

Computersystemen zur Vorhersage toxischer Effekte

Stellungnahme der Beratungskommission der Gesellschaft für Toxikologie in der DGPT zum Einsatz von Computersystemen zur Vorhersage und Beurteilung toxischer Effekte von chemischen Stoffen

B. Simon-Hettich, T. Steger-Hartmann, H. Foth, R. Kahl, P.-J. Kramer, W. Lilienblum, D. Schrenk, T. Schulz, H. Schweinfurth

Die anstehende Neuordnung der europäischen Chemikalienbewertungspolitik (REACH) hat zu kontroversen Diskussionen hinsichtlich der zusätzlich erforderlichen Tierversuche geführt. Um steigende Tierversuchszahlen zu vermeiden sowie zur Zeit- und Kostenersparnis, werden tierversuchsfreie Alternativen zur Erfassung toxischer Effekte befürwortet. Der Entwurf der Verordnung vom 29. Oktober 2003 (EU, 2004) sieht einen Einsatz derartiger Methoden vor, ohne sie allerdings genauer zu spezifizieren.

Insbesondere im Bereich der kosmetischen Mittel wird seit 1986 auf EU-Ebene das Ziel verfolgt, Alternativen zu Tierversuchen zu entwickeln und so das Leiden und die Zahl der benötigten Tiere zu verringern. Tierversuche sollen durch „wissenschaftlich zufrieden stellende Alternativen“ ersetzt werden, wenn sie den Verbrauchern ein „gleichwertiges Schutzniveau“ bieten (EU 1986; EU 1993; EU, 2003). Mit der Änderungsrichtlinie 2003/15/EG werden vom Europäischen Parlament und vom Rat Termine festgeschrieben, bis zu denen Tierversuche gänzlich durch alternative Methoden ersetzt werden sollen: Tierversuche im Zusammenhang mit Toxizität nach wiederholter Verabreichung, Reproduktionstoxizität und Toxikokinetik bis März 2013, alle anderen bis März 2009 (EU, 2003).

Die DGPT-Beratungskommission hat diese aktuelle Diskussion zum Anlass genommen, eine Stellungnahme zum Thema computergestützte Vorhersagesysteme in der Toxikologie zu verfassen. Es soll dabei zunächst eine Begriffsbestimmung vorgenommen werden, um dann auf Einsatzgebiete derartiger computergestützter Vorhersagemodelle, deren Grenzen und mögliche zukünftige Entwicklungen einzugehen.

Begriffsbestimmungen

Neben dem Begriff ‚computergestützte Vorhersagemodelle‘ werden synonym auch die Bezeichnungen „in silico toxicology“ oder „predictive tools“ verwendet. Allen hierunter zusammengefaßten Modellen liegt die Erstellung einer Struktur-Aktivitätsbeziehung (structure-activity relationship, SAR)

zugrunde, also die Feststellung von Zusammenhängen zwischen chemischen Strukturen von Substanzen oder Substanzklassen und den durch diese in biologischen Systemen ausgelösten Effekten. Eine genauere Betrachtung der Funktionsweise der Modelle erlaubt die Differenzierung in Expertensysteme einerseits, welche in ihrer Reinform lediglich qualitative Vorhersagen treffen, (Ja/Nein Antwort) und quantitative Struktur-Aktivitäts-basierte Systeme (QSAR) andererseits, welche die Größe des zu erwartenden Effektes vorhersagen.

Expertensysteme beziehen ihren Namen aus der Tatsache, dass hier Fachwissen in ein computerisierbares Regelwerk überführt wurde. Die Vorgehensweise ist dabei vergleichbar mit der Erstellung eines Gutachtens über eine bestimmte Substanz- oder Strukturklasse: Verfügbare Information wird gesammelt, bewertet und gewichtet und dann nach struktureller Ähnlichkeit der Substanzen sortiert. Danach wird die Strukturinformation in eine maschinenlesbare Sprache übersetzt und mit den entsprechenden Informationen über die Substanzen zu einer Datenbank verknüpft („knowledge base“). Das System ist dann in der Lage neue Substanzen auf das Vorkommen von toxikologisch relevanten Strukturen zu überprüfen. Auf der Basis des gesammelten Vorwissen geben Expertensysteme meistens auch Auskunft über den Mechanismus eines Effektes. Der Nutzer ist daher auf der Basis dieser Informationen in der Lage die Relevanz des Warnhinweises, der vom System gegeben wird, zu bewerten. Eine Aussage über das quantitative Ausmaß eines Effektes liefern Experten-Systeme jedoch in der Regel nicht.

QSAR-Systeme hingegen beruhen auf einer statistischen Korrelation zwischen der chemischen Struktur eines Moleküls und dessen biologischen Effekten. Dabei wird bestimmten chemischen oder physikochemischen Eigenschaften, die aus der Struktur der Substanz abzuleiten sind (sogenannte Deskriptoren wie z.B. Elektronendichteverteilung, Polarität, Oktanol-Wasser-Verteilungskoeffizient u.ä.), das quantitative Ausmass eines biologischen bzw. toxischen Effektes zugeordnet und daraus eine Korrelationskurve oder Regressionsgerade errechnet. Im Gegensatz zu Expertensystemen liefern QSAR-Systeme daher eine quantitative Aussage über das Ausmaß eines Effektes. Eine Übersicht über derzeit marktgängige Systeme beider Kategorien findet sich in Greene (2002) und ECETOC (2003).

Beiden Kategorien ist gemeinsam, dass sie auf Datensätzen basieren, die zuvor aus entsprechenden experimentellen in vitro oder in vivo Prüfungen oder entsprechenden Beobachtungen an Menschen für den jeweiligen toxikologischen Endpunkt gewonnen wurden. Die Qualität dieser Datensätze beeinflusst damit unmittelbar die Güte der Vorhersage. Vorhersagesysteme, die aus Datensätzen minderer Qualität abgeleitet werden, resultieren zwangsläufig in mäßigen oder falschen Vorhersagen.

In der Toxikologie wird in der Regel von in vitro oder tierexperimentellen in vivo Prüfungen auf Effekte am Menschen extrapoliert, d.h. bereits hier wird eine Vorhersage getroffen, deren Trefferwahrscheinlichkeit keine 100 % erreichen wird. Im Falle von Industriechemikalien verbieten sich toxikologische Prüfungen am Menschen in der Regel aus ethischen Gründen. Aus prinzipiellen Überlegungen ist daher davon auszugehen, dass die Trefferwahrscheinlichkeit eines computergestützten Vorhersagemodells im Idealfall lediglich den Wert des entsprechenden experimentellen Systems zur Vorhersage von Effekten am Menschen annehmen kann. In der Realität wird sie mehr oder weniger deutlich darunter liegen.

Welche toxikologischen Endpunkte sind einer Vorhersage zugänglich?

Theoretisch können Datensätze zu jedem toxikologischen Endpunkt einer statistischen Analyse unterworfen werden, mit dem Ziel eine neue Regel für ein Expertensystem bzw. eine quantitative Beziehung zwischen Struktur und Wirkung abzuleiten. Wenn der toxische Effekt idealer Weise auf einen bekannten Mechanismus zurückgeführt werden kann, kann eine solche statistische Analyse erfolgversprechend sein. Wenn der Effekt jedoch von einer Vielzahl unterschiedlicher Mechanismen ausgelöst werden kann, so wird eine statistische Korrelation ohne Berücksichtigung der jeweiligen Mechanismen zumindest bei dem derzeitigen Stand der Vorhersagetechniken zu fehlerhaften Intra- und Extrapolationen und daraus resultierend zu nicht verlässlichen Vorhersagen führen.

Mit den heute verfügbaren kommerziellen Systemen werden beispielsweise bereits akzeptable Vorhersagen für den Bereich Mutagenität erzielt (White et al. 2003). Als „akzeptabel“ sind dabei Vorhersagen zu werten, die für die weiter unten definierten Parameter „Sensitivität“, „Spezifität“ oder „Konkordanz“ Werte über 70% liefern. Vergleichsweise gute Vorhersagen werden insbesondere für mutagene Effekte beobachtet, die auf dem Mechanismus kovalenter DNA-Bindung beruhen, deren Folge im Ames-Test als somatische Mutation detektiert wird. Zwei Gründe können für die gute Prädiktivität im Vergleich zu anderen toxikologischen Endpunkten angeführt werden. Zum einen liegen für diesen einfach durchzuführenden und standardisierten Test umfangreiche Datensätze vor, d.h. eine Vielzahl unterschiedlichster Moleküle wurde charakterisiert. Zum anderen handelt es sich hier um einen toxikologischen Wirkmechanismus, der sich annähernd aus den Moleküleigenschaften ableiten lässt: die DNA-Reaktivität steht im Zusammenhang mit der elektrophilen Natur eines Moleküls bzw. dessen Metaboliten, welche sich wiederum aus der Struktur erschließen lässt. Voraussetzung ist allerdings, dass die Metabolisierungswege der Substanz bekannt sind oder ihrerseits verlässlich vorhergesagt werden können. Die Vorhersage von mutagenen Effekten, die nicht auf DNA-Schädigung durch kovalente Bindung zurückzuführen sind (bei klastogenen oder aneugenen

Substanzen), ist jedoch deutlich weniger verlässlich (Snyder et al. 2004). Hier ist auch der Mechanismus weniger eindeutig von der Struktur ableitbar bzw. in vielen Fällen nicht bekannt.

Ähnlich wie DNA-Reaktivität gibt es noch andere toxische Effekte, die sich auf bestimmte physikalisch-chemische Eigenschaften eines Moleküls zurückführen lassen. Beispielsweise wurden Wasser- oder Fettlöslichkeit, Dampfdruck und Oberflächenaktivität als Parameter identifiziert, die zu lokalen Effekten (Reizung) an Haut oder Auge beitragen (Gerner et al., 2000). Des Weiteren ist Proteinreaktivität eine Ursache für die Bildung von allergenen Haptenen im Rahmen der Hautallergisierung (Kimber et al., 2001). Allerdings spielt hier auch die Hautresorption als zweiter Parameter eine entscheidende Rolle, die wiederum maßgeblich durch die physikalisch-chemischen Eigenschaften Molekülgröße und Fettlöslichkeit bedingt ist. Durch Kombination geeigneter Vorhersagesysteme für Resorption und Reaktivität lassen sich auch für diesen Endpunkt akzeptable Vorhersagen erzielen (Barratt 2001).

Endpunkte wie akute Toxizität oder Letalität, chronische Toxizität, Kanzerogenität, Reproduktionstoxizität oder Organtoxizität sind einer Modellierung nur bedingt zugänglich, da die im Versuch erhobenen Beobachtungen eine Vielzahl unterschiedlicher Variablen betreffen und über verschiedenste Mechanismen ausgelöst werden können. Darüber hinaus ist für solche Endpunkte die Datenlage, d.h. die Zahl verfügbarer valider Studien sowie die Variabilität der untersuchten chemischen Strukturen, limitiert, so dass ein universell anwendbares Vorhersagesystem für solche Endpunkte derzeit nicht verfügbar ist.

Lebertoxizität, zum Beispiel, zeigt sich bereits histopathologisch in verschiedensten Erscheinungsbildern, die kaum in einem generellen Vorhersagemodell abgebildet werden können. Daher existieren dazu bisher keine verlässlichen QSAR Systeme. Für den speziellen Mechanismus der speziesspezifischen Peroxisomenproliferation in der Leber wurden hingegen mittlerweile einige taugliche Struktur-Aktivitätsbeziehungen aufgestellt. Da die Peroxisomenproliferation rezeptorvermittelt ist, ist die Rezeptorbindung eine wesentliche Voraussetzung für das Eintreten des toxischen Effektes. Die Bindungstasche ist ihrerseits durch die bekannten Proliferatoren annähernd charakterisiert, so dass die Aufstellung von Struktur-Wirkungsbeziehungen auf dieser Basis möglich ist (Ashby et al. 1994).

Generell spielt die Reproduzierbarkeit der experimentellen Ergebnisse sowie die experimentelle Variabilität der Effekte eine entscheidende Rolle für die Entwicklung von validen Vorhersagemodellen. Der hohe Grad an Variabilität bei Kanzerogenitätsstudien mit Nagern wird beim Vergleich der Ergebnisse von NTP[1]-Kanzerogenitätsstudien mit Ergebnissen aus nicht-standardisierten Prüfungen

deutlich. Dabei wurde für einen Datensatz von >100 Chemikalien eine Reproduzierbarkeit <60% gefunden (Gottmann et al, 2001). Darüber hinaus bestehen Unterschiede bei der Beurteilung der kanzerogenen Effekte (positives oder negatives Ergebnis) und der biologischen Signifikanz der Effekte zwischen verschiedenen Untersuchern, da hier weniger eindeutige Kriterien als beispielsweise bei der Beurteilung der Mutagenität im Ames-Test vorliegen. Dieses zeigt die Bedeutung der Qualität der experimentellen Daten sowie international einheitlicher Bewertungskriterien für die Auswahl der Studien im Rahmen der Modellentwicklung. Modelle, die auf Daten aus nicht-standardisierten Tests oder auf Daten geringer Qualität basieren, werden kaum verlässliche Vorhersagen liefern können.

In Anbetracht der strukturellen Vielfalt chemischer Verbindungen und der großen Zahl möglicher Mechanismen, die bei Langzeitbehandlung von Nagern zu Tumoren führen können, in Verbindung mit dem Fehlen von Daten für große Bereiche des „chemischen Universums“ ist es fraglich, ob die derzeitigen Daten ausreichend als Basis für ein generelles Vorhersagemodell sind. Für Substanzklassen mit gemeinsamem Wirkmechanismus oder ähnlicher chemischer Struktur (kongenere Chemikalien) ist eine erfolgreiche Modellentwicklung jedoch möglich. Die Entwicklung von SARs für große, heterogene Datensets wird allgemein als sehr schwierig angesehen (Richard & Benigni, 2001; Benigni & Passerini, 2002; Clark & Grootenhuis, 2002).

Wie wird die Vorhersagekraft ermittelt?

Die Beurteilung der Vorhersagequalität erfolgt maßgeblich auf der Basis der Parameter Sensitivität (Verhältnis der korrekt vorhergesagten toxischen Substanzen zur Gesamtzahl der toxischen Substanzen) und Spezifität (Verhältnis der korrekt vorhergesagten nicht-toxischen Substanzen zur Gesamtzahl der nicht-toxischen Substanzen) bzw. auf dem zusammengefassten Parameter Konkordanz (Verhältnis der korrekt vorhergesagten Substanzen zur Gesamtzahl der getesteten Substanzen; synonym als „Richtigkeit“, im englischen auch als „accuracy“ bezeichnet) (ECETOC, 2003; Steger-Hartmann, 2004). Eine Optimierung der Vorhersagesysteme in Richtung Sensitivität hat in der Regel eine verringerte Spezifität zur Folge. Für die Vorhersage von Mutagenität heißt dies, dass zwar mehr Substanzen korrekt als Mutagene identifiziert werden, andererseits aber auch die Zahl derer ansteigt, die fälschlicherweise als Mutagene vorhergesagt werden. Je nach Einsatzgebiet muss entschieden werden, welcher der beiden Parameter optimiert werden soll. In der frühen Arzneimittelentwicklung beispielsweise ist eine hohe Spezifität anzustreben, da hier jede falsch positive Vorhersage zu einem Verlust eines wertvollen potentiellen Entwicklungskandidaten führen würde, der in jedem Falle zu einem späteren Zeitpunkt einer ausführlichen experimentellen Prüfung unterliegt. Im regulatorischen Kontext (Sicherheitsbewertung bisher nicht experimentell überprüfter Substanz,

Verbraucherschutz) ist dagegen eher eine optimierte Sensitivität gewünscht, da jede falsch negative vorhergesagte Substanz, für die nicht die entsprechenden Konsequenzen getroffen werden, ein Risiko für den Anwender oder Verbraucher darstellt.

Sensitivität und Spezifität können nicht als absolute Zahlen für ein einzelnes System angegeben werden, sondern sind abhängig vom Testdatensatz, mit dem das System geprüft wird. Diese Feststellung ist im Prinzip vergleichbar mit der in den achtziger Jahren gemachten Beobachtung, dass die Sensitivität des Ames-Tests für die Vorhersage der Karzinogenität am Nager von 90% auf 45% fiel, nachdem sich die zugrundeliegende NTP-Datenbasis von chemisch-hochreaktiven Substanzen zu Verbindungen verschob, die primär aufgrund ihrer Umweltrelevanz getestet wurden (Ashby und Tennant, 1988). Die in der erwähnten Publikation auf der Basis der NTP-Daten aufgestellten sogenannten „structural alerts“ haben Eingang in die meisten Experten-Systeme gefunden. Eine Validierung eines Systems auf der Basis der NTP-Datenbank führt daher zu akzeptablen Werten für Sensitivität und Spezifität. Wird diese Validierungsübung hingegen mit einem anderen Datensatz, z.B. mit zugelassenen Arzneimitteln, durchgeführt, so fallen die Ergebnisse deutlich schlechter aus (Snyder et al. 2004). Die Qualität eines Vorhersagesystems hängt somit maßgeblich davon ab, inwieweit zu testende Verbindungen bereits durch verwandte Verbindungen im System abgebildet sind („coverage“). Bedauerlicherweise lässt sich dies derzeit noch nicht durch einen Zahlenwert darstellen, der die Vergleichbarkeit von Systemen verbessern würde. Ein Vergleich von Sensitivitäts- und Spezifitätsangaben zwischen Systemen ist daher nur auf der Basis identischer Testdatensätze möglich.

Werden computergestützte Vorhersage-Systeme Tierversuche einsparen?

Eine wesentliche Einsparung von Tierversuchen durch die alleinige Verwendung von computergestützten Vorhersagesystemen ist in naher Zukunft nicht zu erwarten. Wie oben geschildert, ist die Mehrzahl der toxischen Effekte derzeit noch schlecht vorherzusagen. Allerdings ist davon auszugehen, dass computergestützte Vorhersagesysteme für einige mechanistisch klarer umrissene Endpunkte mit guter Datenbasis (ätzende Wirkung, Reizung) in Zukunft ihren festen Platz in einem abgestuften Prüfverfahren haben werden (Gerner & Schlede, 2002; Höfer et al. 2004). Sinn dieser abgestuften Verfahren ist es, eindeutig positiv vorhergesagte und von einem Experten ebenfalls so bewertete Substanzen nicht der experimentellen Testung zu unterziehen, sondern nur mehr die als unwirksam vorhergesagten Substanzen tierexperimentell zu überprüfen, um einen optimalen Verbraucherschutz zu gewährleisten. Damit würde zumindest die Zahl der Tiere verringert.

Was ist der derzeitige Stellenwert von computergestützten Vorhersage-Systemen in der Risikobewertung?

Wie bereits oben geschildert, wird von den EU-Behörden angestrebt, dass für einige toxische Endpunkte computergestützte Vorhersagesysteme ihren festen Platz im Rahmen abgestuften Prüfverfahren bekommen werden. Allerdings bedarf es dazu noch erheblicher Entwicklungsanstrengungen, bis von in silico Methoden als verlässlichen und validierten Ersatzmethoden gesprochen werden kann.

Zusammenstellungen über die derzeit bei internationalen Entscheidungsträgern/Behörden verwendeten (Q)SAR-Modelle sind bei Walker et al. (2002) oder Cronin et al. (2003) zu finden. Eine weite Verbreitung hat die Verwendung von Strukturanalogien bei der Bewertung von Chemikalien gefunden. So werden z.B. in der EU, Korea und USA bei der Anmeldung neuer Chemikalien Analogiebetrachtungen zu strukturell ähnlichen Substanzen anerkannt und können zu einem eingeschränkten (öko)-toxikologischen Prüfprogramm und somit zur Reduktion von Tierversuchen führen. Im Technical Guidance Document on Risk Assessment der EU (ECB, 2003) wird für ökotoxikologische Endpunkte eine Vielzahl von Einsatzmöglichkeiten von (Q)SARs beschrieben. Für toxikologische Endpunkte wird der Einsatz von Struktur-Wirkungsbeziehungen im Wesentlichen auf „expert judgement“ beschränkt, welches dann als zusätzlicher Hinweis gewertet wird. Deutlich weiter geht die dänische Umweltbehörde Danish EPA mit dem Einsatz von QSAR-Systemen. In einem umfangreichen Programm wurden für ca. 47.000 chemische Stoffe Vorhersagen zu toxikologischen Endpunkten (akute und subchronische Toxizität, Mutagenität, Kanzerogenität, Reproduktionstoxizität) mit computerbasierten QSAR-Systemen ermittelt. Basierend auf diesen Vorhersagen wurden für diese Stoffe Einstufungs- und Kennzeichnungsvorschläge erarbeitet (Tyle et al., 2002). Im Gegensatz dazu verwenden japanische Behörden (METI, MHWL) keine Analogien oder (Q)SAR-Betrachtungen bei der Beurteilung der toxikologischen Eigenschaften von neuen Chemikalien.

Zur Erhöhung der Akzeptanz von (Q)SAR-Systemen und zur Minimierung des fehlerhaften Einsatzes von solchen Systemen wurden auf dem von CEFIC/ICCA2 organisierten internationalen Workshop Akzeptanzkriterien für Vorhersagemodelle entwickelt (CEFIC, 2002; Jaworska et al., 2003). Eine Validierung verlässlicher Modelle entsprechend diesen sogenannten Setubal-Kriterien wird als Voraussetzung für eine Verwendung im regulatorischen Bereich angesehen.

Validitäts-/Akzeptanzkriterien für Vorhersagemodelle/(Q)SARs sind:

- Sie beschreiben einen definierten toxikologischen Endpunkt
- Sie verwenden einen eindeutigen und einfach anzuwendenden Algorithmus
- Sie sind idealer Weise mechanistisch basiert
- Sie weisen einen definierten Anwendungsbereich auf („domain of applicability“)
- Sie beinhalten ein Maß für die Datenqualität der zugrundeliegenden Datenbasis
- Die Vorhersagekraft wurde in einer externen Validierung ermittelt

Zur Zeit führt ECVAM3 ein Validierungsprogramm entsprechend den Setubal-Kriterien mit den verfügbaren Vorhersagemodellen durch (ECVAM, 2004). Keines der kommerziell verfügbaren Vorhersagesysteme erfüllt derzeit alle der aufgelisteten Kriterien.

Computergestützte Vorhersagesysteme stellen jedoch in jedem Falle bereits heute eine sinnvolle Ergänzung des toxikologischen Repertoires für die Bewertung von Substanzen dar, die sich auf Grund der vorhandenen geringen Menge oder auf Grund schwieriger Synthese einer experimentellen Prüfung entziehen (z.B. Verunreinigungen von Arzneimitteln). Darüber hinaus können sie Hilfestellung beim Screening und bei der Priorisierung von Substanzen geben, wie dies im Rahmen des europäischen REACH-Programmes erforderlich wird. Es ist zu erwarten, dass die bei der Datensammlung im Rahmen des REACH-Programms anfallenden neuen Datensätze für eine Verbreiterung der Datenbasis und damit für eine Anhebung der Vorhersagekraft bestehender QSAR-Programme genutzt werden können.

Literatur

Ashby J, Tennant RW (1988) Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutation Research* 204:17-115

Ashby J, Brady A, Elcombe CR, Elliott BM, Ishmael J, Odum J, Tugwood JD, Kettle S and Purchase IF (1994) Mechanistically-based human hazard assessment of peroxisome proliferator-induced hepatocarcinogenesis. *Human and Experimental Toxicology* 13 (Suppl.): S1-S117

Barratt MD, Rodford, RA (2001) The computational prediction of toxicity. *Current Opinion in Chemical Biology* 5:383-388

Benigni R, Passerini L (2002) Carcinogenicity of the aromatic amines: from structure-activity relationships to mechanisms of action and risk assessment. *Mutation Research* 511: 191-206

CEFIC (2002) Proceedings of workshop on regulatory acceptance of (Q)SARs for human health and environmental endpoints, Setubal, March 4-6, 2002

Clark DE, Grootenhuis PDJ (2002) Progress in computational methods for the prediction of ADMET properties. *Current Opinion in Drug Delivery & Development* 5: 382-390

Combes RD, Gaunt I, Balls M (2004) A scientific and animal welfare assessment of the OECD health effects test guidelines for the safety testing of chemicals under the European union REACH system. *ATLA* 32: 163-208

Cronin MT, Jaworska JS, Walker JD, Comber MHI, Watts, CD, Worth AP (2003) Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. *Environmental Health Perspectives* 111: 1391-1401

ECB (2003) Technical Guidance Document on Risk Assessment, 2nd ed., European Chemicals Bureau, Ispra (<http://ecb.jrc.it/tgdoc>)

EU (1986) Richtlinie 86/609/EWG des Rates vom 24. November 1986 zur Annäherung der Rechts- und Verwaltungsvorschriften der Mitgliedstaaten zum Schutz der für Versuche und andere wissenschaftliche Zwecke verwendeten Tiere. - Amtsblatt Nr. L 358 vom 18.12.1986, S. 1-28.

EU (1993) Richtlinie 93/35/EWG des Rates vom 14. Juni 1993 zur sechsten Änderung der Richtlinie 76/768/EWG zur Angleichung der Rechtsvorschriften der Mitgliedstaaten über kosmetische Mittel. - Amtsblatt Nr. L 151 vom 23.06.1993, S. 32-37.

EU (2003) Richtlinie 2003/15/EG des Europäischen Parlaments und des Rates vom 27. Februar 2003 zur Änderung der Richtlinie 76/768/EWG zur Angleichung der Rechtsvorschriften der Mitgliedsstaaten über kosmetische Mittel. - Amtsblatt Nr. L 66 vom 11.3.2003, S. 26-35.

EU (2004) <http://europa.eu.int/comm/enterprise/reach/index.htm>

ECETOC (2003) (Q)SARs: Evaluation of the commercially available software for human health and environmental endpoints with respect to chemical management applications, Technical Report No. 89

ECVAM (2004) <http://ecvam.jrc.it/index.htm>

Gerner I, Zinke S, Graetschel G, Schlede E (2000) Development of a decision support system for the introduction of alternative methods into local irritancy/corrosivity testing strategies. Creation of fundamental rules for a decision support system. ATLA 28: 665-698

Gerner I, Schlede E (2002) Introduction of in vitro data into local irritation/corrosion testing strategies by means of SAR considerations: assessment of chemicals. Toxicology Letters 127: 169-175

Gottmann E, Kramer S, Pfahringer B, Helma C (2001) Data quality in predictive toxicology: reproducibility of rodent carcinogenicity experiments. Environmental Health Perspectives 109: 509-514

Greene N (2002) Computer systems for the prediction of toxicity: an update. Advanced Drug Delivery Reviews. 54: 417-431

Höfer T, Gerner I, Gundert-Remy U, Liebsch M, Schulte A, Spielmann H, Vogel R, Wettig K (2004) Animal testing and alternative approaches for the human health risk assessment under the proposed new European chemicals regulation. Archiv. Toxicol. 78: 549-564

Jaworska JS, Comber M, Auer C, Van Leeuwen CJ (2003) Summary of a workshop on regulatory acceptance of (Q)SARs for human health and environmental endpoints. Environmental Health Perspectives 111: 1358-1360

Richard AM, Benigni R (2001) AI and SAR approaches for predicting chemical carcinogenicity: survey and status report. SAR QSAR Environ Res 13: 1-19

Snyder RD, Pearl GS, Mandakas G, Choy WN, Goodsaid, F, Rosenblum IY (2004) Assessment of the sensitivity of the commercial programs DEREK, TOPKAT, and MCASE in the prediction of the genotoxicity of pharmaceutical molecules. Environmental and Molecular Mutagenesis 43: 143-158

Steger-Hartmann (2004) Computer-gestützte Vorhersagesysteme. In: Reichl, Schwenk (Eds.): Regulatorische Toxikologie; Springer-Verlag, 97-102

Tyle H, Larsen HS, Wedebye EB, Sijm D, Pedersen Krog T, Niemelä J (2002) Identification of potential PBTs and vPvBs by use of QSARs. Danish EPA, SHC/TS 2-3/029

Walker JD, Carlsen L, Hulzebos E, Simon-Hettich B (2002) Global government applications of analogues, SARs and QSARs to predict aquatic toxicity, chemical or physical properties, environmental fate parameters and health effects of organic chemicals. SAR and QSAR in Environmental Research 13: 607-616

White AC, Mueller RA, Gallavan RH, Aaron S, Wilson AGE (2003) A multiple in silico program approach for the prediction of mutagenicity from chemical structure. Mut. Res. 539: 77-89

1 NTP: National Toxicology Program

2 CEFIC/ICCA: European Chemical Industry Council/International Council of Chemical Associations

3 European Centre for the Validation of Alternative Methods