600Z92002

federal register

Friday
June 5, 1992

Part V

Environmental Protection Agency

Draft Report: A Cross-Species Scaling Factor for Carcinogen Risk Assessment Based on Equivalence of mg/kg$^{3/4}$/Day; Notice

SM-X3

ENVIRONMENTAL PROTECTION AGENCY

[FRL-4139-7]

Draft Report: A Cross-Species Scaling Factor for Carcinogen Risk Assessment Based on Equivalence of mg/kg$^{3/4}$/Day

AGENCY: U.S. Environmental Protection Agency.

ACTION: Request for comments on the draft report: A Cross-Species Scaling Factor for Carcinogen Risk Assessment Based on Equivalence of mg/kg$^{3/4}$/day.

SUMMARY: Three Federal regulatory agencies, the Environmental Protection Agency, the Food and Drug Administration, and the Consumer Product Safety Commission, are today asking for public comments on the draft report: A Cross-Species Scaling Factor for Carcinogen Risk Assessment Based on Equivalence of mg/kg$^{3/4}$/day.

The report is intended to serve as the basis for a common and unified science policy among these three agencies on a default methodology for determining equivalence of doses—to be used when existing agent-specific data are insufficient for a case-by-case determination—when extrapolating results of rodent carcinogen bioassays to humans.

The public is invited to comment, and public comments will be considered in final revision of the report and in the final adoption of science policies by the participating agencies on cross-species extrapolation of equivalent doses in assessing potential human risks from putative chemical carcinogens.

Commenters are asked to focus on the report's discussion of several issues: (1) The bearing of empirical data on carcinogenic potencies in experimental animals and in humans to the appropriate choice of a dose-scaling methodology; (2) the use of allometric scaling as a means for suggesting appropriate dose scaling methods; (3) the appropriate use of pharmacokinetic and other data in defining a default methodology and particularly in supplanting such default assumptions with case-specific, data-based analysis of dose equivalence; (4) distinguishing the contributions of pharmacokinetic and pharmacodynamic factors to species differences in a carcinogen's potency; and (5) the advisability of adopting the proposed dose-scaling methodology as a common default methodology for the participating agencies.

The complete text of the draft report is published as the last section of this notice.

DATES: The draft document is being made available for public review and comment until August 4, 1992. Comments must be in writing and must be postmarked by August 4, 1992.

INSPECTION AND COPYING: This notice, references, supporting documents, and other relevant materials are available for inspection and copying from the ORD Public Information Shelf at the EPA Headquarters Library, 401 M Street, SW., Washington, DC, Telephone: (202) 260-5926 or FTS: 260-5926. The Library is open daily between the hours of 8 a.m. and 5:30 p.m., except weekends and holidays.

ADDRESSES: Comments may be mailed or delivered to: Project Officer for Cross-Species Scaling Factor Report, c/o Technical Information Staff, Office of Health and Environmental Assessment, U.S. EPA (RD-689), 401 M Street, SW. (room 3703), Washington, DC 20460.

FOR FURTHER INFORMATION CONTACT: Dr. Lorenz Rhomberg, Human Health Assessment Group, Office of Health and Environmental Assessment, U.S. EPA (RD-689), Washington, DC 20460, Telephone: (202) 260-5723 or FTS: 260-5723.

SUPPLEMENTARY INFORMATION: This document reports a consensus reached by representatives of the U.S. Environmental Protection Agency (EPA), the Food and Drug Administration (FDA), and the Consumer Product Safety Commission (CPSC) in discussions conducted under the auspices of the Interagency Pharmacokinetics Group, a workgroup of Federal scientists dealing with issues of common interest arising in the application of pharmacokinetics to chemical health risk assessment. The report is a product of the Interagency Pharmacokinetics Group. It comprises an analysis of empirical and theoretical aspects of the cross-species dose-scaling question, together with an argument for adopting the method of scaling daily administered doses by body mass raised to the ¾ power to achieve presumed equivalence in lifetime carcinogenic risk in different mammalian species. These recommendations have been reviewed and endorsed by the EPA, the FDA, and the CPSC.

If such a policy is adopted, it would replace the current practices in carcinogenic risk assessment of scaling daily administered amounts by body mass (as at FDA) or by body surface area (as at EPA and CPSC). The consensus recognizes that there is considerable scientific uncertainty around any scaling method; it does not claim to have overturned these previous methods with one of superior scientific validity or reduced uncertainty. Rather, in view of the benefits of having the major practitioners of carcinogen risk assessment in the Federal government adhere to a single, consistent methodology, the proposal provides a common default procedure to encourage consistent analyses in cases where agent-specific information is insufficient to suggest appropriate dose-equivalencies on a case-by-case basis. Such case-specific information is always to be preferred to the default methodology proposed herein, and its development and appropriate use are encouraged. Since the scaling methodologies in current use by the agencies participating in this proposal are within the span of scientific uncertainty surrounding the cross-species scaling question, it is not proposed to retroactively change or adjust any risk assessments completed under current policies.

This document has undergone a preliminary interagency review under the auspices of the Ad Hoc Working Group on Risk Assessment of the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET). This request for public comment and a concurrent external scientific peer review will contribute to the development of a final report on this topic. This final report of the Interagency Pharmacokinetics Group will provide the basis for a recommendation of a uniform, default science policy on interspecies scaling for carcinogen risk assessment, to be endorsed by the FCCSET Working Group and used by a broad segment of Federal agencies.

Dated: May 22, 1992.

F. Henry Habicht II,

Deputy Administrator.

Contents

I. Introduction
II. Approaches to Choosing a Cross-Species Scaling Factor
A. Empirical Approach
B. Allometric Approach
1. Species Differences in Pharmacokinetics
2. Species Differences in Pharmacodynamics
3. Toxicological Equivalence
4. A Physiological Time Approach to Toxicological Equivalence
III. Discussion
IV. Conclusions
V. References

A Cross-Species Scaling Factor for Carcinogen Risk Assessment Based on Equivalence of mg/kg$^{3/4}$/Day

I. Introduction

As a matter of necessity, the potential for a chemical agent to cause toxic reactions in humans is often

investigated by exposing and observing the reactions of experimental animals, usually rats and mice. This practice rests on the high degree of physiological, biochemical, and anatomical similarity among mammalian species; the biological reactions in the experimental animals may be taken as evidence that humans might show similar responses to the same chemical exposures. When the objective is to use the animal data to predict the degree or probability of response in humans—that is, when the aim is quantitative extrapolation—one must define the dose levels for humans and animals that are expected to produce the same degree of effect. For this, it is necessary to take into account the pronounced difference in *scale* between the tested model organisms and humans. That is, even if fundamental similarity is presumed, one must allow for the fact that humans are much larger than experimental rodents and will experience chronic exposure to a toxicant for a longer lifetime.

Defining such "toxicologically equivalent" doses has been problematic. Alternatives that have found use include scaling daily administered amounts by body weight or by body surface area; scaling cumulative lifetime intake by body weight; equating exposures to contaminated air, food, or water according to the concentration of toxic agent; and others. Despite considerable study and debate (Pinkel, 1958; Freireich et al., 1966; Mantel and Schneiderman, 1975; Rall, 1977; Hoel, 1977; Hogan and Hoel, 1982; Calabrese, 1983, 1987; Crump et al., 1985; Davidson et al., 1986; Gillette, 1987; Vocci and Farber, 1988; Hill et al., 1986), no alternative has emerged as clearly preferable, either on empirical or theoretical grounds. The various Federal agencies conducting chemical risk assessments have developed their own preferences and precedents for cross-species scaling methodology. This variation stands among the chief causes of variation among estimates of a chemical's potential human risk, even when assessments are based on the same data.

The variety of cross-species scaling methods in use correctly reflects the uncertainty about the best procedure, but the resulting disagreement in risk estimates results in some awkwardness in the regulatory arena. Increasingly, regulatory procedures are being mandated that establish decision points contingent on whether a certain human risk level is to be expected according to "generally accepted" risk assessment procedures. Variation in methodology frequently leads to ambiguity as to whether regulatory action should take place. It has therefore become important to resolve differences in cross-species scaling assumptions.

A second impetus for reexamining the scaling question comes from the increasing availability of comparative pharmacokinetic information on toxic agents. Pharmacokinetic analysis uses data on absorption of agents into the body, distribution among the tissues, metabolic activation or detoxification, and elimination to develop a picture of the disposition of a dose by the body and consequent exposure of the actual target tissues of toxic action. Pharmacokinetic differences among species clearly contribute to the magnitude of equipotent doses. However, the appropriate use of such information for the dose equivalency question hinges on resolving the role of pharmacokinetics compared to that of species differences in the magnitude of toxic reaction to a given degree of target-tissue exposure (i.e., "pharmacodynamics"). Distinguishing the roles of these two aspects of potency scaling has been hampered by imprecisely articulated rationales for the various methods.

In view of the above considerations, the Federal agencies with primary responsibility for conducting chemical risk assessments have endeavored to define a uniform cross-species scaling methodology and rationale for use when extrapolating results of rodent carcinogen bioassays to humans. Discussions and debate on the issues have been held under the auspices of the Interagency Pharmacokinetics Group (IPG), an ongoing workgroup of Federal scientists that deals with issues of common interest arising in the application of pharmacokinetics to risk assessment. The present report is a product of the Interagency Pharmacokinetics Group, and represents a statement of the consensus recommendation resulting from these discussions.

The consensus is that, in the absence of adequate information on pharmacokinetic and sensitivity differences among species, doses of carcinogens should be expressed in terms of daily amount administered per unit of body mass raised to the ¾ power. Equal doses in these units (i.e., in $mg/kg^{3/4}$/day), when experienced daily for a full lifetime, are presumed to produce equal lifetime cancer risks across mammalian species. This proposed scaling method has the advantage of being intermediate between the two currently used methods (scaling daily amount by body mass or by body surface area). It is not merely a compromise; it is as well supported by the empirical data on carcinogen potencies in animals and humans as the methods it would replace. It also has an explicit rationale (the concept of species-independent "physiological time") that may be derived from principles of interspecific allometric variation in anatomy, physiology, and pharmacokinetics. That is, it can be interpreted as a correction for readily observable scale differences among species as their essentially similar biology varies in a regular quantitative way as a function of size.

The consensus does not pretend to have solved the underlying scientific issues. Former methodologies have not been shown to be in error; the consensus should not be construed as overturning previous assumptions and replacing them with one of superior scientific validity. Rather, the consensus achieves the benefits of having all Federal risk assessments adhere to a single, consistent methodology that is in accord with current scientific knowledge on the scaling question. Moreover, the method corresponds to a fully articulated rationale with explicitly stated assumptions about the roles and interactions of various underlying determinants of carcinogenic potency. This aids in consistent and scientifically appropriate application. Furthermore, as information is gained on how the biology of carcinogenesis varies among species, it will be clearer how the arguments and previous presumptions should be modified to accommodate these new insights.

The balance of this document reviews the evidence and arguments that may be adduced to address the question of cross-species scaling of equally carcinogenic doses, and outlines the support for the recommended position of equipotent doses in terms of $mg/kg^{3/4}$/day.

II. Approaches to Choosing a Cross-Species Scaling Factor

There are two broad and complementary approaches to choosing a cross-species scaling factor. The first is empirical; one may seek cases in which human epidemiologic data allow a direct estimate of an agent's potency, and then investigate the success of various scaling methods in predicting that potency from animal data. The second approach is theoretical, and is grounded in the principles of allometry, which is the study of the regular variation in features of anatomy and physiology as a function of overall body size. The strategy for this second

approach is to develop a scientific rationale for a particular scaling factor by investigating the allometric variation of the biological features and processes that influence and underlie carcinogenic potency.

Clearly, in many cases there will be agent-specific ways in which humans and experimental animals differ in a nonsystematic fashion. These may include metabolic activation or detoxification, interaction with key receptors or target molecules, and others. Such factors create unpredictable deviation from the general pattern of scaling, and must be discovered and accounted for on a case-by-case basis. The factor proposed here is a *default* scaling factor, by which is meant one that is to be applied in the absence of adequate case-specific information. Lacking such information, one provisionally assumes that the agent in question is an example of a "typical" or "average" chemical that follows a general pattern of cross-species potency differences. This presumption may be modified as information becomes available, but the default assumptions still serve as the benchmark against which the new information is evaluated.

A. Empirical Approach

This approach attempts to find a factor value that is empirically successful in producing good estimates of potency in humans from data on potencies in other species. The underlying reason why such a factor works is a secondary consideration. The advantage of an empirical approach is that, by directly examining carcinogenic potencies (rather than influences on potency, such as pharmacokinetics), all relevant factors are included. The disadvantage is that the data are few and of low resolution. One must hope that the agent-specific factors, mentioned above, average out to give a good estimate of the general relationship.

A number of studies have sought general scaling factors empirically. Freireich et al. (1966), testing and extending the suggestion of Pinkel (1958), examined maximum tolerated doses (MTDs) of 18 antineoplastic drugs in mice, rats, hamsters, dogs, monkeys, and humans. LD_{10}s were used for rodents, and were presumed to be an equivalent level of toxicity to an MTD. Doses from experiments of different length were reexpressed in terms of an exposure regimen of 5 consecutive days, on the assumption that cumulative dose is proportional to effect. The authors concluded that, when doses were expressed as mg/m^2 body surface area/day, good predictions of human MTDs were obtained from all animal species, but that body weight scaling of doses overpredicted human MTDs (i.e., underpredicted potency in humans) by a margin that increased as one extrapolates from smaller and smaller species. Since an MTD is intended to be a dose causing no lethality, while an LD_{10} causes 10% lethality, the equivalence of these two end points can be questioned. Antineoplastic drugs typically have very steep dose-response curves, however, and survival near the MTD is maintained by close monitoring and intervention, which the rodent LD_{10} determination lack.

Collins et al. (1986, 1990) have found that the human MTD for 16 antineoplastic drugs is well predicted on average by the mouse LD_{10} when doses are expressed as mg/m^2 of body surface area. (If the MTD is considered to be a less severe end point, in such comparisons potencies in the larger species are overestimated vis-à-vis those in rodents; a bias would then be created that would increase the apparent success of surface area scaling compared to scaling by body weight.) That is, if these endpoints of acute toxicity are taken as equivalent, scaling doses in proportion to surface area tends to equalize toxicity across species. Moreover, Collins et al. (1990) compared the blood levels (in terms of the areas-under-the-curve of concentration in plasma as it declines over time, or "C x T") that correspond to equally toxic administered doses and found that these were an even better predictor, in that they displayed less case-by-case variation. These results illustrate three points that are returned to in Section B, below: (1) Scaling administered doses in this way tends to equalize blood levels across species; (2) areas-under-the-curve of blood concentration can serve as a predictive measure of the toxic response to a dose, even across species; and (3) obtaining pharmacokinetic data on internal dose measures can increase the precision of the cross-species prediction of equivalently toxic doses by accounting for case-by-case variation.

Travis and White (1988) reanalyzed the Freireich et al. (1966) data set and nearly doubled the number of drugs by adding a similar data set of Schein et al. (1979). Instead of simply examining the success of prevously proposed scaling methods, they used regression techniques empirically to determine the optimal power of body weight to achieve the best fitting allometric relationship of MTDs across species. For both data sets individually and for the combined data set, a power of 0.72 to 0.74 led to the best cross-species predictions. In the analysis of the combined data, a power of unity (body weight scaling) was clearly rejected at the 95% level of significance, and a power of 2/3 (surface area scaling) was barely rejected. The authors discuss the history of empirical studies of allometric variation in a number of physiological features, primarily basal metabolism, and arque that their result is part of a general empirical support for scaling by the 3/4 power of body weight.

The difficulty with applying these studies to the present question is that they address acute systemic toxicity of a rather narrowly defined type rather than carcinogenesis. Although dose-scaling for different toxic end points should have some features in common (notably pharmacokinetics), it is not altogether clear how lifelong risks that accumulate over time (such as cancer risk) should relate to short-term toxicity dependent only on immediate insults to target tissues.

Some empirical studies of comparative potencies of carcinogens in different species have been done. Such studies face the difficulty of precisely determining potencies in humans based on epidemiologic data. There is also some ambiguity in defining potencies in animals, owing to the variations in rout of exposure, sex and strain differences, varying experimental designs, and so on. Nonetheless, such studies represent the direct investigation of the question at hand.

The National Academy of Sciences (NAS, 1975) examined the potencies of six carcinogenic agents in bioassays using mice and rats and from human epidemiologic studies. They recommended as a dose measure cumulative lifetime amount of agent administered (in mg) per kg body weight. Such scaling is more "conservative" (i.e., predictive of higher human risk from animal results) than either surface area scaling or body weight scaling (from which it differs by a factor of 35, owing to the lack of adjustment for differences in length of lifetime). The NAS conclusion was not based on formal quantitative comparison with surface area scaling ($mg/kg^{2/3}/day$) or body weight scaling.

The paucity of carcinogen potencies in humans known directly from epidemiologic data limits the precision of such comparisons. Crouch and Wilson (1979) instead investigated dose scaling between rats and mice in about 70 ingestion cancer bioassays from the National Cancer Institute testing program. They measured potency by the parameter of a fitted one-hit dose-response model (in units of risk per mg/

kg/day), focusing on the tumor site/type producing the greatest potency (excluding testicular tumors in Fisher 344 rats, and skipping cases in which potency was less than twice sensitivity in either species). A geometric mean of potencies in each sex (which were highly correlated) was used. Interspecies comparisons were based on the best-fitting line of unit slope on a plot of the logarithm of potency in rats against the logarithm of potency in mice. The intercept of such a line gives the geometric mean of the factor by which the rat potency must be divided to give the mouse potency. Body weight scaling predicts a factor of one (i.e., equal risk per mg/kg/day in both species) while surface area scaling predicts a factor of about 2.1 to 2.3, depending on the exact body weights. (For comparison, the scaling by $mg/kg^{3/4}/day$, as advocated herein, predicts a ratio of about 1.8 or 1.9.) The results depend on the strain of rat used. In the 17 cases of comparison between Osborne-Mendel rats and B6C3F1 mice the mean ratio of potencies was 0.40; these rats were somewhat less sensitive than mice, contrary to the expectations of both scaling methodologies. When Fischer 344 rats were compared to the same mouse strain (18 cases) a mean ratio of 4.5 was obtained, indicating that rats were even more sensitive than surface area scaling would expect. (A geometric mean of these two ratios is 1.3. To attempt definition of a general mammalian cross-species allometric relationship using only two species is fraught with pitfalls, especially when they are as close in size as are rats and mice. Nonetheless, for the purposes of this discussion one may note that, using typical body weights—70 kg for a human, 40 g for a mouse, 467 g for a rat of unspecified strain, 500 g for an Osborne-Mendel rat, and 360 g for a Fischer rat—the ratio of 1.3 implies scaling by body weight to the 0.89 power.)

Crouch and Wilson (1979) also examined ratios of rodent potency to epidemiologically derived human potency, comparing "insofar as possible" studies with the same route of exposure and duration in fraction of a lifetime. Owing to imprecision in the epidemiologically based human estimates, no precise curve fitting was attempted, but the authors state that humans appear to be more sensitive to a mg/kg/day dose by about a factor of 5 compared to either rats or mice. (Using the typical body weights listed previously, a factor of 5 corresponds to scaling doses by a power of body weight of 0.7 and 0.8 based on the rat and mouse results, respectively.)

A similar comparison of rats and mice, based on an expanded base of 187 NCI bioassays, was conducted by Crouch (1983). (Despite the larger original database, there were only a few more chemicals in the final analysis, apparently owing to more stringent requirements for significance of portency estimates.) Again, the rat strain influenced the results: for Osborne-Mendel rats the mean ratio was 0.63 while for Fischer 344 rats it was 2.29. (A geometric mean of these two ratios is 1.20.) Separate analysis of males and females changed these ratios only slightly. An analysis irrespective of rat strain yielded a ratio of 1.62. (Using the typical body weights listed previously, rations of 1.20 and 1.62 imply scaling by body weight to the 0.92 and 0.80 power, respectively.)

Gaylor and Chen (1986) examined data on rats, mice, and hamsters in the extensive database of Gold et al. (1984) on TD_{50}s, the dose (in mg/kg/day) leading to a halving of the actuarially adusted percentage of tumor-free animals at the end of a standard lifespan. The tumor site/type showing highest potency (i.e., lowest TD_{50}) was chosen to represent the species, and only agents with responses in both species were included. For 190 compounds administered in the diet, the geometric mean ratio of TD_{50}s in rats and mice was $0.455=1/2.20$. That is, rats were on average about 2.2-fold more sensitive. (Using the typical body weights listed previously, this corresponds to scaling by body weight to the 0.68 power.) Ratios for other routes of exposure varied somewhat, although based on much lower sample sizes than the ingestion results cited above. By gavage, 32 compounds had a mean ratio 1/1.32, in drinking water 10 compounds had a mean ratio of 1.45 (i.e., rats were less sensitive), and by inhalation 7 compounds had a mean ratio of 1/11.2 (i.e., rats were much more sensitive).

Chen and Gaylor (1987) investigated NCI/NTP cancer bioassays of compounds administered orally to rats and mice. They compared "virtually safe doses" (VSDs), defined as doses associated with a lifetime cancer risk of one in a million. These were determined by the method of Gaylor and Kodell (1980), i.e., a linear extrapolation was conducted from an upper bound on a fitted multistage model dose-response curve. Thus, both the rat and mouse VSDs are in some sense "upper bounds." Chemicals were included if judged by the NTP to be positive in at least one species, and when in only one, if there was at least a positive trend in the other species for the same tumor site/type. Unlike the studies mentioned above, Chen and Gaylor (1987) focused on Correspondence of VSDs at the *same* site and sex across species. VSDs were expressed in terms of concentration (parts per million [ppm]); as discussed further in the following section on allometry, since intakes of contaminated media (air, food, water) tend to be proportional to body surface area, the expectation from surface area scaling is that VSDs expressed in ppm would be about equal across species, while body weight scaling would expect a ratio of rat to mouse VSDs to be slightly greater than 2. Again, the results depend on the strain of rat used: For Fischer 344 rats the mean ratio is 1.15, for Osborne-Mendel rats it is 1.68, and for Sprague-Dawley rats it is 1.78. Ignoring rat strain gives a mean ratio of 1.27. These results are intermediate between the expectations of surface area and body weight scaling. For ease of comparison with other studies, one may convert these ratios from a ppm basis to a mg/kg/day basis using empirically based daily food and water consumption patterns in rats and mice (for food, 5% and 13% of body weight for rats and mice, respectively, and for water, 7.8% and 17% [U.S. EPA, 1984]). On a mg/kg/day basis, the rat:mouse VSD ratios are 0.44–0.53 for Fischer rats, 0.647–0.771 for Osborne-Mendel rats, and 0.69–0.82 for Sprague-Dawley rats. (The range reflects using rat:mouse ratios of water and food consumption, respectively, which differ slightly.) Using the typical body weights listed previously, and assuming a weight of 540g for Sprague-Dawley rats, these ratios correspond to scaling doses by body weight to the 0.63–0.71 power (when based on Fischer rats, which constituted most of the cases), 0.83–0.90 power (when based on Osborn-Mendel rats), and 0.86–0.92 (when based on Sprague-Dawley rats).

Metzger et al. (1989) expanded Crouch's (1983) earlier data set by including all 264 cases from the Gold et al. (1984) database in which a significant TD_{50} was obtained in an oral study of rats and mice (of any strain), i.e., including studies that were not in the NCI/NTP database. A best-fitting line of unit slope showed a TD_{50} ratio of 1.46 between mice and rats. This is intermediate between the ratio of 1.0 expected from body weight scaling and 2.5 from suface area scaling (using the authors' assumptions about body weights—this implies a power of body weight of 0.86).

A major study of animal-to-human extrapolation of cancer potencies was carried out by Allen et al. (1987), and reported on by Crump et al. (1987, 1989) and Allen et at. (1988). Twenty-three chemicals were identified that permitted quantitative evaluation of potency in humans and in animals. "Risk-Related Doses" (RRDs) were calculated, defined as the average daily dose per kg of body weight that would be expected to result in an extra cancer risk of 25% over a lifetime. Chemicals were included even if RRD estimates were "infinite" for one species, as happens when no carcinogenic effect is observed. Unlike the studies reviewed above, the Allen et al. (1987) study considered a large number of alternative ways of representing the potency in animals as well as various methods for extrapolating the resulting RRDs to humans. Alternative sets of "risk assessment assumptions" restricted the animal database according to various criteria of experimental design, route of exposure, and tumor type. Different levels of averaging results over experiments, sex, and species were tried. Finally, different methods for combining the multiple animal results on a given chemical into a single measure of its "potency in animals" were examined. This complexity allows an admirably comprehensive look at animal-to-human extrapolation, but it also makes manifest a problem that is latent in the other extrapolation studies: The performance of a scaling factor depends on how the animal potency is characterized. A factor that tends to overpredict human risk can be "rescued" by a method for characterzing animal potency that tends to produce a low estimate, and vice versa.

When the objective is to examine alternative dose-scaling factors, it would seem that the best approach is to examine analyses that aim at broadly based and unbiased estimates of the potency in animals. Risk assessment practices such as using upper bounds on dose-response curves and extrapolating from the most sensitive sex and species of animal are explicitly conservative; they may be appropriate science policies for regulatory purposes, but when the issue is empirically to choose a best-performing scaling factor, they introduce a bias, favoring a less conservative factor to compensate for their conservatism and restore a good prediction of the known human potency.

To compare potencies, Allen et al. (1987) fit a line of unit slope to the data of epidemiologically observed log RRD in humans plotted against the predicted human log RRD based on the animal data and the chosen scaling methodology. The intercept of this line gives an average ratio of the observed to predicted potency, with a ratio of unity indicating unbiased prediction. The analyses discussed prominently in the Allen et al. (1987, 1988) and Crump et al. (1987, 1989) reports show that body weight scaling leads to a ratio of approximately one to somewhat less than one depending on the particular suite of risk assessment assumptions chosen (i.e. slightly underpredicting human risk), while surface area scaling overpredicts human risk several-fold.

These results are sometimes cited as tending to support mg/kg/day scaling, but such a conclusion should be tempered. The particular choice of risk assessment assumptions (among many examined) in the widely cited analysis is the one with results least favorable to surface area scaling; most of the alternatives discussed by Allen et al. (1987) show that body weight scaling underestimates human risks by about the degree to which surface area overestimates it. Moreover, these analyses contain a bias of the sort outlined above—the animal potency for a chemical is characterized by the median of the *lower bounds* on the RRDs for the various animal data sets rather than on best estimates. At present it is unresolved how much the use of central estimates of animal risk to predict central estimates of human risk—a more appropriate analysis for resolving the scaling factor—would shift the results toward favoring surface area scaling.

Two additional studies of comparative cancer potencies should briefly be mentioned, both favoring a somewhat more conservative scaling factor. Raabe et al. (1983) compared bone cancer risks from radium in watch dial painters (who ingested radium by tipping brushes on their tongues) and in beagle dogs exposed to radium by injection. Doses were measured as dose to bone of deposited radium, so this camparison can be seen as lacking the pharmacokinetic component of cross-species differences. Potency was measured by the relative mean degree of life-shortening as a function of does. The authors argued that a cumulative lifetime radiation dose per unit of bone seemed to give good correspondence between human and dog. This result could be related to mg/kg/lifetime scaling for chemical agents.

Kaldor et al. (1988) examined carcinogenic potency of five antineoplastic drugs, using potencies derived from bioassays in rodents and from the secondary tumors the drugs caused in human cancer patients. They argued that potency seemed to be related to total cumulative lifetime exposure per kg of body weight.

The empirical evidence on cross-species scaling of carcinogen potencies can be summed up as follows. The correlation of agents' potencies across species is clearly and strongly demonstrated. This correlation extends to humans, so far as is ascertainable from the limited number of agents for which potencies can be estimated epidemiologically. There is a remarkable agreement among studies that the dose-scaling methods in current use span a range that appears approximately correct. The resolution of the data available at present, however, does not permit a clear choice between surface area and body weight scaling. Empirically chosen scaling factors tend to fall in between these two choices in most cases, but the specific results depend on the laboratory strains used, route of administration, details of the methods for characterizing the carcinogenic potency in animals, and the statistical methods used in curve fitting. The data seem consistent in indicating that body weight scaling somewhat underestimates risks in larger species. The exception is when Osborne-Mendel or Sprague-Dawley rats are compared to B6C3F1 mice, in which comparison the rats are seen to be less affected even by doses scaled to body weight. The preponderance of data are from Fischer 344 rats, however, and this is the strain used in most modern bioassays.

Several points should be borne in mind while interpreting the empirical scaling data. First, although several studies are reviewed, they overlap considerably in their databases; the individual studies are not independent tests. Second, the specific results of a study depend on details of the methodology. The Allen et al. (1987) study showed that whether potencies were averaged over sexes, whether both benign and malignant tumors were counted, whether projections were made for specific tumor sites or for the most potent site, and other such factors could swing the analysis toward favoring one scaling method or another. It is hard confidently to identify and isolate the specific contribution of dose scaling among the many factors that contribute to the final predictions of human risk. Third, the epidemiologically based human potencies that serve as "targets" for the animal-based extrapolations are themselves very uncertain and, as in the animal data, dependent on the specifics of the methodology used in their

estimation. As a result of this and of the previous point, the comparability of animal- and human-based potencies may be problematic. (For example, potencies calculated from human data are usually based on cancers that were the cause of death following partial lifetime exposure, while animal-based estimates usually reflect incidental as well as fatal tumors arising after full lifetime exposure.) A final point to be borne in mind is that the report empirically derived factors represent averages over large numbers of cases. Although the means vary over a narrow range, the individual chemicals show ratios of potencies in different species that span orders of magnitude. Most of the rat-to-mouse comparisons were within an order of magnitude of the average scaling relationship, but several agents showed a 100-fold difference. Variances of rodent-to-human potency ratios were higher, reflecting the uncertain determination in humans and the lack of standardized experimental design. The existence of this scatter of cases around the mean helps to define the limits to the resolution of any scaling method and emphasizes the importance of case-to-case variation. Moreover, it provides some insight into the distribution of uncertainty in the cross-species dose extrapolation step of risk assessment.

Despite these shortcomings, the empirical data support the general practice of scaling rodent potencies to humans, and show that, on average, the current methods perform satisfactorily. Certainly, any method that produces average results an order of magnitude higher or lower than the range represented by body weight and surface area scaling would be in contradiction to the empirical data. The data suggest that a scaling factor in between the surface area and body weight scaling can be considered to have empirical support.

B. Allometric Approach

The complement to the empirical investigation of potency scaling is a more theoretical approach that seeks to identify the biological factors whose variation underlies the variation in a carcinogen's potency across species, and then attempts to adjust for their effect. Clearly, these factors are numerous and, for the most part, poorly understood. Fortunately, there are some rather simple and general quantitative patterns in the variation of many features of anatomy and physiology across differently sized mammalian species, representing broad trends in the way the essentially similar mammalian system operates in large and small editions. Although specific processes acting on specific chemicals can (and do) deviate from these broad trends, it is argued below that the general patterns can provide a benchmark that expresses the expectation about a chemical's carcinogenic potency in small mammals such as experimental rodents and larger ones such as humans. This expectation can be refined (or refuted) by case-specific biological and mechanistic data, when available, showing how the actual processes of metabolism and carcinogenesis differ from the presumptions of the broad trend analysis that serves as the default.

The aim of a dose-scaling methodology is to estimate administered daily doses to experimental rodents and humans that result in equal lifetime cancer risks. That is, the scaled doses are intended to be "toxicologically equivalent." It is useful to recognize two components to this equivalence. The first, which might be termed "pharmacokinetic equivalence," concerns adjustment of the administered dose to a rodent or human so that the corresponding tissues that constitute the targets of the agent's toxicity receive similar exposures to the toxin. The second, or "pharmacodynamic equivalence," relates to the relative tissue doses that, when experienced daily for a lifetime, yield equal lifetime cancer risks. This latter aspect includes, but goes beyond the question of "sensitivity" to address species differences in the operation of the carcinogenic processes as they relate to tissue does. For both the pharmacokinetic and the pharmacodynamic component, scaling questions arise and the problem of defining "equivalence" must be faced.

By way of illustration, consider a hypothetical agent with rather simple pharmacokinetics (first order elimination from a single compartment) given by intravenous injection to a mouse and a human. As shown in Figure 1, such a compound will demonstrate an almost instantaneous peak in its blood concentration, followed by exponential decline. If the administered doses are equal in terms of mg/kg body weight, the peak concentrations are the same in the mouse and the human, but the mouse rids itself of this body burden faster, owing to its more rapid metabolism and elimination compared to the human. As a result, the area under the curve (AUC) of blood concentration as it declines with time is much less in the mouse. If the amount injected is properly adjusted, as illustrated in Figure 2, a concentration profile can be achieved in which the initial peak blood concentration is much less in the human, and yet is balanced by the compound's longer persistence to generate an AUC equal to that of the mouse.

BILLING CODE 6560-50-M

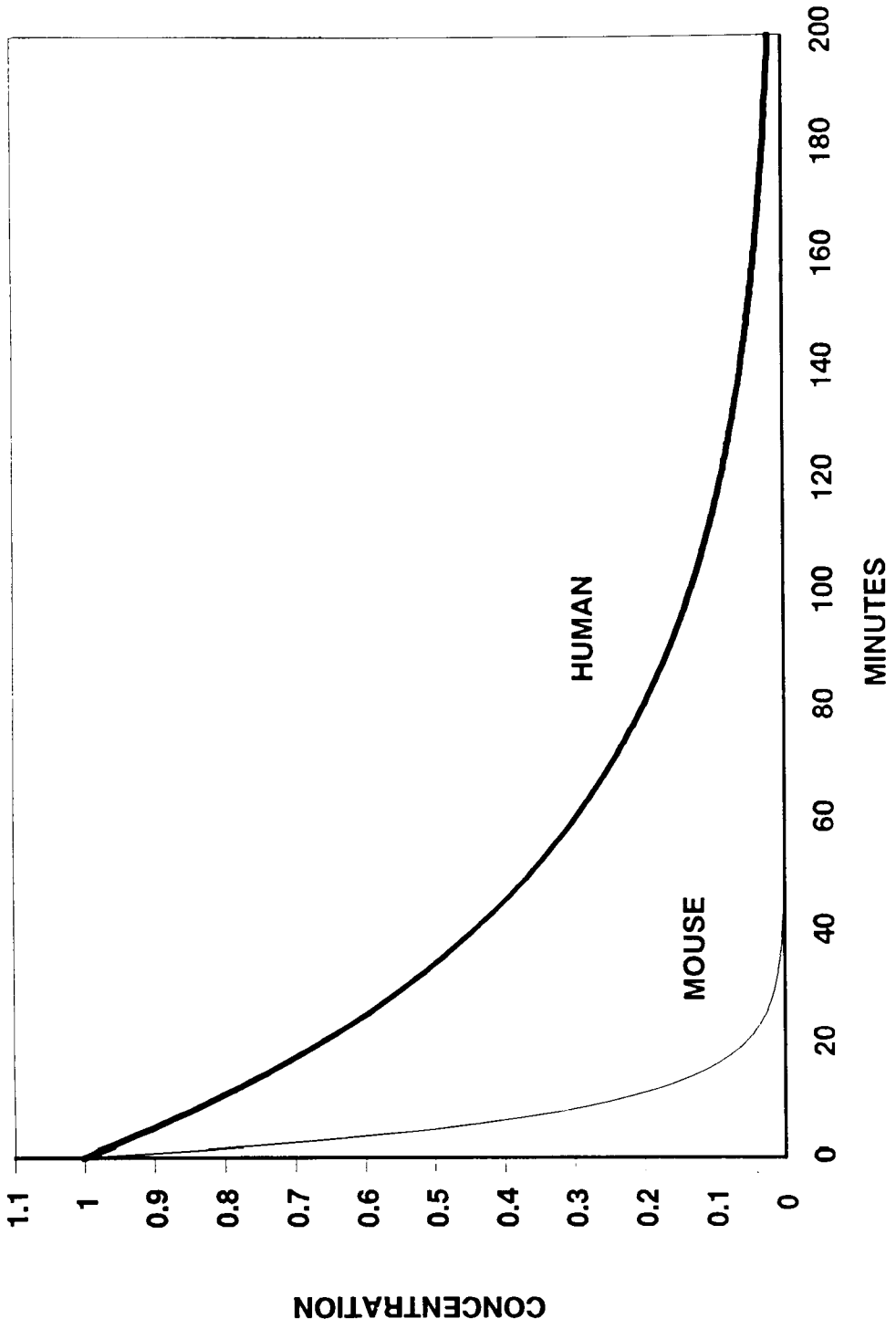

Figure 1. Blood concentration following injection of a dose scaled to body weight in a mouse (light line) and a human (heavy line). The human has an area under the curve that is 7-times greater.

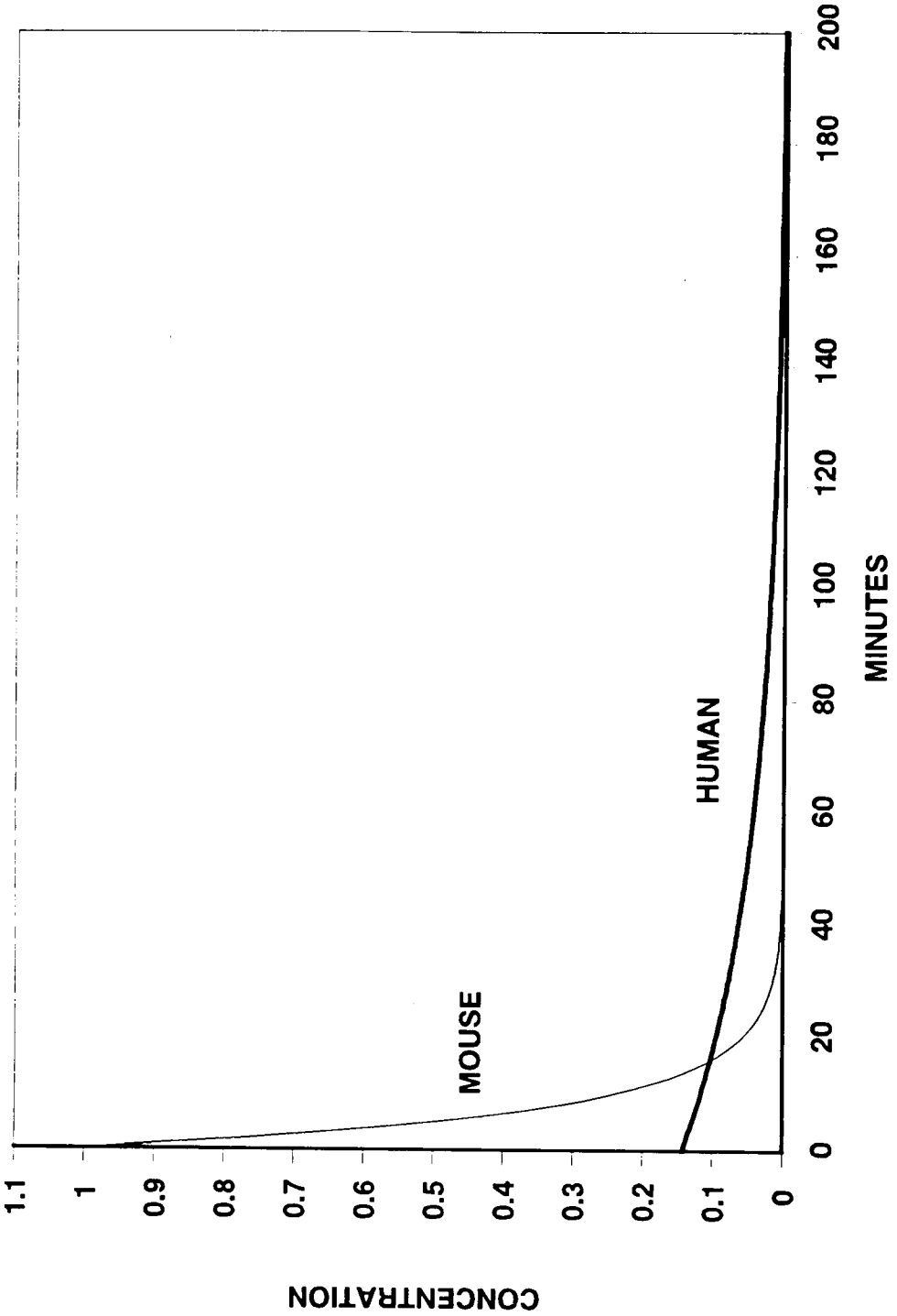

Figure 2. The injection amount scaled in proporation to $W^{3/4}$. The initial concentration in the mouse (light line) is 7-times higher than in the human (heavy line), but the AUCs are equal.

BILLING CODE 6560-50-C

This example illustrates two points: that knowledge of a compound's pharmacokinetics can suggest scaling of administered doses so as to equalize the exposure of internal targets of toxicity, and that "equal" internal exposure requires further definition. The area under the concentration curve encompasses both the amount of a compound that is present and the duration of its presence, providing a measure of the compound's opportunity to interact with the targets of toxicity. Moreover, since the AUC is the integral of concentration X time—that is, the "sum" of many momentary concentration levels—dividing the AUC by the time interval over which it is measured gives the average concentration during that interval. As such, the AUC is more representative of the target organ's total exposure to the agent than is the peak concentration. The AUC provides a measure of the agent's opportunity to participate in critical reactions at the target site. For example, for DNA-reactive compounds, the AUC is predictive of the rate of generation of DNA adducts (Hattis, 1990), while for moderate levels of receptor mediated carcinogens it tends to be proportional to average receptor occupancy. For such reasons, pharmacokinetic equivalence is usually defined in terms of equality of AUCs.

If this hypothetical chemical is assumed to be a carcinogen, an added difficulty in defining pharmacodynamic equivalence is also readily apparent. It should be remembered that equally carcinogenic doses are defined in terms of exposures repeated *every day* over a full lifetime. An adjusted daily dose that yields pharmacokinetic equivalence for one day's exposure of the target organ (as illustrated in Figure 2) is repeated for 2 years in the lifetime of a mouse, but 70 years in a human's. Furthermore, if the agent's stress on the physiological system at any given moment is not proportional to its concentration, the fact that the pharmacokinetically "equivalent" equal AUCs are achieved from different time-patterns of target organ exposure (as seen in Figure 2) could affect the carcinogenic consequences. These and other issues will be discussed at greater length further on in this document; they are raised here to emphasize that pharmacokinetic equivalence need not lead to carcinogenic equivalence without first employing further scaling considerations.

Clearly, actual pharmacokinetic and pharmacodynamic processes will be more complex than the simple considerations mentioned above would indicate. Nevertheless, there are some well recognized general trends in species differences (e.g., the higher metabolic rate in small mammals, the longer tumor latency in humans via-a-vis experimental rodents) that clearly influence the appropriate scaling of doses of carcinogens, and for which we should attempt to account in our scaling rationale (Boxenbaum, 1982, 1983; Schmidt-Nielsen, 1970, 1975, 1984; Travis et al., 1990; Ings, 1990). An analysis of the effects of major general trends in cross-species physiological differences not only helps guide our choice of appropriate scaling factors, but it provides the benchmark against which increasingly available case-specific data on the complex details of pharmacokinetics and carcinogenesis may be compared. Without such a framework, the impact of data on a single component—metabolic activation of a carcinogen in a target tissue in mice and humans, for example—is difficult to guage (U.S. EPA, 1987a,b). The analysis presented below is not a definitive solution to the cross-species scaling problem. Rather, it is presented as an attempt to accommodate present knowledge about the major quantitative trends in comparative anatomy and physiology into a scaling rationale with explicity stated assumptions.

The scaling of the myriad physiological processes that underlie the processing of carcinogens and their toxic effects can be drawn together into a single scheme by referring to the concept of *physiological time*. This concept proposes that quantitative differences across mammalian species in physiological processes can be seen largely as the consequence of fundamentally similar anatomical and biochemical machinery operating at different *rates* in differently sized species, smaller species having faster physiological "clocks." By correcting for these differences in size and time one can express dose-response problems in terms of a single scale-free mammalian system in which scaled doses should yield equal responses. (It is this very similarity, after all, that leads us to use experimental animals as surrogates for humans in risk assessment.) In the sections that follow, the issues of pharmacokinetic and pharmacodynamic equivalence are considered in turn.

1. Species Differences in Pharmacokinetics

The physiological time concept emerges from the study of the allometry of key physiological and anatomical variables that affect pharmacokinetics. Allometry studies the variation in features (and the consequences of that variation) as a function of body size and some other parameters. Most quantitative features that vary among mammals are well described by the so-called allometric equation,

$$Y = a\,W^{b},$$

where b is the power of body weight (W) to which attribute Y maintains a constant proportionality, a. A review of the large literature on this subject is beyond the scope of the present paper. The reader is referred to a number of excellent reviews (Adolph, 1949; Kleiber, 1932, 1961; Lindstedt and Calder, 1976, 1981; Schmidt-Nielsen, 1970, 1975, 1984).

The key point for the present argument is that there is great regularity in the value of b for certain classes of attributes relevant to pharmacokinetics (Travis et al., 1990). Volumes and capacities (blood volume, volumes of distribution, organ sizes, lung capacity, etc.) tend to remain in approximately constant proportion to body size (i.e., $b \approx 1.0$) in large and small mammals.

Rates, in contrast, tend to maintain proportionality with body weight to the 3/4 power (i.e., $b \approx 0.75$). Such rates include cardiac output, minute volume, basal metabolic rate and oxygen consumption, glomerular filtration rate, and many others. Consumption rates also tend to scale this way, including daily intakes of food, air, and water. A rate that scales in this way becomes smaller per unit weight (or volume) in larger animals. For example, a human has a total cardiac output (mL/min) about 300 times greater than a mouse, but in proportion to the human's 2000-times more massive body, the rate of blood delivery per gram of tissue is approximately seven-fold smaller (in terms of mL/min/g).

Several authors have suggested that this consistent scaling of rates of physiological processes leads to a useful concept of *physiological time* (Dedrick et al., 1970; Dedrick, 1973; Boxenbaum, 1982, 1983, 1984, 1986; Lindstedt and Calder, 1981; Mordenti, 1986; Lindstedt, 1987; Travis et al., 1990). A mouse is carrying out the same set of physiological processes as a human, but each process proceeds at a rate some 7-times faster. The various processes stay in proportion to one another, but all of them are relatively sped up in smaller species. If one scales the units of *time* by dividing them by the fourth root of body mass (i.e., $\text{min} \cdot W^{-1/4}$, correcting the physiological time scale) then the time-course of physiological processes becomes congruent across species. If time were measured according to some internal, physiological standard (such as

heartbeats, breaths, blood circuit times, clearance half-lives, etc.), rather than in minutes, then the rates of pharmacokinetic processes, the time course of disposition of a dose, and even life milestones and lifespan would all be about equal across species. (As discussed more fully below, humans tend to be an outlier in the relationship of lifespan to $W^{1/4}$, living longer than expected. Some authors have addressed this by including brain weight as a second factor in the allometric equation [Boxenbaum, 1986].)

This concept is illustrated by the simple example introduced in the previous section (shown graphically in Figure 1)—a single intravenous dose of a compound to a mouse and a human, and its subsequent blood concentration as it is removed from a single body compartment. (The simplicity is for illustration; the argument can be shown to hold for more complex pharmacokinetic models as well, e.g., Travis et al., 1990.) If doses are scaled to body weight (mg/kg) then initial concentrations are equal, but the blood level takes much longer to decline in the human, owing to slower processing of the compound. The human has a bood volume (which is proportional to body weight) some 2000-fold higher than the mouse, but the compound must be cleared from this volume by processes (metabolism and/or excretion) that operate only 300-fold faster (or seven-fold slower per unit blood volume). As a result, the human has an area under the blood concentration curve (or AUC) that is 7-fold higher. The AUC has units of [conc.]•[time], e.g., (mg/L)•min.

There are two kinds of scaling one could imagine to accommodate the species difference in pharmacokinetic behavior. The first has already been illustrated in Figure 2; one could give a smaller initial dose to the human—one that is seven-fold smaller in terms of mg/kg *but equal in terms of mg/kg*$^{3/4}$. The initial concentration is lower, but this is balanced by the slower removal to give the same AUC as seen in the mouse.

Alternatively, one could give the same initial mg/kg dose, but scale the *time* axis, expressing time in "physiological time units" (i.e., minutes divided by $W^{1/4}$). This is illustrated in Figure 3. Such graphs are sometimes called "Dedrick plots," following the demonstration of Dedrick et al. (1970) that scaling time in this way leads to congruity of methotrexate pharmacokinetics among several species. The mouse and human curves are identical on such a graph, falling to the same concentration after the same amount of *physiological* time has elapsed. (Of course, it still takes 7-times more minutes in a human for a given interval of physiological time to elapse. The AUC in the usual chronological time units is still bigger in the human, but in units of [conc.]•[physiological time] it is equal.)

BILLING CODE 6560-50-M

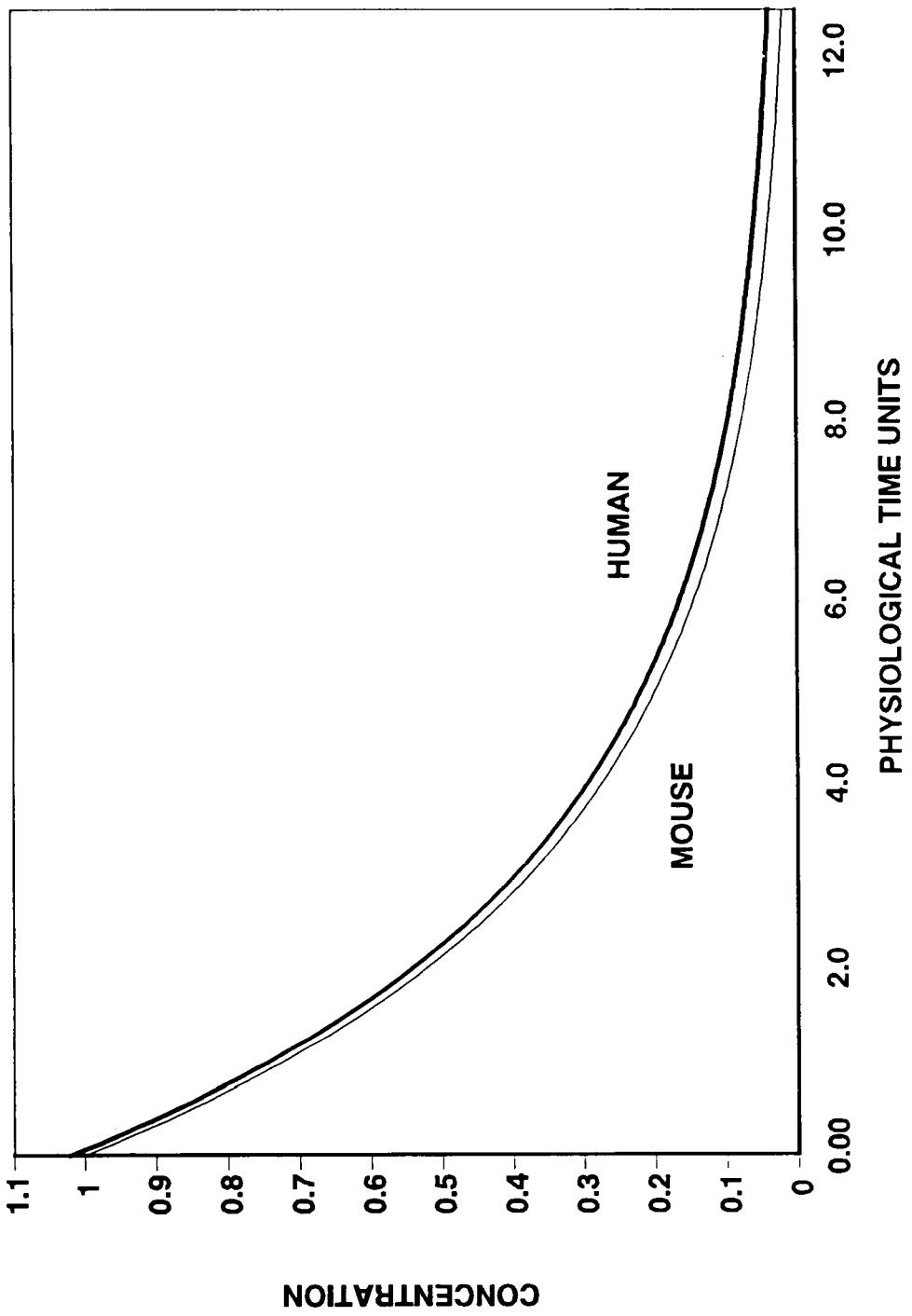

Figure 3. The human and mouse curves are superimposed when the time axis is expressed in units of physiological time, i.e., min. $\cdot W^{-1/4}$.

BILLING CODE 6560-50-C

It can be shown that these two scaling approaches—shrinking doses or stretching the time scale—give equivalent ways of dealing with scale differences as long as saturable pharmacokinetic processes do not figure prominently (O'Flaherty, 1989). For example, consider the slightly more complex case of repeated dosing. Figures 4 and 5 show blood concentration versus time curves for bolus dosing repeated at regular intervals. If dosing is daily (i.e., inter-dose intervals are equal for animal and human in clock time, as in Fig. 4) then scaling the bolus amount by $W^{3/4}$ achieves an equal area under the curve after a given number of days, as well as an equal average steady-state blood concentration. Alternatively (Fig. 5), one can give equal mg/kg doses spaced according to equal intervals of physiological time (e.g., daily in the mouse and every seven days in the human) to achieve the same end.

BILLING CODE 6560-50-M

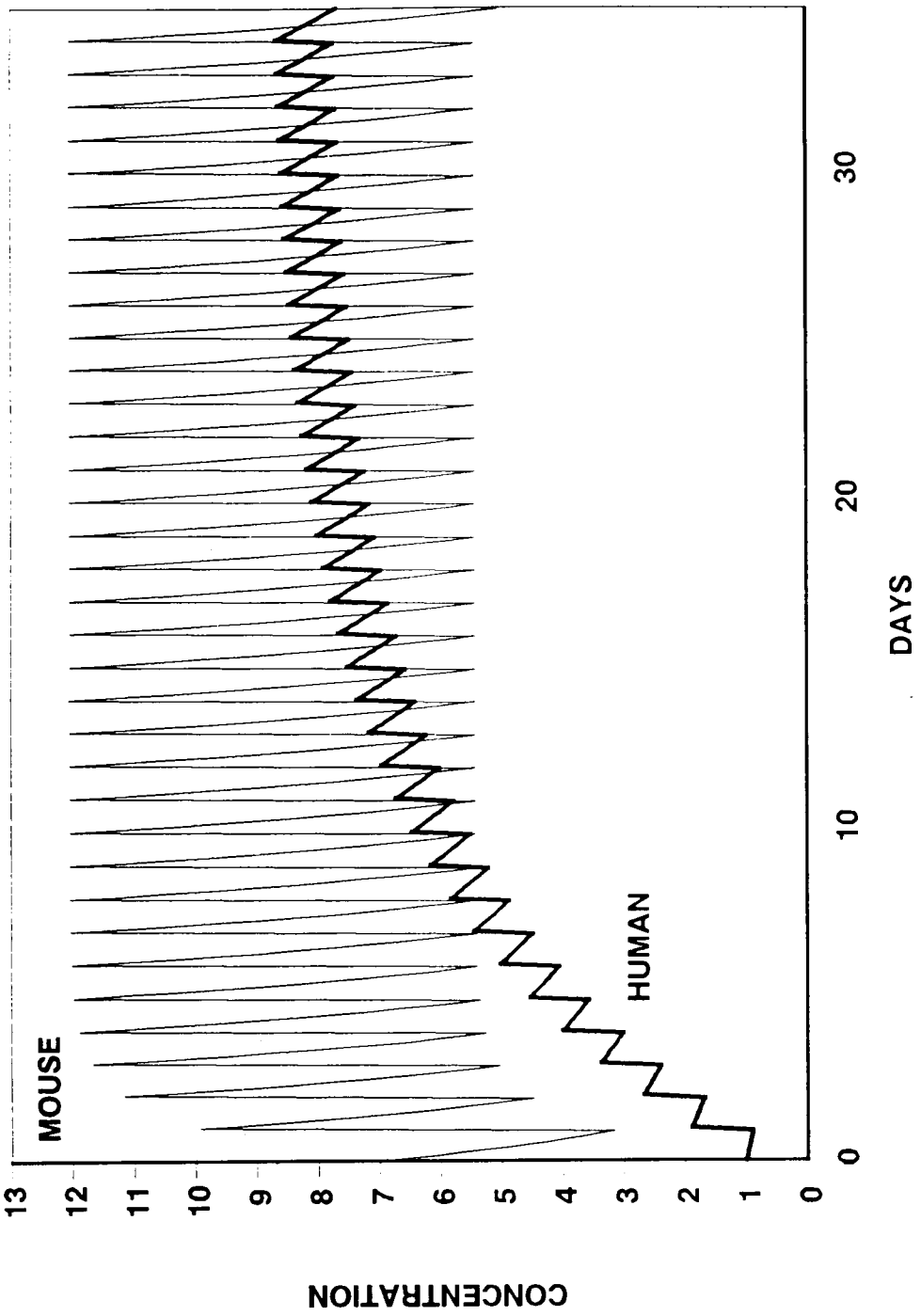

Figure 4. Repeated daily doses scaled to $W^{3/4}$ in the mouse (light line) and the human (heavy line).

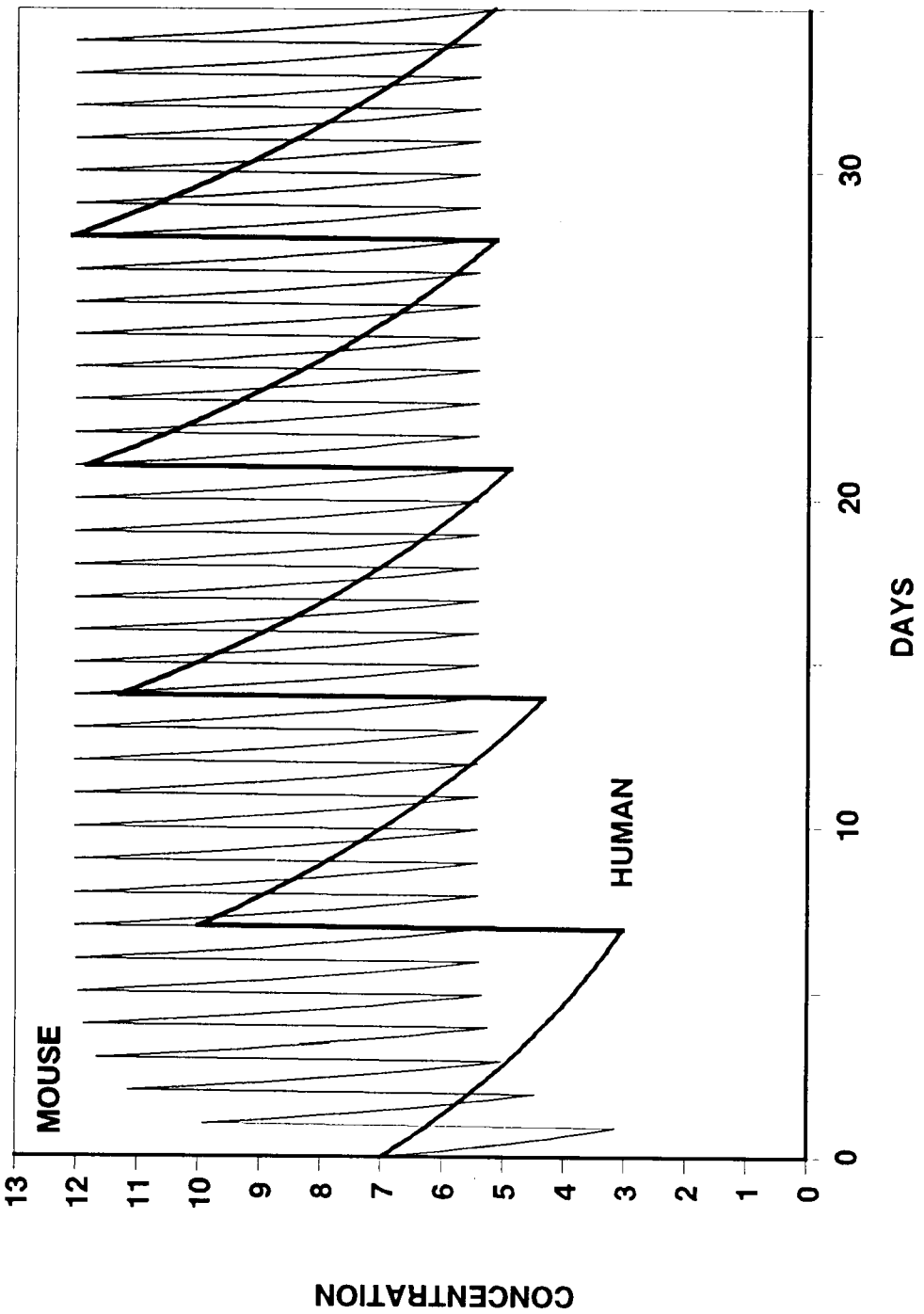

Figure 5. Repeated dosing of amounts scaled to body weight, repeated daily in the mouse (light line) and every 7 days in the human (heavy line).

BILLING CODE 6560-50-C

The foregoing examples are of course simplified and hypothetical, designed to illustrate the principles of allometric variation in physiological rates and volumes and their impact on the relation of administered dose to the degree of "internal" exposure. The same principles, however, can be shown to apply to much more complex pharmacokinetic systems as well, including multicompartment models, multiple routes of uptake and elimination, and multiple metabolic pathways causing carcinogenic activation and/or detoxification. The arguments have been most extensively developed by Mordenti (1986), O'Flaherty (1989), and Travis et al. (1990). The complete elaboration of the allometry of pharmacokinetics is too complex to detail here, but a few important points should be made.

First, the ability to predict the pharmacokinetic consequences of variation in the dozens of parameters that affect a chemical's uptake, distribution, processing, and elimination rests on the *regularity* in their cross-species variation and the *congruence* of these patterns for certain classes of parameters (rates, volumes, etc.). If physiological features varied haphazardly across species, or if all features had independent allometric patterns unrelated to one another, then no dose scaling method could be defined ($W^{3/4}$ or any other) to approximate pharmacokinetic equivalence without first knowing the compound's pharmacokinetics in detail.

Owing to their importance, it is well briefly to examine the starting assumptions that form the basis of the allometric, "physiological time" concept and its predictions. They are: (a) Volumes and capacities (organ sizes, blood volumes) retain proportionality to W; (b) the absolute rates of physiological processes are proportional to $W^{3/4}$; these rates include cardiac output, minute volume, glomerular filtration, and the rates of specific metabolic steps; (c) physicochemical and thermodynamic properties of compounds (solubilities in various tissues) are equal in all species; and (d) for metabolic pathways with saturable metabolism, the Michaelis constant (the substrate concentration at which half the maximum reaction velocity is achieved) is invariant, while the maximum velocity scales as $W^{3/4}$. A corollary to points (a) and (b) is that when rates are figured relative to body size (or to a volume, or in terms of concentration rather than absolute amount), they scale as $W^{3/4}/W = W^{-1/4}$, as illustrated by the cardiac output example shown earlier.

Most of the above assumptions are well supported by data on comparative anatomy and physiology, as detailed in the allometry references cited previously. Collectively, they embody the concept of a basically similar mammalian physiological and anatomical plan that varies primarily in scale from one species to another. The most problematic issue is the scaling of rates of individual metabolic transformation reactions as $W^{3/4}$. Not only are there few data on such scaling, but some individual metabolic enzyme activities are shown to vary rather haphazardly across species (e.g., Gillette, 1987; Calabrese 1986a,b). Several points should be made, however. First, there are data that support the proposition of $W^{3/4}$ scaling in specific cases (e.g., Reitz et al., 1988). Second, overall metabolic rate (O_2 consumption, resting metabolic rate) clearly scales as $W^{3/4}$; indeed, this is the issue around which physiological allometry was developed. Scaling an individual metabolic step in this way corresponds to keeping it in proportion to general metabolism, which seems the best default. Third, daily intake of natural toxins (the usual targets of carcinogen-metabolizing enzymes) depends on intake of air, water, and food (which all scale as $W^{3/4}$). That is, scaling detoxification processes in proportion to their anticipated load also predicts $W^{3/4}$ scaling.

Consideration of these points leads to the view that $W^{3/4}$ scaling of the rates of individual metabolic transformation reactions can be viewed as a benchmark around which different species (and individuals within a species) vary from instance to instance. Such variation does not invalidate the general scaling argument, nor does it provide evidence for any different scaling factor. Rather, the variation simply illustrates that any single conception of cross-species scaling can accommodate only the general trends, not the diversity of particular instances. Clearly, when data on metabolic conversion are available in a particular case, they should be used in preference to the $W^{3/4}$ default. In fact, instances of chemical-, dose-, and species-specific variation in metabolic transformation of a chemical may constitute the principal reason for deviation from the allotmetric default assumptions herein laid out. Accordingly, empirical determination of such metabolic variation constitutes the most important pharmacokinetic data that can be brought to bear on the estimation of target tissue exposures.

A second major point to bear in mind about the allometric analysis of pharmacokinetics is that the cross-species consequences of variation in the many physiological parameters depend not on the individual parameters, but on their interrelation. It is misleading simply to examine the scaling of one component (say, metabolic activation) in isolation. One must remember that the many quantitative differences across species are having their influences simultaneously; it is their *interactions* and *net results* that determine the consequences for doses to the tissues. For example, metabolic rates alone are a less important determinant of the fraction of a dose that is metabolically activated than is the ratio of metabolic activation rates to rates of other competing processes (such as renal clearance) that remove a compound from the body.

The third major point is that, despite the variety and diversity of underlying pharmacokinetic processes that may obtain from one case to another, the allometric analysis of pharmacokinetics makes rather general and simple predictions about how administered doses should relate to target tissue exposures in experimental rodents and humans. These predictions are:

For a given dosing pattern in which amounts are scaled to body weight, the tissue exposures (as measured by areas under the concentration curve) tend to be bigger in larger species by the ratio of human to animal body weight to the 1/4 power (which amounts to almost seven-fold for mouse-to-human scaling and not quite four-fold for rat-to-human scaling). If the administered amounts are kept in proportion to $W^{3/4}$ (rather than to W) the doses tend to be "pharmacokinetically equivalent" in the sense of yielding similar areas under the curve of concentration over time. Since daily intakes of air, food, and water tend to be in proportion to $W^{3/4}$ across species, calling exposures to environmental media equivalent on a ppm basis (i.e., when they are equally contaminated) produces essentially the same expectation of pharmacokinetic equivalence as scaling by $W^{3/4}$ (Hattis, 1991).

In fact, all variables containing [time] in their units will scale in a way that leads to the human value being bigger by the ratio of body weights to the 1/4 power. If these variables are reexpressed in terms of "physiological time units," i.e., [time] $\bullet W^{-1/4}$, then their values are equal across species.

The above conclusions apply to parent compound and to metabolites, since (in this generalized scheme)

metabolites are also subject to scale-affected clearance processes. In humans a metabolite may be formed more slowly, but the amount that is formed persists longer, resulting in similar AUCs as seen in rodents. The pharmacokinetic equivalence applies not only to an agent's concentration in blood, but also to concentrations in any specified organ or tissue. Thus, the scaling applies to the AUC of the ultimate carcinogenic species (be it parent compound or metabolite) at the particular site in the body that constitutes the target of carcinogenesis (presuming the target site to be the same across species).

The proportion of the administered dose that ends up having any particular ultimate fate (e.g., being excreted unchanged, being metabolized by a particular biochemical pathway at a particular site, being excreted as a conjugate in the urine, etc.) is predicted to be the same independent of species. That is, if a mouse given 10 mg/kg of an agency ends up metabolizing 4 mg/kg into a form that has an AUC in the spleen of 100 (mg/L)•min, then the allometric prediction for a human given 10 mg/kg is that 4 mg/kg will be metabolized, but the AUC in the spleen will be 700 (mg/L)•min, owing to the metabolite's slower clearance.

A difficult situation arises when the active carcinogen is neither the parent compound nor a stable metabolite, but rather a very reactive metabolite, perhaps an intermediate formed ephemerally during the course of metabolic transformation. If this reactive compound is removed by spontaneous reaction (rather than further enzymatic processing) and if such spontaneous reaction is so rapid that the moiety never leaves the tissue in which it is formed, then the removal rate may no longer be species-dependent; instead, it may hinge only on physicochemical properties of the reactant and its milieu. In such a case, without species differences in persistence, the AUC of the reactive moiety in its tissue of formation may be proportional to the amount formed. Such AUCs would tend to be equalized when doses are scaled to body weight, rather than to $W^{3/4}$ (Travis, 1990).

It may be well to reiterate at this point that the reason for constructing these general allometric arguments is to predict the AUC of the proximate carcinogenic agency at its site of action in those cases (which constitute the majority of cases at present) for which no better means exists to determine relative target tissue doses in rodents and humans. Clearly, if better means exist to characterize target tissue exposures, they should take precedence. Pharmacokinetic modeling of a particular compound may demonstrate that the allometric presumptions are in error. Two possible causes of such error are: (a) species differences in metabolic processing that do not adhere to the rule of proportionality to $W^{3/4}$, and (b) saturation of metabolism in one but not the other species as a result of comparing markedly different dose levels or dosing regimens. The importance of the "reactive metabolite" scenario outlined in the previous paragraph is best determined by case-specific characterization of metabolic activation and its effects. Macromolecular adducts may be particularly useful in this regard since, under certain circumstances (including negligible repair), their accumulation in a tissue would be expected to be proportional to the AUC of the adduct-forming moiety in that tissue.

It must be conceded that, in actuality, mice and rats are not simply scale-model humans; certain particular characteristcs (metabolism among them) do not necessarily vary in a simple way with body size. However, the long-standing toxicological practice of using rodent exposures to toxic agents as surrogates for the human experience rests on the belief that, to a first approximation, the similarities that stem from a shared mammalian anatomy and physiology outweigh the differences. The species differences in size, uptake rates, basal metabolism, blood flows, organ sizes, and so on are clearly important to acknowledge in any dosimetric scheme. The allometric arguments adduced here attempt to construct a logical and consistent framework for investigating cross-species dosimetry. This framework provides a basis for articulating the expected consequence of those broad general patterns of cross-species difference in size scale and time scale that we understand, while providing rebuttable default positions for those aspects, such as chemical-specific metabolism, that are less well understood.

2. Species Differences in Pharmacodynamics

The overall aim of dose scaling is to achieve toxicological equivalence across species. The foregoing section discussed pharmacokinetic equivalence. For such results to be useful for carcinogen risk assessment—that is, to complete the equation of exposure and tumorigenic response—it remains to determine what toxicological consequences to expect from given target tissue exposures in humans and animals. As argued earlier, the principles of pharmacodynamic equivalence are far from self-evident.

The issues about pharmacodynamic equivalence fall into three categories. First, the appropriate measures of "delivered dose" would seem to depend on details of the mechanism of toxic action, details that are frequently poorly understood. In the foregoing section, scaling of administered doses was discussed in terms of tendency to equalize the AUC, an integrated measure of target tissue concentration. Although this is a frequent and widely accepted measure of a target organ's exposure to a toxin (Voisin, et al., 1990), its use as a measure of carcinogenic equivalence of doses rests on the presumed proportionality of the rates of toxicological reactions to the AUC. If the underlying reactions that comprise the process of carcinogenicity are markedly nonlinear with target-tissue concentration, if they include capacity-limited steps or magnitudes below which significant stress on the system is absent, then proportionality of toxic response to the AUC (or to any other easily characterized summary measure of target-tissue exposure) becomes problematic. Thus, use of the AUC as an "equivalent" tissue dose should be regarded as a default that corresponds to the presumption that the processes constituting carcinogenicity operate in proportion to the concentration of the carcinogen at the target. In particular applications, this assumption should be critically examined, and relevant data brought to bear, if possible.

The second issue returns to the question of scale. For corresponding organs bathed in an equal concentration of carcinogen, a human will have many more target cells exposed than a rodent, only one of which need be transformed to found a tumorigenic clone. Moreover, during the course of a full lifetime under this dosing regime, a human's cells will be exposed for much longer and undergo many more cell divisions (NAS, 1975; U.S. EPA, 1987a). Although this would seem to suggest a much larger sensitivity to carcinogens in larger species, the empirical evidence shows instead a rough lifetime-to-lifetime equivalence across species of both the magnitude of spontaneous cancer risk and the age pattern of its appearance. When arguments from first principles lead to answers that are clearly off track, it indicates that key factors have not been brought into consideration. In this case, the role of species differences in repair processes may enter. Also, the number of cells (or cell divisions) at risk may be less different among species

than presumed, owing to slower turnover, stem cell populations that are not proportional to tissue volume, or other factors. The point is raised here simply to emphasize that size and timespan differences across species may have key roles in comparative pharmacodynamics just as they do in comparative pharmacokinetics, although the particulars are not clear at present. In the face of this difficulty, it has been the ususal practice to assume lifetime equivalence when projecting carcinogenesis patterns across species, an assumption that has held up well in experience. This point will be returned to below.

The third issue in pharmacodynamic equivalence also parallels one in pharmacokinetics—that of the uniqueness and species-specificity of carcinogenic responses that tends to obscure overall trends and patterns. The pharmacodynamic reasons for differences in sensitivity of potential target organs among species are perhaps more obscure than the pharmacokinetic reasons, but they surely exist. As with the case-by-case particulars of pharmacokinetic processes, the idiosyncratic and species-specific variations in responsiveness to carcinogenic stimuli create an unavoidable envelope of uncertainty around the predictions of a scaling methodology that can only characterize the average behavior of carcinogens overall. When data are available that enable the investigator to incorporate knowledge of species differences in the carcinogenic reactions to a given level of target-tissue dose, they should be considered in the analysis and incorporated when appropriate.

Although certain pieces of the puzzle of cellular and molecular biology that underlie carcinogenesis are known, and despite rapid progress, it not yet possible to undertake a detailed analysis of the magnitudes and causes of species differences in the carcinogenic process. At present, there can be no empirical and allometric characterizations of general cross-species trends, as has been done in this report for the pharmacokinetic part of the equation. One can, however, make use of the observation of general lifetime-equivalence, noted above, to suggest how the insights of cross-species patterns in pharmacokinetics might be applied to the question of toxicological equivalence.

3. Toxicological Equivalence

When experimental animals and humans are exposed to a chemical in such a way that they experience equal areas-under-the-curve of the proximate carcinogenic agent (be it the parent compound, a metabolite, or a reactive intermediate of metabolism) at the target of toxic action, then they will have their susceptible tissues exposed to equal *average* concentrations of the carcinogen over the exposure period. Over the course of a full lifetime of exposure, the lifetime average target-tissue concentrations are equal (although the total accumulated AUC is larger in humans, by virtue of their longer lives). The earlier discussion of pharmacokinetics argued that, if daily administered doses are scaled in proportion to $W^{3/4}$ (or if exposures of equal duration are equated on a ppm basis), such equality of resulting AUCs tends to result across mammalian species.

If the empirical principle of lifetime-to-lifetime equivalence is applied, then a possible presumption is that such pharmacokinetically equivalent lifetime exposures (in terms of equal average concentrations of the carcinogen at its target) should be equivalent in the degree of lifetime cancer risk they engender (although other interpretations of the consequences of pharmacokinetic equivalence are possible). That is, it may be assumed that equal carcinogen concentrations at the target lead to equal degrees of impact at the cellular level which, if continued for a lifetime, yield equal lifetime probabilities that a tumor will be caused in that target organ.

The reasons for approximate lifetime equivalence in the carcinogenic process among species of different body size and lifespan are not clear. One can, however, rationalize this observation by extending the concept of physiological time from pharmacokinetic processes to cover pharmacodynamic processes as well. The following section explores this approach.

4. A Physiological Time Approach to Toxicological Equivalence

It is helpful to begin by considering the case of "zero" dose, i.e., by examining background or spontaneous carcinogenesis. Although the common cancer types differ somewhat, humans and experimental animals have roughly similar lifetime cancer rates. Moreover, the latency periods are greatly different in animals and humans, but in a way that is roughly proportional to lifetime. Age-specific incidences are also roughly parallel when time is measured not in years, but on a lifetime scale (Cutler and Semsei, 1989). If these equivalencies were not so, we would either never see tumors in experimental animals (since they would die of other causes before the 20-to-40 year latency was completed), or we would find humans to be overwhelmed with spontaneously arising tumors during childhood. These results from spontaneous carcinogenesis appear to be paralleled by chemically induced cancers, in that such cancers also arise and progress on a "lifetime" time scale in experimental animals and humans.

The above results suggest that carcinogenesis proceeds more slowly in larger animals, in a way that makes its progress roughly constant per lifetime, rather than per unit of clock time. This is in accord with the current risk assessment practice of equating lifetime cancer incidences in humans and rodents. It would seem that the concept of physiological time—that large animals carry on their life processes at an overall slower pace than smaller ones—proves as useful in examining pharmacodynamics as it does for pharmacokinetics. As argued in the previous section, the rates of the underlying pharmacokinetic processes tend to operate in proportion to a size-dependent physiological time "clock," which allows appropriate scaling to explain and correct for species differences in pharmacokinetic end points." In the case of carcinogenesis, the component physiological features and processes are less easily observed, but the "pharmacodynamic end point" can be seen in the above-mentioned cross-species patterns of spontaneous carcinogenesis. In sum, not only may "pharmacokinetic time" vary among species in a regular way, "pharmacodynamic time" may do so as well. Total lifespans of different species generally scale in rough proportion to $W^{1/4}$ (Sacher, 1959; Lindstedt and Calder, 1976, 1981). (In terms of the physiological time concept, the "processes of living" that proceed at a rate proportional to $W^{3/4}$—or on a per kg basis, to $W^{-1/4}$—go slower in a larger animal, and so take chronological time in proportion to $W^{1/4}$ to go "to completion.") Hence, the two physiological time scales are quite similar. However, humans live longer than their allometric prediction by about a factor of five.

The above discussion of pharmacodynamics suggests that carcinogenesis (in common with other physiological processes) proceeds more slowly in humans than in rodents, in a way that tends to be equivalent on a lifetime basis. Together with the pharmacokinetic results outlined earlier—namely, that scaling daily administered doses in proportion to $W^{3/4}$ tends to result in "pharmacokinetically equivalent"

exposures to corresponding organs and equal steady-state concentrations of agents and their metabolites—this suggests that administered doses of carcinogens be considered equal in lifetime risk when expressed in units of mg/kg $^{3/4}$/day. One possible interpretation of this line of reasoning is that tissues experiencing equal average concentrations of the carcinogenic moiety over a full lifetime should be presumed to have equal lifetime cancer risk. Under the arguments on pharmacokinetic allometry set out earlier, such equality of average concentrations would tend to be produced by daily administered doses scaled in proportion to $W^{3/4}$. However, if the pharmacokinetically equivalent doses can be obtained by experimental means, under this line of reasoning, such results could replace the allometric presumptions, and equal risks would be expected when average daily AUCs are equal (or equivalently, when average concentrations are equal). If the default allometrically based assumptions about pharmacokinetics are adhered to by a particular compound, the introduction of data in place of assumptions will leave the answer unchanged. Other interpretations of the question of the cross-species toxicological equivalence of delivered doses are possible, and the issue remains one on which further insight would be helpful.

If we use a scale of pharmacodynamic time based on the equivalence of lifetimes, then the 35-times larger exposure of human tissues to carcinogens that results from a lifetime of doses scaled by mg/$W^{3/4}$/day results in an equal lifetime cancer risk because the affected physiological processes of carcinogenesis themselves are operating more slowly (by assumption, 35-times more slowly). A given span of clock time that a tissue spends under a given concentration regime yields less risk in a human (since the tissue has spent less "pharmacodynamic time" exposed).

It should be clear that not every empirical measure of "internal dose" is equally informative about species differences. As noted earlier, the amount of a dose metabolically activated, for example, may be equal in a mouse and a human, but the human's AUC of metabolite at the target may be much larger. If an empirical measurement or modeled result is to be used as a surrogate for "internal dose" in a cross-species extrapolation, its value in animals and humans should be compared to the predictions of the default assumptions of allometrically scaled pharmacokinetics (which should be aided by a full analysis of the uncertainties in the available data and of reasonably likely alternative pharmacokinetic modeling approaches). With this kind of analysis, it is possible to judge whether those default assumptions have actually been contradicted by data for the case at hand.

Once again it should be stressed that the arguments set out here are intended as defaults. They attempt to gauge the expected effect of known major cross-species trends in the rates and magnitudes of the underlying physiological processes, both in the internal disposition of a dose and its subsequent carcinogenic effect. Just as the pharmacokinetic presumptions may be able to be replaced with sufficiently validated case-specific modeling, the pharmacodynamic presumptions may be replaced with suitable biologically based dose-response models. The true pharmacodynamic situation is clearly more complex than represented here. In particular, there may be dose-rate effects, in which higher concentrations have more-than-proportionally stronger effect (Hattis, 1990). The effect of one moment's exposure may also depend on age or on the degree of exposure earlier in life. Such effects have no generalizable patterns, however, and cannot serve as a basis for default scaling of effects. Again, we seek a simple default principle to guide our expectations, while allowing for the use of case-specific experimental or epidemiologic insights (when available) to improve the estimate based on the simplifying assumptions.

It should also be pointed out that this scheme, with its explicit treatment of time, pharmacokinetics, and pharmacodynamics, provides a conceptual framework for examining such crucial emerging issues as risks from partial lifetime exposures, potencies in children vis-à-vis adults, and other similar questions. Failing to provide such an explicit argument from stated assumptions dooms a scaling factor to be inapplicable to such questions and provides no means for incorporating biological insights, such as data on pharmacokinetics and mechanism of action, when they are available.

III. Discussion

This proposal aims at arriving at a very broad generalization about carcinogen exposures that can be considered of equal risk in experimental animals and humans—one that can be applied to potentially carcinogenic chemicals lacking adequate information on pharmacokinetics and mechanisms of action. It attempts to provide a rational basis for a *prima facie* characterization of potential risks in humans, consistent with our empirical knowledge of carcinogen potencies in animals and humans and with the known general consequences of species variation in body size and the rates of physiological processes.

To achieve this wide applicability and generality, it is necessary to rely on simplified, broad patterns and trends of biological variation, while bypassing many details and causes of case-by-case variation. This is not to deny the importance of these details, nor to denigrate the value of case-specific data that show species- or dose-related differences in uptake, metabolism, or physiological actions of putative carcinogenic agents. To the contrary, the intention is to provide a framework for the use of such data, allowing (and indeed, encouraging) one to go beyond the *prima facie* case based on overall trends to address the impact of specific knowledge about the chemical and its actions.

The empirical data on carcinogen potencies estimated in various animal species and in humans demonstrate the large variability involved. Although scaling doses by $W^{3/4}$, as proposed herein, characterizes the trend fairly well, individual chemicals may deviate from this overall pattern by two orders of magnitude or more in either direction. In the case of the allometric arguments, there are dozens of points in the chain of inference where one could raise counterexamples to simplifying assumptions, arguing that the generalized $W^{3/4}$ scaling method thereby would over- or underestimate human risks for that case. For example, Gillette (1985) lists a number of physiological factors with high variability that would influence the accuracy of extrapolation of a dose's toxicity to an exposed human, not the least of which is the 20-to-50-fold variation among individual humans in their ability to take up and metabolize an agent and to repair any resulting damage.

The existence of such underlying variation means that the extrapolation of chemically induced risks observed in one circumstance (say, in a mouse lifetime cancer bioassay) to another (say, to people exposed to environmental pollutants) needs to be carefully and properly interpreted. Clearly, the projection of an equivalent dose is not merely a conversion of units, with the resulting human dose achieving an equal factual standing to the original animal observation. The projection is an

hypothesis, formulated in the face of uncertainty. In the most basic case—when there is little additional information that may be brought to bear—this hypothesis is framed in terms of the general features of anatomical and physiological differences among species that should affect all chemicals. It represents a best guess based on general principles and the recognition of overall trends. This best guess is surrounded by an envelope of considerable uncertainty, owing to the dozens of particulars that make each chemical's disposition and toxic effects in various species unique, despite the overall trends. When applicable pharmacokinetic and mechanistic insights into the particular chemical and its actions are available, they can (and should) be used to refine the projections by identifying and accounting for these chemical-specific factors.

Every projection of human equivalent dose, no matter how sophisticated, will have associated with it both uncertainty and variability. The uncertainty concerns whether the scaling method employed has correctly embodied and utilized the information at hand (be it general cross-species trends over all chemicals or case-specific insights from pharmacokinetics and mechanistic studies). The variability arises because even a sophisticated projection, when applied to a population of cases, will at best predict the mean of an array of actual values that reflect the myriad individual factors that no analysis can completely take into account. The "true" dose of equivalent risk will vary among exposed humans according to how each individual deviates from the overall human norm, owing to genetic factors, environmental influences, age, sex, lifestyle, and countless details of personal history.

The goal of a cross-species scaling methodology, then, is not to arrive at "true" values of equivalent does under all circumstances (for this is impossible, even in principle). Rather, it is to embody correctly and without bias the impact of the information at hand, providing rational estimates that take into account what is known, recognizing that true values will vary around this estimate as a result of case-by-case particulars, many of which are either unknown to vary among the individuals for whom the projections are being made.

The proposed scaling of daily administered doses of putative carcinogens by $W^{3/4}$ is intended to be such an unbiased projection; i.e., it is to be thought of as a "best" estimate rather than one with some conservatism built in to assure that any error is on the side of being overly protective. It should not be interpreted as a "safety factor" or other intentional bias designed to "err on the side of safety." Thus, it is to be expected that some individual compounds will have their human potencies overestimated by this procedure, while others will have them underestimated.

This having been said, it must be said, it must be acknowledged that there is considerable uncertainty about the best scaling method to achieve this unbiased projection. In particular, the empirical data on comparative carcinogen potencies are also compatible with both body weight and surface area scaling, the methodologies that we propose to abandon in favor of $W^{3/4}$ scaling. The $W^{3/4}$ scaling is chosen both to achieve unity of default methods and because it can be related to an explicit rationale based on allometric variation of the underlying anatomy and physiology. Former methodologies have not been shown to be false, however, and it is considered that risk assessments conducted under these methodologies are not in need of revision on account of any agreement to utilize a common methodology in the future.

The utility of the "physiological time" concept for understanding the patterns of cross-species differences in a carcinogen's action lies in its simplicity and generality. Because organ volumes tend to share a common pattern of allometric variation, while rates of physiological processes share another, the general predictions of cross-species differences is independent of specific hypotheses about target organs or mechanisms of action. One could, for instance, envisage an alternative allometric formulation that, rather than relying on overall patterns for unspecified organs in all mammals, focuses instead on the details of specific organs (common target organs or sites of metabolic transformation, say) in specific laboratory animal strains and in humans. For example, instead of relying on the approximation that breathing rates vary as $W^{3/4}$, one could make precise measurements of rates in B6C3F1 mice and in the humans whose risks are being evaluated. The utility of such an approach for a *default* scaling factor is doubtful, however, since the generality of the argument is lost, and the analysis becomes contingent on the details of the specific physiological hypothesis being elaborated. If such specificity is possible in an individual instance, it should become part of the case-specific pharmacokinetic and pharmacodynamic analysis that overrides the default methodology.

It is sometimes suggested that there should be more than one "default" scaling methodology, with different generalized procedures to be applied to different classes of chemical carcinogens. At present, it is not clear how such division of cases would be made, however, nor what the consequences on a generalized method should be. For example, tissue area-under-the-curve of the toxic moiety would seem to be the best *prima facie* dosimeter for the effects of both genotoxic and non-genotoxic carcinogens on their target organs. Similarly, the general allometric arguments for how AUCs are expected to vary across species apply both to agents active as the parent compound and to those requiring metabolic activation.

A possible exception to this pattern has been mentioned earlier. The generalized allometric pattern assumes that the rate of clearance of a metabolite from the target site of toxic action, like other rates, scales in proportion to $W^{3/4}$. If a compound acts through a very reactive metabolite that is spontaneously and fully deactivated by purely physical-chemical processes within the target tissue itself, then the rate of detoxification may be species-independent, and the AUC may be more related to the amount metabolized, which by default is expected to retain proportionality to body mass (Travis, 1990). Such a situation is not only plausible, it may be frequent. There is no particular indication from the empirical data, however, that different rules apply to metabolically activated compounds. Moreover, since the reactive intermediate scenario breaks the symmetry of the physiological time argument, it is difficult to know exactly what the carcinogenic consequences should be. This remains an important problematical area that requires future attention. For the present, however, there do not seem to be grounds for specifying when and how one should alter the default proposal.

The analysis presented herein is oriented around scaling doses so as to yield equal areas under the carcinogen's concentration curve at the target site. This definition of equivalence of target "doses" is in line with common practice. The AUC provides a measure of the agent's opportunity to interact with the target. Equal AUCs over a fixed time interval correspond to equal average concentrations of the agent during that interval. It should be borne in mind, however, that other measures of target

tissue dose might be more appropriate for specific mechanisms of carcinogenicity. For example, if a critical concentration must be reached or if there is a nonlinear dependence of toxic stress on concentration of the agent. Such alternative have no generalizable consequences or patterns, however, and there is no evident way to bring them into a default methodology. When case-specific pharmacokinetic analysis is undertaken, careful attention should also be paid to the measure of target tissue dose that is being considered to yield equivalent lifetime carcinogenic effect, and alternatives should be examined.

When AUCs from daily exposures are equal, then average concentrations of the agent at the target sites are equal. And when dosing producing equal daily average concentrations is continued for a lifetime, then average lifetime concentrations are equal. If one presumes that such average lifetime concentrations yield equal cancer risk, then the argument follows common practice and is in accord with the general finding that age-specific tumor incidence patterns tend to be congruent across species when expressed on a lifetime scale. (Other presumptions about the impact of such equal concentrations can be held, however.) The underlying biological basis for lifetime equivalence, and the conditions under which it might be violated, are not clear at present. This is an area in need of further investigation, and increased understanding will be key to determining how to scale the results of cell-kinetically based models of carcinogenesis from animal models to humans.

It should be borne in mind that the arguments for scaling doses by $W^{3/4}$ have been cast in very general terms to reflect constant, low-level, lifetime dosing and consequent lifetime cancer risks. Care should be taken when applying the methodology to specific exposure scenarios that deviate from this pattern. For example, the allometric arguments are adduced for variation among mammals. Other groups of animals have their own characteristic allometric patterns, but they are different than the mammalian ones. To extrapolate across classes of vertebrates with the proposed methodology, for example, would violate the basic presumption of the variation in a basically similar anatomical and physiological plan among differently sized mammals.

The allometric patterns relied on by the present argument represent variation *among species* for adult organisms. Allometric patterns among variously sized *individuals* of the same species can (and generally do) differ from the pattern seen from one species to another. The metabolic and lifespan patterns across species do not really describe variation among differently sized humans, for example. In other words, the scaling arguments presented here do not necessarily apply for the adjustment of doses to larger and smaller humans. In such cases, it is probably preferable to use mg/kg scaling (although the difference between this and $W^{3/4}$ scaling is minor). Similarly, the allometric patterns describing the changes within an individual as he or she grows and matures from child to adult generally differ from both the cross-species pattern and from the variation among differently sized adults. Compared to adults, children do have faster metabolic rates and greater intakes of food, water and air per unit of body weight, but these relations are not well described by proportionality $W^{3/4}$, as they are across species. Moreover, children also have proportionally faster rates of cell division (i.e., both pharmacokinetic and pharmacodynamic time are accelerated compared to adults). This a complex and problematic issue that is beyond the scope of the present document. It is deserving of further study. At present, it seems most reasonable to follow current practice, i.e., to scale doses for adults and children (and for differently sized adults) on a mg/kg basis. For similar reasons, the present scaling arguments provide no special insight into the problem of partial lifetime exposures.

Finally, it should be borne in mind that the scaling arguments are made for similar levels and patterns of exposure in animals and humans. When experimental animals are exposed to much higher levels than humans (as is common in carcinogenicity bioassays) there is the possibility of saturation of metabolism in animals that is not shared with human exposures. Such effects will obscure the usual pattern of equivalence of internal doses projected on the assumption of similar exposure regimes. In other words, dose scaling cannot solve the high-to-low-dose extrapolation problem, which must be addressed by other means. Case-specific pharmacokinetic analysis can, however, provide very valuable insight into differences in target tissue doses between rodents at high bioassay exposures and humans at much lower exposures.

IV. Conclusions

This notice is an announcement of a consensus reached by the Environmental Protection Agency, the Food and Drug Administration, and the Consumer Product Safety Commission to consider that lifetime cancer risks will be presumed to be equal when daily amounts administered are in proportion to body weight raised to the 3/4 power. It should be reiterated that former methodologies have not been shown to be in error, and this agreement should not be construed as overturning those practices with one of superior scientific validity.

The empirical data on comparative carcinogenic potencies in different species support the general practice of scaling rodent potencies to humans, and show that, on average, current methods perform rather well. The data are not of sufficient resolution, however, to distinguish between surface area and body weight dose scaling. The data are fully consistent with the proposal contained herein for scaling by body weight to the 3/4 power.

Theoretical support for scaling carcinogen doses by the 3/4 power of body weight is available from analysis of the allometric variation of key physiological parameters across mammalian species. Such an analysis has the benefit of providing an articulated rationale for the scaling methodology and of setting out the underlying assumptions explicitly.

V. References

Adolph, E.F. 1949. Quantitative relations in the physiological constitution of mammals. *Science* 109:579–85.

Allen, B.C., A.M. Shipp, K.S. Crump, B. Kilian, M.L. Hogg, J. Tudor, and B. Keller. 1987. *Investigation of cancer risk assessment methods*. (3 volumes plus summary report). Prepared for U.S. Environmental Protection Agency under contract to Research Triangle Institute, U.S. EPA Contract N° 68–01–6807. National Technical Information Service N° PB88–127113.

Allen, B.C., K.S. Crump, and A.M. Shipp. 1988. Correlation between carcinogenic potency of chemicals in animals and humans. *Risk Anal.* 8:531–61.

Boxenbaum, H. 1982. Interspecies scaling, allometry, physiological time, and the ground plan of pharmacokinetics. *J. Pharmacol. Biopharm.* 10:201–27.

Boxenbaum, H. 1983. Evolutionary biology, animal behavior, fourth-dimensional space, and the raison d'être of drug metabolism and pharmacokinetics. *Drug Metab. Rev.* 14:1057–97.

Boxenbaum, H. 1984. Interspecies pharmacokinetic scaling and the evolutionary-comparative paradigm. *Drug. Metab. Rev.* 15:1071–121.

Boxenbaum, H. 1986. Time concepts in physics, biology, and pharmacokinetics. *J. Pharmacol. Sci.* 75:1053–62.

Calabrese, E.J. 1983. *Principles of Animal Extrapolation.* John Wiley & Sons. New York.

Calabrese, E.J. 1987. Extrapolation from animal data. *In:* Tardiff, R.G., and J.V. Rodricks (eds.). 1987. *Toxic Substances and Human Risk: Principles of Data Interpretation.* Plenum Press. New York.

Calabrese, E.J. 1986a. Animal extrapolation and the challenge of human heterogeneity. *J. Pharmacol. Sci.* 75:1041–6.

Calabrese, E.J. 1986b. Comparative biology of test species. *In:* Hill, T.A., R.C. Wands, and R.W. Leukroth, jr. (eds.). 1986. *Biological Bases for Interspecies Extrapolation of Carcinogenicity Data.* Prepared for Food Safety and Applied Nutrition, Food and Drug Administration, Department of Health and Human Services, Washington, D.C. under contract to Federation of American Societies for Experimental Biology, FDA contract N° 223–83–2020.

Chen, J.J., and D.W. Gaylor. 1987. Carcinogenic risk assessment: Comparison of estimated safe doses for rats and mice. *Environ. Health Perspect.* 72:305–9.

Collins, J.M., D.S. Zaharko, R.L. Dedrick, and B.A. Chabner. 1986. Potential role for preclinical pharmacology in phase I clinical trials. *Cancer Treat. Rep.* 70:73–80.

Collins, J.M., C.K. Grieshaber and B.A. Chabner. 1990. Pharmacologically guided phase I clinical trials based upon preclinical drug development. *J. Natl. Cancer Inst.* 82:1321–6.

Crouch, E.A.C. 1983. Uncertainties in interspecies extrapolations of carcinogenicity. *Environ. Health Perspect.* 50:321–7.

Crouch, E., and R. Wilson. 1979. Interspecies comparison of carcinogenic potency. *J. Toxicol. Environ Health,* 5:1095–118.

Crump, K., B. Allen, and A. Shipp. 1989. Choice of dose measure for extrapolating carcinogenic risk from animals to humans: an empirical investigation of 23 chemicals. *Health Physics* 57(Sup.1):387–93.

Crump, K.S., A Silvers, P.F. Ricci, and R. Wyzga. 1985. Interspecies comparison for carcinogenic potency to humans. *In:* Ricci, P.F. (ed.) *Principles of Health Risk Assessment.* Prentice-Hall, Englewood Cliffs, N.J.

Crump, K., B. Allen, and A. Shipp. 1987. An investigation of how well human carcinogenic risk from chemical exposures can be predicted by animal data, with emphasis upon selection of dose measure for extrapolation from animals to humans. Presentation at Twenty-Sixth Hanford Life Sciences Symposium; Modeling for Scaling to Man: Biology, Dosimetry, and Response. Richland, Washington. October 20–23, 1987.

Cutler, R.G., and I. Semsei. 1989. Development, cancer and aging: possible common mechanisms of action and regulation. J. Gerontol. 44:25–34.

Davidson, I.W.F., J.C. Parker, and R.P. Beliles. 1986. Biological basis for extrapolation across mammalian species. *Reg. Toxicol. Pharmacol.* 6:211–37.

Dedrick, R.L. 1973. Animal scale-up. *J. Pharma. Biopharm.* 1:435–61.

Dedrick, R.O., K.B. Bischoff, and D.S. Zaharko. 1970. Interspecies correlation of plasma concentration history of methotrexate (NSC–740). *Cancer Chemother. Rep. Pt. 1,*54:95–101.

Freireich, E.J., E.A. Gehan, D.P. Rall, L.H. Schmidt, and H.E. Skipper. 1966. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey and man. *Cancer Chemother. Rep.* 50:219–44.

Gaylor, D.W., and J.J. Chen. 1986. Relative potency of chemical carcinogens in rodents. *Risk Anal.* 6:283–90.

Gaylor, D.W., and R.L. Kodel. 1980. Linear extrapolation Algorithm for low dose risk assessment of toxic substances. *J. Environ. Pathol. Toxicol.* 4:305–12.

Gillette, J.R. 1985. Biological variation: the unsolvable problem in quantitative extrapolation from laboratory animals and other surrogate systems to human populations. *In:* Hoel, D.G., R.A. Merrill, and F.P. Perera (eds.) *Risk Quantitation and Regulatory Policy, Banbury Report 19.* Cold Spring Harbor Laboratory, Cold Spring Harbor, L.I., N.Y.

Gillette, J.R. 1987. Dose, species, and route extrapolation: general aspects. *In:* National Research Council. *Pharmacokinetics in Risk Assessment: Drinking Water and Health, Vol. 8.* National Academy Press, Washington, DC.

Gold, L.S., C.B. Sawyer, R. McGaw, G.M. Buckman, M. DeVecidna, R. Levinson, N.K. Hooper, W.R. Hevendor, L. Bernstein, R. Peto, M.C. Pike, and B.N. Ames. 1984. A carcinogenic potency database of the standardized results of animal bioassays. *Environ. Health Perspect.* 58:9–319.

Hattis, D. 1990. Pharmacokinetic principles for dose rate extrapolation of carcinogenic risk from genetically active agents. *Risk Anal.* 10:303–16.

Hattis, D. 1991. Use of biological markers and pharmacokinetics in human health risk assessment. *Environ. Health Perspect.* 89:230–8.

Hill, T.A., R.C. Wands, and R.W. Leukroth, Jr. (eds.). 1986. *Biological Bases for Interspecies Extrapolation of Carcinogenicity Data.* Prepared for Food Safety and Applied Nutrition, Food and Drug Administration, Department of Health and Humas Services, Washington, DC, under contract to Federation of American Societies for Experimental Biology, FDA contract N° 223–83–2020.

Hoel, D.G. 1977. Some problems in low-dose extrapolation. *In:* Hiatt, H.H, J.D. Watson, and J.A. Winsten (eds.) *Origins of Human Cancer: Book C, Human Risk Assessment. Cold Spring Harbor Conferences on Cell Proliferation, Vol. 4.* Cold Spring Harbor Laboratory, Cold Spring Harbor, L.I., N.Y.

Hogan, M., and D.G. Hoel. 1982. Extrapolation to man. *In:* Hayes, A.W. (ed.) *Principles of Toxicology.* Raven Press, New York.

Ings, R.J.M. 1990. Interspecies scaling and comparisons in drug development and toxicokinetics. *Xenobiotica* 20:1201–31.

Kaldor, J.M., N.E. Day, and K. Hemminki. 1988. Quantifying the carcinogenicity of antineoplastic drugs. *Eur. J. Can. C.* 24:703–11.

Kleiber, M. 1932. Body size and metabolism. *Hilgardia* 6:315–53.

Kleiber, M. 1961. *The Fire of Life: An Introduction to Animal Energetics.* Wiley, New York.

Lindstedt, S.L. 1987. Alloemtry: body size constraints in animal design. *In:* National Research Council. *Pharmacokinetics in Risk Assessment: Drinking Water and Health, Vol. 8.* National Academy Press, Washington, DC.

Lindstedt, S.L., and W.A. Calder. 1976. Body size and longevity in birds. *Condor* 78:91–4.

Lindstedt, S.L., and W.A. Calder. 1981. Body size, physiological time, and longevity of homeothermic animals. *Quart. Rev. Biol.* 56:1–16.

Mantel, N., and M.A. Schneiderman. 1975. Estimating safe levels, a hazardous undertaking. *Cancer Res.* 35:1379–86.

Metzger, B., E. Crouch, and R. Wilson. 1989. On the relationship between carcinogenicity and acute toxicity. *Risk Anal.* 9:169–77.

Mordenti, J. 1986. Man versus beast: Pharmacokinetic scaling in mammals. *J. Pharmacol. Sci.* 75:1028–40.

National Academy of Sciences (NAS). 1975. *Pest Control Volume 1: An Assessment of Present and Alternative Technologies.* National Academy Press, Washington, DC.

O'Flaherty, E.L. 1989. Interspecies conversion of kinetically equivalent doses. *Risk Anal.* 9:587–98.

Pinkel, D. 1958. The use of body surface area as a criterion of drug dosage in cancer chemotherapy. *Cancer Res.* 18:853–6.

Raabe, O.G., S.A. Book, and N.J. Parks. 1983. Lifetime bone cancer dose-response relationships in beagles and people from skeletal burdens of ^{226}Ra and ^{90}Sr. *Health Phys.* 44, Supl.1:33–48.

Rall, D.P. 1977. Species differences in carcinogenesis testing. *In:* Hiatt, H.H, J.D. Watson, and J.A. Winsten (eds.) *Origins of Human Cancer: Book C, Human Risk Assessment. Cold Spring Harbor Conferences on Cell Proliferation, Vol. 4.* Cold Spring Harbor Laboratory, Cold Spring Harbor, L.I., N.Y.

Reitz, R.H., A.L. Mendrala, and F.P. Guengerich. 1988. In vitro studies of methylene chloride (MEC) metabolism in human and animal tissues: Use in physiologically-based pharmacokinetic (PB–PK) models. *The Toxicologist* 621.

Sacher, G.A. 1959. Relation of lifespan to brain weight and body weight. *In:* Wolstenholme, G.E.W., and M. O'Conner (eds.) *The Lifespan of Animals.* Little Brown, Boston.

Schmidt-Nielsen, K. 1970. Energy metabolism, body size, and problems of scaling. *Fed. Proc.* 29:1524–32.

Schmidt-Nielsen, K. 1975. Scaling in biology: the consequences of size. *J. Exp. Zool.* 194:287–308.

Schmidt-Nielsen, K. 1984. *Scaling: Why is Animal Size so Important?* Cambridge University Press, Cambridge.

Schein, P.S., R.D. Davis, S. Carter, J. Newman, D.R. Schein, and D.P. Rall. 1979. The evaluation of anticancer drugs in dogs, and monkeys for the prediction of quantitative

toxicities in man. *Clin. Pharmacol. Therapeut.* 11:3–40.
Travis, C.C. 1990. Tissue dosimetry for reactive metabolites. *Risk Anal.* 10:317–21.
Travis, C.C., and R.K. White. 1988. Interspecific scaling of toxicity data. *Risk Anal.* 8:119–25.
Travis, C.C., R.K. White, and R.C. Ward. 1990. Interspecies extrapolation of pharmacokinetics. *J. Theor. Biol.* 142:285–304.
U.S. EPA. 1984. Health assessment document for epichlorohydrin. EPA–600/8–83–03F. Available from National Technical Information Service, Springfield, VA. PB85–132363/AS.
U.S. EPA. 1987a. Technical analysis of new methods and data regarding dichloromethane hazard assessment. EPA/600/8–87/029A (Review Draft, June 1987). Available from National Technical Information Service, Springfield, VA. PB87–228557/AS.
U.S. EPA. 1987b. Update to the health assessment document and addendum for dichloromethane (methylene chloride): pharmacokinetics, mechanism of action, and epidemiology. EPA/600/8–87/030A (Review Draft, July 1987). Available from National Technical Information Service, Springfield, VA. PB87–228565/AS.
Vocci, F., and T. Farber. 1988. Extrapolation of animal toxicity data to man. *Regul. Toxicol. Pharmacol.* 8:389–98.
Voisin, E.M., M. Ruthsatz, J.M. Collins, and P.C. Hoyle. 1990. Extrapolation of animal toxicity to humans: interspecies comparisons in drug development. *Regul. Toxicol. Pharmacol.* 12:107–116.

[FR Doc. 92–13207 Filed 6–4–92; 8:45 am]
BILLING CODE 6560–50–M

☆U.S. Government Printing Office : 1992 - 312-014/40149