

## Briefe an den Herausgeber

### Anmerkungen zum Beitrag von Reinhard Länge et al. 'Kriterien zur Bewertung der Qualität und Validität von toxikologischen und ökotoxikologischen Studien für regulatorische Fragestellungen' [UWSF 18 (1) 49–54 (2006)]

Andreas Gies

Umweltbundesamt, Abteilung Risikobeurteilung, D-06844 Dessau ([andreas.gies@uba.de](mailto:andreas.gies@uba.de))

Den Autoren des Diskussionsbeitrages gebührt zweifelsohne Lob. Der Glaube der Öffentlichkeit in die Angemessenheit der Bewertung von Stoffen ist zurzeit tief erschüttert. Gerade auf dem Gebiet der Umwelthormone, das sich wissenschaftlich rasch entwickelt, scheint wenig sicher zu sein und Vieles in der Diskussion. Die Zweifel an der Bewertung von Stoffen durch die Produzenten, die Vermarkter und die Behörden wird noch durch unangenehme Begleitumstände gesteigert, die immer wieder durch die Presse gehen:

- Länge et al. erwähnen die gefälschte Studie, die Arnold et al. (1997) vor neun Jahren zur synergistischen Wirkung von östrogenen Stoffen veröffentlichten.
- Jüngst wurde bekannt, dass die Tabakindustrie jahrelang Teile der wissenschaftlichen Gemeinde täuschte und selektiv Forschung zu Gesundheitsgefahren veröffentlichte, um ihnen genehme Bewertungen des Tabakkonsums zu erreichen (Diethelm et al. 2005).
- Ende letzten Jahres zahlte die Chemiefirma Dupont in den USA eine Strafe von 16,5 Millionen Dollar, weil sie gegenüber der US-Umweltbehörde EPA Gesundheitsrisiken verschwieg, die durch Perfluoroktansäure (PFOA) entstehen können (US EPA 2005).
- David Michaels (2005) beschreibt jüngst eindrucksvoll im Scientific American, wie Industrieunternehmen zur Verteidigung ihrer Produkte planmäßig und erfolgreich versuchen, die wissenschaftliche Unsicherheit zu vergrößern.

Lediglich der Fall der Fälschung von Arnold konnte durch die Mechanismen, die sich die wissenschaftliche Gemeinde zur Qualitätskontrolle selber gegeben hat, aufgeklärt werden. Die mangelnde Reproduzierbarkeit dieser Arbeit zwang zu ihrer Rücknahme.

Andere Mechanismen, wie das Verschweigen von Studienergebnissen oder die nichtadäquate Interpretation der Ergebnisse unterliegen derzeit noch keiner Qualitätskontrolle. Hier die Diskussion zu eröffnen, ist mehr als zeitgemäß.

Beim ersten Lesen des Artikels drängt sich jedoch die Frage auf: Warum definieren die Autoren eine Fülle von Kriterien, die in ihrer Mehrzahl den längst etablierten Regeln guter wissenschaftlicher Praxis entsprechen? Zählt man nach, so enthält der Artikel über 50 Einzelforderungen, angefangen damit, dass das Versuchstier charakterisiert und die

Exposition gemessen sein sollte, bis zu der Forderung nach adäquater statistischer Behandlung der Daten. Warum bedarf es also eines solchen Artikels? Normalerweise sollte es doch ausreichend sein, darauf hin zu weisen, dass Studien, die zu regulatorischen Zwecken eingesetzt werden, den Prinzipien der guten wissenschaftlichen Praxis entsprechen sollten?

Aber so einfach können es sich die Autoren aus vier verschiedenen Industriefirmen nicht machen. Der Großteil der von ihren Firmen eingereichten Unterlagen für die Zulassung und Anmeldung von Stoffen erfüllt diese Kriterien nicht. Die bei den Behörden eingereichten Studien sind im Allgemeinen nicht veröffentlicht und unterliegen somit auch nicht der wichtigen Qualitätskontrolle durch die offene Diskussion in der wissenschaftlichen Gemeinschaft. Die Erstellung dieser Studien findet zwar bisher ausnahmslos nach den Kriterien der Guten Laborpraxis (GLP) statt, doch dies ist lediglich ein Garant dafür, dass wichtige technische Kriterien eingehalten und eine umfangreiche Dokumentation durchgeführt wurden. Die wichtigen, von den Autoren selber aufgestellten Qualitätskriterien wie Relevanz und Reproduzierbarkeit, können im Regelfalle nicht durch die wissenschaftliche Gemeinde überprüft werden.

Die Autoren versuchen daher die Frage der Qualität und der Validität lediglich für die Studien zu diskutieren, die von der unabhängigen wissenschaftlichen Gemeinschaft hervorgebracht wurden und im Regelfall nicht der Interpretationshoheit durch die Industrie unterliegen.

Hinter der Fülle von Forderungen nach Selbstverständlichem verbergen sich indessen einige Aussagen, die ohne Zweifel einer vertieften Diskussion bedürfen. Sinnvoll wäre es zum Beispiel gewesen, Kriterien anzugeben, wie die Reproduzierbarkeit von Studien überprüft werden soll. Dies wäre eine große Hilfe für die tägliche Bewertungsarbeit. Betrachtet man das aktuelle Beispiel der Bewertung des Umwelthormons Bisphenol A, so ist die Frage von Reproduzierbarkeit von Studien entscheidend. Die Autoren erwähnen die Widersprüche, die zwischen den Studien von Saal et al. (1997 und 1998) und denen von Ashby et al. (1999) und Cagen et al. (1999) bestehen. Die erste Autorengruppe fand Effekte bei niedrigen Nagern verabreichten Dosen, die Letzteren fanden keine. Die Folgerung im Diskussionsbeitrag ist: "Wissenschaftlich kann daher dieser

postulierte Effekt nicht als nachgewiesen betrachtet werden". Verschwiegen wird dabei, dass Ashby und Cagen nicht versuchten, die Studien von vom Saal zu reproduzieren. Sie benutzten andere Tierstämme und arbeiteten nach einem anderen Protokoll. Verschwiegen wird auch, dass bis heute mindestens 23 weitere Studien Niedrigdosiseffekte von Bisphenol A bei oraler Gabe an Nagern nachgewiesen haben.

Das Umweltbundesamt hat schon vor Jahren Kriterien aufgestellt, die herangezogen werden sollten, wenn Reproduzierbarkeit untersucht wird. Die Einbeziehung des Erstautors in die Wiederholungsstudie sollte selbstverständlich sein. Das schrittweise Vorgehen, um zuerst vergleichbare Effekte bei den Negativkontrollen und den Positivkontrollen zu erreichen und dann erst in die Prüfung mit den zu betrachtenden Stoffkonzentrationen zu gehen, erscheint uns unabdingbar. Aber sowohl im Falle der toxikologischen Untersuchung von Bisphenol A, als auch der ökotoxikologischen Untersuchung, war es regelmäßig die Industrie, die sich gegen ein solches Vorgehen wendete.

Ebenso wäre es sinnvoll gewesen, Kriterien aufzustellen für die Relevanz und statistische Signifikanz von Ergebnissen. Schauen wir uns wieder das Beispiel des Bisphenol A an: Bei der ökotoxikologischen Bewertung wurden signifikante Effekte auf die Spermienogenese von Fischen (Sohoni et al. 2001) nicht für die Risikobewertung herangezogen, weil diese angeblich keine Relevanz für die Population haben. Bei der humantoxikologischen Bewertung wurden statistisch signifikante Effekte in den Folgegenerationen bei der entscheidenden Nagerstudie (Tyl et al. 2002) verworfen, weil diese durch die Autoren als relevant betrachtet wurden. In beiden Fällen wurden diese Kriterien nach Abschluss der Studie aufgestellt. Zweifelsohne ist es notwendig, Relevanzkriterien vor Beginn der Studie festzulegen und nicht erst nach ihrem Abschluss, wenn die Ergebnisse nicht genehm sind; sonst drängt sich die Befürchtung auf, dass hier unangenehme Ergebnisse retuschiert wurden.

Diese wenigen Beispiele zeigen, wie wichtig es ist, Qualitätskriterien für die gute Bewertungspraxis aufzustellen. Den Autoren gebührt der Verdienst, dass sie hierzu den ersten Schritt getan haben. Auch wenn dieser nicht in die richtige Richtung ging, hat er jedoch Bewegung in das Thema gebracht. Die Diskussion sollte an dieser Stelle fortgesetzt werden.

**Anmerkung:** Der Autor erklärt, keine Interessenkonflikte zu haben.

## Literatur

- Arnold SF, Klotz DM, Collins BM, Vonier PM, Guillette LJ Jr, McLachlan JA (1996): Synergistic activation of estrogen receptor with combinations of environmental chemicals. *Science* 272, 1489–1492
- Ashby J, Tinwell H, Haseman J (1999): Lack of effects for low dose levels of bisphenol A and diethylstilbestrol on the prostate gland of CF1 mice exposed in utero. *Regul Toxicol Pharmacol* 30, 156–166
- Cagen SZ, Waechter JM Jr, Dimond SS, Breslin WJ, Butala JH, Jekat FW, Joiner RL, Shiotsuka RN, Veenstra GE, Harris LR (1999): Normal reproductive organ development in CF-1 mice following prenatal exposure to Bisphenol A. *Toxicol Sci* 50, 36–44
- Diethelm PA, Rielle JC, McKee M (2005): The whole truth and nothing but the truth? The research that Philip Morris did not want you to see. *Lancet* 366, 86–92
- Länge R, Caspers N, Ensenbach U, Pallapies D, Zok S (2006): Kriterien für die Bewertung der Qualität und Validität von toxikologischen und ökotoxikologischen Studien für regulatorische Fragestellungen. *UWSF - Z Umweltchem Ökotox* 18 (1) 49–54
- Michaels D (2005): Doubt is their Product. *Scientific American*, June 2005, 96–105
- Sohoni P, Tyler CR, Hurd K, Caunter J, Hetheridge M, Williams T, Woods C, Evans M, Toy R, Gargas M, Sumpter JP (2001): Reproductive effects of long-term exposure to Bisphenol A in the fathead minnow (*Pimephales promelas*). *Environ Sci Technol* 35, 2917–2925
- Tyl RW, Myers CB, Marr MC, Thomas BF, Keimowitz AR, Brine DR, Veselica MM, Fail PA, Chang TY, Seely JC, Joiner RL, Butala JH, Dimond SS, Cagen SZ, Shiotsuka, RN, Stropp GD, Waechter J (2002): Three-generation reproductive toxicity study of dietary bisphenol A in CD Sprague-Dawley rats. *Toxicol Sci* 68, 121–146
- US EPA (2005): EPA Settles PFOA Case Against DuPont for Largest Environmental Administrative Penalty in Agency History. Press release: <<http://yosemite.epa.gov/opa/admpress.nsf/68b5f2d54f3eefd28525701500517fbf/fdcb2f665cac66bb852570d7005d6665!OpenDocument>>, Zugang: 26. Jan. 2006
- Vom Saal FS, Timms BG, Montano MM, Palanza P, Thayer K, Nagel SC, Dhar MD, Ganjam VK, Parmigiani S, Welshons WV (1997): Prostate enlargement in mice due to fetal exposure to low doses of estradiol or diethylstilbestrol and opposite effects in high doses. *Proc Natl Acad Sci* 94, 2056–2061
- Vom Saal FS, Cooke PS, Buchanan DJ, Palanza P, Thayer K, Nagel SC, Parmigiani S, Welshons WV (1998): A physiological based approach to the study of bisphenol A and other estrogenic chemicals on the size of reproductive organs, daily sperm production and behavior. *Toxicol Indust Health* 14, 239–260