendeten Faktoren für systemisch toxische Effekte (subakut/chronisch: Faktor 6). Der rechnerische Wert für männliche Fertilität beträgt 1,7 (Multiplikation des Faktors für subakut/subchronisch x Faktor subchronisch/chronisch = 1,4 x 1,2). Die folgende Tabelle fasst die Ergebnisse zusammen:

Tab. 1.2 Veränderung der Effektdosis bzw. -konzentration mit zunehmender Expositionsdauer (männliche Reproduktionstoxizität)

Vergleichszeitraum	Geometrischer Mittelwert (90-Perzentil, Anzahl der Datensätze)
Zwei subakute Zeiträume (meist 2 vs. 4 Wochen)	1,3 (90 Perzentil: 2,0; n = 21)
Subakut/subchronisch	1,4 (90 Perzentil: 2,8; n = 20)
Subchronisch/chronisch	1,2 (90 Perzentil: 1,8; n = 5)
Subakut/chronisch	2,9 (n = 3)

Auch bei dieser Auswertung zeigt sich eine unbefriedigende Datenlage, die insbesondere hinsichtlich chronischer Expositionsdauer keine qualifizierte Aussage ermöglicht.

Empfindlichster spezifischer Endpunkt

Verschiedene reproduktionstoxische Endpunkte erwiesen sich bei Untersuchungen zur männlichen Fertilität als unterschiedlich empfindlich.

In einer Auswertung von 2- und 4-Wochenstudien hinsichtlich der spezifischen Endpunkte, die beim LOAEL am häufigsten betroffen waren, erwiesen sich Veränderungen in der Testeshistologie als besonders empfindlich, gefolgt von Veränderungen im Epididymisgewicht, der Epididymishistologie und dem Testesgewicht. Zur quantitativen Analyse wurden die LOAEL-Werte aller untersuchten Parameter paarweise miteinander verglichen. Ein Vergleich mit den Endpunkten Kopulationsverhalten, Implantations- bzw. Trächtigkeitsindex, fetale Mortalität, die nach Angabe anderer Autoren um ca. den Faktor 3 weniger empfindlich als der spezifische Endpunkt Testeshistologie sind, war wegen fehlender Untersuchungen zu diesen Endpunkten nicht möglich. Die Datenlage wird insgesamt als zu unsicher bewertet, um daraus einen Faktor abzuleiten, der z.B. bei Studien zur Fertilität (Ein- oder Mehrgenerationenstudien) ohne vertiefte histologische Auswertung der Testes eingesetzt werden sollte.

Liegen dagegen subakut-(4-Wochen) Studien zur Untersuchung von Effekten auf die männliche Fertilität vor, bei denen der Endpunkt Testeshistologie nicht untersucht wurde, aber die Gewichte der Reproduktionsorgane erfasst sowie eine histologische Untersuchung der Epididymis durchgeführt wurde, scheint die Verwendung eines zusätzlichen Faktors für den nicht getesteten sensitivsten Endpunkt nicht gerechtfertigt zu sein (ermittelte Faktoren meist im Bereich von 1).

Liegt bei Subakutstudien alleine eine Auswertung der Spermienparameter oder der Fertilität vor, ohne Angaben zu Epididymis- und/oder Testeshistologie und Gewichten der Reproduktionsorgane, kann nicht mit Sicherheit angenommen werden, dass der NOAEL für männliche Fertilität richtig erfasst wurde. Es liegen aus unserer Sicht jedoch nicht genügend Daten vor, um einen Faktor für mangelnde Untersuchungstiefe mit ausreichender Genauigkeit zu quantifizieren.

Entwicklungstoxizität

Qualitativer Speziesvergleich

Entwicklungsschäden schließen alle schädlichen Wirkungen auf die Entwicklung der Nachkommenschaft ein, die während der Schwangerschaft verursacht werden und sich prä- oder postnatal manifestieren. Spezifische Endpunkte der Entwicklungstoxizität lassen sich generell in folgende vier Kategorien unterteilen:

- Funktionale Defizite (z.B. geistige Retardierungen, neurologische Auffälligkeiten, Verhaltensauffälligkeiten)
- Intrauterine Wachstumsverzögerungen (IUGR intrauterine growth retardation)
- Angeborene Fehlbildungen (Terata)
- Tod (Abort, Totgeburten oder Tod der Neugeborenen)

Die Effekte, die man bei einer Studie zur Entwicklungstoxizität zum Zeitpunkt der Geburt am besten beobachten kann, sind Letalität, Fehlbildungen und Wachstumsverzögerungen. Diese Effekte sind offensichtlich nicht vollständig unabhängig voneinander. So erreichen z.B. 75% der Embryonen und Feten mit strukturellen Anomalien nie einen Status, in dem sie lebensfähig wären. Bei ca. 33 – 50% der Kinder mit

einer intrauterinen Wachstumsretardierung treten Lernstörungen auf. Zum Teil bleiben solche Effekte unerkannt infolge früher Resorptionen oder Fehlgeburten oder werden nicht erkannt, weil der entsprechende Effekt nicht untersucht wurde.

Der Einfluss des Zeitpunkts der Exposition und der Dosis auf die einzelnen Effekte ist unterschiedlich. Für einige Substanzen scheint bei steigender Dosis ein Kontinuum an zunehmender Toxizität aufzutreten: Wachstumsverzögerung – Fehlbildung – Letalität. Andere Stoffe führen dagegen zu Wachstumsverzögerungen und Embryoletalität ohne teratogen zu wirken. Funktionale Defizite lassen sich im Gegensatz zu den anderen Effekten nur weitaus schwieriger und auch meist erst zu späteren Zeitpunkten postnatal nachweisen.

Ein qualitativer Vergleich der Entwicklungstoxizität in verschiedenen Spezies erbrachte folgende Ergebnisse:

- Entwicklungstoxische Effekte treten meist in mehreren Spezies auf. Positive Ergebnisse werden dabei häufig in Ratten und/oder Kaninchen erhalten.
- Manche Substanzen führen bei verschiedenen Spezies zu vergleichbaren Effekten (z.B. Dimethylformamid, Diethylhexylphthalat), andere wiederum zeigen unterschiedliche Effekte bei den verschiedenen Spezies (z.B. Thalidomid).
- Je nach Substanz reagieren manche Spezies besonders empfindlich, während andere Spezies gar keine entwicklungstoxischen Effekte aufweisen.
- Ursächlich für gravierende Speziesunterschiede in der Empfindlichkeit sind häufig qualitative und quantitative Unterschiede im Metabolismus.
- Daraus ergibt sich, dass ein negatives Ergebnis in nur einer Versuchstierspezies noch keine sicheren Rückschlüsse auf das Fehlen entwicklungstoxischer Effekte beim Menschen zulässt.
- Werden dagegen weder bei der Ratte noch beim Kaninchen entwicklungstoxische Effekte beobachtet, ist mit hoher Wahrscheinlichkeit ein solcher Effekt auch beim Menschen nicht zu erwarten.
- Ein entwicklungstoxischer Effekt in nur einer Spezies sollte als Hinweis darauf gewertet werden, dass die Substanz auch entwicklungstoxisch beim Menschen wirken kann, sofern nicht Speziespezifitäten im Wirkmechanismus, der Kinetik oder andere biologische Parameter eine Übertragbarkeit auf den Menschen ungerechtfertigt erscheinen lassen.
- Eine generelle Aussage, welche Spezies am besten die humane Situation widerspiegelt, kann derzeit nicht getroffen werden. Das Kaninchen scheint in vielen Punkten besser mit dem Menschen vergleichbar als Nager. Die alleinige Verwendung von Kaninchen führt jedoch häufig zu falsch-negativen Aussagen.

Quantitativer Speziesvergleich

Da die Art der entwicklungstoxischen Effekte in verschiedenen Spezies sehr unterschiedlich sein kann, wurden für die quantitative Interspeziesextrapolation die NOAEL- bzw. LOAEL-Werte für die jeweils empfindlichsten Effekte in den einzelnen Spezies miteinander verglichen.

Bei der Stoffauswahl wurden primär solche Substanzen einbezogen, die in der EU als entwicklungstoxisch Kategorie 1 oder 2 eingestuft sind und bei denen Untersuchungen in verschiedenen Spezies möglichst im gleichen Labor durchgeführt wur-

den. Dabei wurde darauf geachtet, dass die Studien hinsichtlich Expositionspfad, -dauer und Untersuchungstiefe möglichst übereinstimmten. Die vorliegenden Daten wurden dahingehend überprüft, ob bei entwicklungstoxischen Substanzen ein Scaling nach Grundumsatz bei der Interspeziesextrapolation angemessen ist oder andere Gesetzmäßigkeiten gelten.

Entsprechend der Empfehlungen der Testguidelines wurden die meisten Daten in den beiden Spezies Kaninchen und Ratte erhoben. Für eine Reihe von Substanzen liegen aber auch Untersuchungen in Ratten und Mäusen vor. Wegen der phylogenetischen Nähe, des geringen Größenunterschieds und der zum Teil hohen Intraspeziesvariabilität wurde bei der Auswertung auf einen Vergleich Ratte/Maus verzichtet. Vergleiche mit größeren Spezies sind vor allem bei den Substanzen möglich, die sich beim Menschen als entwicklungstoxisch erwiesen haben und deshalb intensiv untersucht wurden wie z.B. Retinoide, Thalidomid und Valproinsäure.

Es wurden insgesamt 74 Speziesvergleiche zu 37 Substanzen (15 Arzneimittel und 22 Chemikalien) nach oraler Applikation und 11 Speziesvergleiche zu 8 Chemikalien nach inhalativer Applikation durchgeführt. Die folgende Tabelle zeigt die Datenbasis und das zusammengefasste Ergebnis der Auswertung:

Tab. 1.3 Quantitativer Interspeziesvergleich zu entwicklungstoxischen Effekten: Gegenüberstellung von LOAEL- und/ oder NOAEL-Werten

	Oraldaten* (ohne humanteratogene Pharmaka ^{a)})	Oraldaten* (humanteratogene Pharmaka ^{a)})	Alle Oraldaten*	Alle Inhalationsdaten
Substanzen n =	32	5	37	8
Speziesvergl. n =	40	34	74	11
NOAEL-/LOAEL-Quotienten n =/n =	25/38	0/34	25/72	8/9
AM	37,53	21,17	31,80	2,20
SD	242,58	46,45	197,00	1,80
GM	0,88	4,86	1,60	1,35
GSD	8,46	6,49	8,98	3,17
50-Perz.	0,75	4,40	1,23	1,43
90-Perz.	8,57	50,94	33,38	4,95

a: Retinol, all-trans-Retinsäure, 13-cis-Retinsäure, Thalidomid, Valproinsäure

* nach Transformation über allometrisches Scaling nach kalorischem Grundumsatz

Eine Auswertung alleine der Daten zu 5 Pharmaka, die sich beim Menschen als teratogen erwiesen haben (3 Retinoide, Thalidomid und Valproinsäure) zeigte, dass durchschnittlich eine Scaling nach Grundumsatz nicht zutreffend ist, sondern für diese Substanzen vielmehr große Speziesunterschiede bestehen. Für die übrigen Substanzen ist ein Scaling nach Grundumsatz durchschnittlich jedoch offensichtlich für die Interspeziesextrapolation geeignet (geometrischer Mittelwert der NOAEL- bzw. LOAEL-Quotienten aus erwartetem und tatsächlich beobachtetem Wert ca. 1,5 über alle Oral- und Inhalationsdaten hinweg). Einzelne Stoffe weisen jedoch eine darüber hinausgehende Variabilität auf, für deren Berücksichtigung ein zusätzlicher Faktor notwendig wäre (bis Faktor ca. 30 bei 90 Perzentil).

Zeitextrapolation

Das Auftreten entwicklungstoxischer Effekte ist abhängig vom Zeitpunkt (kritische Fenster), der Dauer und der Dosis der Exposition. Es wurden deshalb verschiedene Studiendesigns entwickelt, mit deren Hilfe bestimmte Endpunkte adäquat untersucht werden können, z.B. Studien zur Untersuchung pränataler Entwicklungstoxizität zur Erkennung von Terata, intrauteriner Wachstumsretardierungen oder Letalität; Ein- und Mehrgenerationenstudien zur Erkennung von Effekten auf die Fertilität, die zum Teil auch prä- und postnatale Entwicklungsstörungen offenbaren; Untersuchungen zur Entwicklungsneurotoxizität. Da die Expositionszeiten in diesen Studien nicht nur unterschiedlich lang sind sondern auch unterschiedliche Phasen umfassen, ist es bei einem Vergleich dieser Studien häufig schwierig zu unterscheiden, ob eine eventuell auftretende Wirkungsverstärkung Folge der längeren Expositionsdauer oder bedingt durch die andere Expositionsphase ist. Trotz der genannten Schwierigkeiten wurden Studien unterschiedlichen Designs und mit verschiedener Expositionsdauer miteinander verglichen. Dabei zeigte sich, dass

- meist keine Wirkungsverstärkung mit zunehmender Expositionszeit zu beobachten war, wenn sowohl die kurzzeitige als auch die länger andauernde Exposition die empfindliche Zeit während der Gestation umfasste,
- auch in Mehrgenerationenstudien in späteren Generationen meist keine Wirkungsverstärkung gegenüber früheren Generationen zu beobachten war; ausgenommen bei akkumulierenden Substanzen wie z.B. Cadmium und Blei.

Demnach sind bei einer Studie, bei der die Exposition nicht die gesamte Trächtigkeit überdeckt, grundsätzlich Bedenken gegeben, ob quantitativ und qualitativ alle spezifischen Endpunkte der Entwicklungstoxizität erfasst wurden. Ein Zeitextrapolationsfaktor ist jedoch nicht anzugeben. Bei akkumulierenden Substanzen, für die keine Mehrgenerationenstudien vorliegen, erscheint die Verwendung eines Zeitextrapolationsfaktors sinnvoll. Bei der vorliegenden Datenbasis kann dieser jedoch nicht mit hinreichender Sicherheit quantifiziert werden.

Empfindlichster spezifischer Endpunkt

Verschiedene Versuchstierspezies reagieren offensichtlich mit unterschiedlicher Empfindlichkeit auf verschiedene entwicklungstoxische Agenzien. Vergleicht man die Häufigkeit, mit der bestimmte Effekte beim LOAEL in den von uns ausgewerteten Studien auftreten, ergibt sich tendenziell eine bestimmte Rangreihe für Kaninchen bei oraler Schadstoffaufnahme, die sich jedoch von derjenigen bei Inhalation unterscheidet und die von der Rangreihe für Ratten divergiert. Ein paarweiser Vergleich der LOAEL-Daten für die einzelnen Effekte zeigte jedoch, dass keiner dieser Endpunkte sich als besonders empfindlich erwies. Einzeldaten deuten darauf hin, dass möglicherweise der Effekt „funktionale Defizite“ bei niedrigeren Dosen auftritt als andere und Letalität eher unempfindlich ist. Da es sich dabei aber nur um Einzelfallauswertungen handelt, ist für eine abschließende Bewertung eine größere Datenbasis notwendig. Beim gegenwärtigen Stand scheint kein Faktor für die ungenügende Untersuchung eines nicht hinreichend sensitiven Endpunkts erforderlich zu sein, wenn die Exposition während der gesamten Gestationszeit erfolgte.

Mit Hilfe der hier erarbeiteten Grundlagen können methodische Details für die Festlegung von ARW-Werten für reproduktionstoxische Stoffe mit Wirkung auf die männliche Fertilität und Entwicklungstoxizität diskutiert werden.