Review

Nonlinearity and thresholds in dose–response relationships for carcinogenicity due to sampling variation, logarithmic dose scaling, or small differences in individual susceptibility

W.K. Lutz\textsuperscript{a,}\textsuperscript{*}, D.W. Gaylor\textsuperscript{b}, R.B. Conolly\textsuperscript{c}, R.W. Lutz\textsuperscript{d}

\textsuperscript{a}Department of Toxicology, University of Würzburg, 9 Versbacher Street, DE-97078 Würzburg, Germany
\textsuperscript{b}Gaylor and Associates, Eureka Springs, AR 72631, USA
\textsuperscript{c}CIIT Centers for Health Research, Research Triangle Park, NC 27709-2137, USA
\textsuperscript{d}Seminar for Statistics, Swiss Federal Institute of Technology, Zürich, Switzerland

Received 15 July 2004; revised 21 January 2005; accepted 21 January 2005
Available online 27 June 2005

Abstract

Nonlinear and threshold-like shapes of dose–response curves are often observed in tests for carcinogenicity. Here, we present three examples where an apparent threshold is spurious and can be misleading for low dose extrapolation and human cancer risk assessment. Case #1: For experiments that are not replicated, such as rodent bioassays for carcinogenicity, random variation can lead to misinterpretation of the result. This situation was simulated by 20 random binomial samplings of 50 animals per group, assuming a true linear dose response from 5\% to 25\% tumor incidence at arbitrary dose levels 0, 0.5, 1, 2, and 4. Linearity was suggested only by 8 of the 20 simulations. Four simulations did not reveal the carcinogenicity at all. Three exhibited thresholds, two showed a nonmonotonic behavior with a decrease at low dose, followed by a significant increase at high dose (“hormesis”). Case #2: Logarithmic representation of the dose axis transforms a straight line into a sublinear (up-bent) curve, which can be misinterpreted to indicate a threshold. This is most pronounced if the dose scale includes a wide low dose range. Linear regression of net tumor incidences and intersection with the dose axis results in an apparent threshold, even with an underlying true linear dose–incidence relationship. Case #3: Nonlinear shapes of dose-cancer incidence curves are rarely seen with epidemiological data in humans. The discrepancy to data in rodents may in part be explained by a wider span of individual susceptibilities for tumor induction in humans due to more diverse genetic background and modulation by co-carcinogenic lifestyle factors. Linear extrapolation of a human cancer risk could therefore be appropriate even if animal bioassays show nonlinearity.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Dose–response relationship; Nonlinearity; Threshold; Hormesis; Bioassay for carcinogenicity; Cancer risk assessment; Pharmacodynamic modeling

Contents

Introduction .................................................. S566
Case #1: Random variation in bioassays for carcinogenicity ........................................ S566
Case #2: Introduction of nonlinearity by logarithmic dose scaling ................................. S567
Case #3: Homogeneous tolerance distribution in animal experiments ........................... S567
Conclusions .................................................. S568
References .................................................. S569

\textsuperscript{*} Corresponding author. Fax: +49 931 201 48446.
E-mail address: lutz@toxi.uni-wuerzburg.de (W.K. Lutz)

0041-008X/S - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.taap.2005.01.038
Introduction

Deviation from linearity in dose–response relationships for tumor induction in rodent bioassays for carcinogenicity is important in the context of extrapolation to low dose. Strongly sublinear (up-bent) curves and apparent thresholds may allow for a rejection of the linear-no threshold (LNT) default assumption and for a discussion of threshold doses and safety factors to derive tolerable exposure levels. This appears to be appropriate if mechanistic considerations can explain the threshold-like shape of the dose–response curve (Lutz, 1998). Here, we draw your attention to situations where an apparent deviation from linearity has no mechanistic support. The first two examples are trivial mathematically but are often overlooked. The third refers to extrapolations when epidemiological data in humans are available.

Case #1: Random variation in bioassays for carcinogenicity

Sampling variability can produce nonlinear shapes even if the true dose response is linear. This is particularly important for bioassays for carcinogenicity, because this test is usually not repeated. We simulated 20 bioassays by random sampling from a binomial distribution, with 50 animals per group and a true linear dose–response relationship with 5–7.5–10–15–25% tumor incidence at arbitrary units of dose 0–0.5–1–2–4. Fig. 1 shows the results. Among the 20 simulations, 4 did not reveal significant carcinogenicity of the chemical (nos. 1, 7, 10, 19), and only 8 showed a more or less linear dose response (nos. 3, 5, 9, 12, 14, 15, 16, 20). Three could be interpreted as threshold-like (nos. 4, 11, 13), 2 showed even a nonmonotonic dose response, with a tumor incidence below control at the lowest dose, followed by a significant increase at high dose (nos. 2, 4). This simulation shows that the hypothesis of a threshold should be tested appropriately. A statistical procedure for this purpose has been suggested recently (Lutz et al., 2002). It is based on the simplest case with the analysis of three data points: (i) the control group, (ii) the highest dose that did not give rise to a significant increase in tumor incidence (NOEL), and (iii) the lowest dose that resulted in a statistically significant increase (LOEL). Based on a minimum chi-square procedure, the probability $P$ was estimated that the observed tumor incidences could be a chance result of an underlying true linear dose response. If this null hypothesis was rejected on a 5% level, the lower limit of a 95% confidence interval for the threshold dose, i.e., the lower bound breakpoint of a hockey stick, was estimated. The procedure runs on the free statistics software package ‘R’ (http://www.r-project.org) and the source file is available at http://www.stat.math.ethz.ch/~lutz/publ/RTPlintest.php. When this test was applied to the 3 simulations that showed the most prominent threshold appearance in Fig. 1 (nos. 4, 11, 13), linearity could

Fig. 1. Simulation of dose–response relationships for a standard rodent bioassay for carcinogenicity. Groups of 50 animals were sampled randomly from a binomial distribution, assuming a true linear dose response with a tumor incidence of 5% in the control group (dose 0) and tumor incidences of 7.5, 10, 15, and 25% at doses 0.5, 1, 2, and 4, respectively. TBA/50, number of tumor-bearing animals in groups of 50 animals. The dashed line indicates the true dose response.
not be rejected for any. The lowest P value came close to significance (0.052 for no. 11, tested for the incidence triplet at doses 0, 2, and 4), which in fact matches the level of false positives expected for 20 simulations.

Case #2: Introduction of nonlinearity by logarithmic dose scaling

Plotting doses or concentrations on a logarithmic scale is a standard procedure for the analysis of pharmacological and toxicological experiments, because the dose steps used are often spaced widely and in logarithmic steps (e.g., concentrations $10^{-6} - 10^{-5} - 10^{-4} - 10^{-3}$ molar; dose steps 0–1–2–4–8). If extrapolation to dose zero is an issue in connection with a toxicological low dose risk assessment, logarithmic transformation of the dose scale can be strongly misleading.

Fig. 2 shows four representations of the same dose-incidence relationship, which is a linear increase from 5% to 15%, 25%, and 45% tumor incidence at dose levels 0, 1, 2, 4. Panel A shows the arithmetic dose scale, panels B and C show the results of logarithmic transformation of the dose axis, with two different dose ranges. The linear dose-response relationship becomes sublinear (panel B). A threshold appears when the dose scale is extended farther to the low dose end, in an attempt to approach the asymptote of the 5% control tumor incidence (horizontal line). This is necessary because log (0) is not defined and the control tumor incidence cannot be represented as a data point. To circumvent this problem, control incidences are sometimes deducted from the dosed groups and increases above control plotted, which results in panel D. Now, it is tempting to make a linear regression for the high dose points and define the intersection with the dose axis as the “threshold dose”, i.e., the dose below which there is no effect. For the example shown in Fig. 2, a “threshold dose” of 0.68 dose units results ($10^{-0.17}$), although panel A gives no justification of deviation from linearity below dose 1.

This type of analysis has recently been suggested by W.J. Waddell as a general procedure for the analysis of bioassay data. Since the control group lies at $-\infty$ on the dose axis, this procedure inextricably produces an intersection with the dose axis and a positive value for a ‘threshold dose’ if the slope of the regression line is positive. Furthermore, it does not consider any statistical criteria. For instance, Waddell derived a ‘threshold dose’ of 28 ppm for the induction of liver tumors by 2-acetylaminofluorene in mouse liver from the 24-month data (converted from his x-intercept at $10^{19.1}$ molecules 2-AAF per kg body weight per day in Fig. 9 in reference Waddell, 2003), while the 30-ppm group of animals exhibited a significant increase in tumor incidence above control (55 tumor-bearing animals in 900 treated mice vs. 9 in 383 controls; $P = 0.0046$ with Fisher’s exact test).

Case #3: Homogeneous tolerance distribution in animal experiments

In a comprehensive review on dose–response relationships for carcinogens (Zeise et al., 1987) the authors said: “The several examples in rodents, even though for high dose data, suggest that nonlinearity is common.” This statement was supported later with the analysis of 315 NCI-
NTP bioassay results, showing that tumor site data were more often consistent with a quadratic response than with a linear response (Hoel and Portier, 1994).

As opposed to the situation in animal experiments, Zeise et al. also stated: “There are few examples of significant nonlinear dose–response relations in humans.” How can this discrepancy be explained? First of all, the number of epidemiological studies that include reliable dose information is much smaller than the number of bioassays. Second, the dose range tested in animals usually includes toxic levels, which increases the probability of synergistic superimposition of carcinogenic mechanisms (Lutz, 1998). Third, the dose estimations in the human studies are associated with a large error, which increases the confidence intervals and decreases the probability of finding a significant higher-order term for the dose response.

Here, we call your attention to another factor that affects the shape of the dose–incidence curve, namely differences in individual susceptibility to the carcinogen (Lutz, 2001). For a group of animals of an inbred strain treated with a dose of a carcinogen that would ultimately result in 100% tumor manifestation, we expect the rate of the process of carcinogenesis to be quite similar for all individuals. They are similar genetically, eat and drink the same, and breathe the same air. On a time axis, this means that tumor manifestation in all individuals within a dose group is expected to occur within a relatively narrow window of time. This has been shown a long time ago with the induction of palpable liver tumors in an inbred strain of rats treated with diethylnitrosamine (Druckrey et al., 1963). In a human population, genetic and lifestyle-dependent factors generate a much wider susceptibility distribution, which results in a flatter increase of the cumulative tumor incidence with time.

The idea is illustrated schematically in Fig. 3, on the basis of hypothetical model assumptions and six groups exposed at dose levels 0–1–2–3–4–5. Time-to-tumor was assumed to follow cumulative normal distributions, with median times-to-tumor span a wider time window, which has been shown a long time ago with the induction of palpable liver tumors in an inbred strain of rats treated with diethylnitrosamine (Druckrey et al., 1963). In a human population, genetic and lifestyle-dependent factors generate a much wider susceptibility distribution, which results in a flatter increase of the cumulative tumor incidence with time.

The idea is illustrated schematically in Fig. 3, on the basis of hypothetical model assumptions and six groups exposed at dose levels 0–1–2–3–4–5. Time-to-tumor was assumed to follow cumulative normal distributions, with median times-to-tumor assumed to decrease in steps of 10 weeks from 130 to 80 weeks with the five dose steps. In panel A, the standard deviation (SD), used here as the measure of the variation between individuals, was assumed to be 10 weeks. The control group is represented by the flattest curve, the curve between individuals, was assumed to be 10 weeks. The most susceptible individual is represented by the steepest curve, the curve between individuals, was assumed to be 20 weeks, respectively. While a clearly sigmoid shape is seen for the homogenous strain (SD = 10 weeks), a much more linear curve is obtained with the groups that were attributed a larger interindividual variability (SD = 20 weeks). A similar difference could be postulated for observations in an animal bioassay vs. epidemiological data in humans.

Genetically determined risk factors may include a heterozygous state for inactivated tumor suppressor genes or activated proto-oncogenes, reduced capacity of DNA repair or fidelity of DNA replication, or an unfavorable situation with respect to the activity of enzymes involved in the metabolic activation and detoxification of carcinogens. Synergistic effects of exposures to other DNA-damaging carcinogens and tumor-promoting factors, particularly by smoking and caloric overnutrition, might also render people more susceptible to incremental exposure to additional genotoxic or epigenetic carcinogens. Individual rates of the process of carcinogenesis are therefore spread much wider in humans as compared with rodents in standard bioassays. Individual times-to-tumor span a wider time window, which results in a more linear dose–incidence relationship.

A nonlinear dose–response relationship observed in a bioassay for carcinogenicity in animals therefore does not necessarily mean that this nonlinearity prevails also in the human population. If epidemiological data in humans are available, linear extrapolation to low dose could still be appropriate. Mechanistic considerations should then focus on the question about the factors that render some individuals more susceptible than others. For formaldehyde, as an example, carcinogenicity in the rat is associated with cytotoxicity to the nasal epithelium. For humans, this could mean that chronic irritation of the nasal mucosa could increase the susceptibility against formaldehyde. Avoiding the co-carcinogenic risk factor(s) might then be a more effective tumor-preventive measure than a general reduction of the formaldehyde reference concentration.

Conclusions

The three cases illustrate that not all ‘demonstrations’ of nonlinearity or thresholds in dose–response relationships for carcinogenicity are real.

- Beware of experiments that have not been replicated.
- Do not consider plots on a logarithmic dose scale if the analysis includes extrapolation to low dose.
Consider the different susceptibility distributions for the discussion of nonlinear dose–incidence curves in animal experiments with regards to a human cancer risk assessment.

References


