Towards Confidence Bounds for IRIS Values that Do Not Perpetuate the Underestimation of Risk and False Security about Safety

EPA IRIS Workshop

October 16, 2014; Arlington, VA

Adam M. Finkel, Sc.D., CIH

Fellow and Executive Director, Penn Program on Regulation, Univ. of Pennsylvania Law School

Professor of Environmental and Occupational Health, Rutgers School of Public Health

Member, *Science and Decisions* Committee, National Research Council (2006-2008)

afinkel@law.upenn.edu

Six Themes for this Talk:

1. *IRIS Must Not Abandon its Primary Purpose as it Seeks to Evolve a Secondary Purpose.*

2. *EPA Must Not Ignore the Central Methodologic Recommendation of the NAS "Silver Book"—Move Away from Bright Line Reference Values (RfC/RfD) for Non-Carcinogens, in Favor of Estimating Dose-Response Functions.*

3. *EPA Can and Should Expeditiously Develop Common-Sense Procedures for Estimating and Communicating the Parameter Uncertainty in IRIS Potency Values*

4. EPA Can And Should Adopt the NAS "Silver Book" Recommendations to (Finally!) Take Account of Interindividual Variability in Susceptibility to Carcinogenesis.

5. The Problem of Model Uncertainty is More Vexing—EPA Should Publish Multiple Potency Estimates when Two or More Fundamentally Irreconcilable Models are Sufficiently Plausible—but should not Abandon its Evidence-Based Default Models while it Awaits Sufficient Evidence to the Contrary.

6. The Workshop Question About the Problem of Estimates being "Overly Conservative" Reveals Inappropriate Bias—Current IRIS Estimates are in Some Important Respects not "Conservative" AT ALL.

"The slow pace of IRIS threatens public health… Rough-and-ready estimates are often sufficient for policy-making, and are better than nothing. IRIS should include information from private groups and other governments, and apply available techniques for calculating the risks of chemicals for which there are little data."

- George Gray and Josh Cohen (*Nature,* 9/6/2012, pp. 27-28)

1-Bromopropane: Ample Data, no IRIS Entry

1999– reproductive LOAEL (animals): 200 ppm

1999– nominated for NTP bioassay by OSHA

1999– Swiss circuit board maker ceases use of 1-BP: "there is a weight of evidence that should sound warning bells to any thinking person."

2002-04– case reports of irreversible neuropathy in workers at ≅ 100 ppm

2004– human LOAEL (loss of vibratory sense in toes): 1.1 ppm

2009– NTP bioassay published; 9/50 female mouse lung tumors (1/50 controls) at 62.5 ppm [$q_1^* \cong 2 \times 10^{-3}$ per ppm]

2010– "60 female workers in four 1-BP factories demonstrated dose-dependent neurological and hematological effects of 1-BP exposure with a LOAEL of 1.28 ppm for loss of vibration sense in toes"

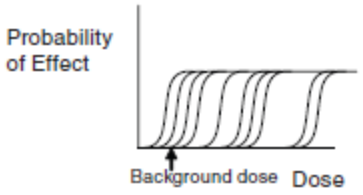
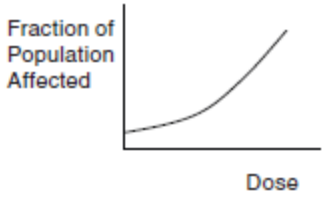
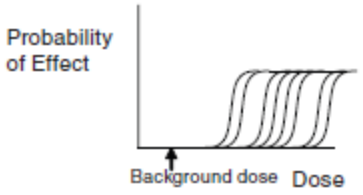
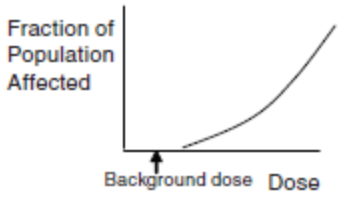
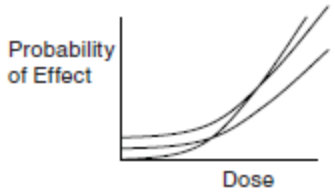
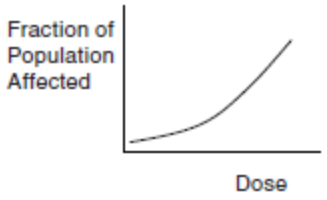
Conceptual Models for Low-Dose-Response	Individual Dose-Response	Population Dose-Response
1. An individual's: Nonlinear The population: Linear		
2. An individual's: Nonlinear The population: Nonlinear		
3. An individual's: Linear The population: Linear		

FIGURE 5-10 Examples of conceptual models to describe individual and population dose-response relationships.

BOX 5-2 Conceptual Model 1: Default Linear Low-Dose Extrapolation for Phosgene

1. Assume uncertainty in all parameters can be characterized by a lognormal distribution, with standard deviation represented by σ.
2. BMD_{10} (human equivalent concentration) = 170 $\mu g/m^3$, with 95%-tile lower bound 30 $\mu g/m^3$ variability in animal BMD, with a difference between lower 95% bound and median of 5.7-fold (because 5.7=170/30):

   $\sigma_{\text{Animal BMD}} = \log(5.7)/1.645 = 0.46$

   (Division by the 95% confidence bound is 1.645 standard deviations from the median in the standard normal distribution.)
3. The human equivalent concentration accounts for cross-species differerence in pharmacokinetics but not pharmacodynamics.

   Assume, as in Hattis et al. 2002, that $\sigma_{\text{logA}\rightarrow\text{H}} = 0.42$
4. Median human POD:

   Adjust for subchronic to chronic study length, as in Hattis et al. 2002, by a factor of 2:

   170 $\mu g/m^3 \div 2 = 85\ \mu g/m^3$

   Assume the uncertainty ($\sigma^2_{\text{logSC}\rightarrow\text{C}}$) in the adjustment, as in Hattis et al. 2002:

   $\sigma_{\text{logSC}\rightarrow\text{C}} = \log[2.17] = 0.34$
5. Uncertainty in the human POD ($\sigma_{\text{logHuman POD}}$):

   $\sigma^2_{\text{logHuman POD}} = \sigma^2_{\text{logAnimal BMD}} + \sigma^2_{\text{logA}\rightarrow\text{H}} + \sigma^2_{\text{logSC}\rightarrow\text{C}}$

   $\sigma^2_{\text{Human POD}} = 0.46^2 + 0.42^2 + 0.34^2 = 0.71^2$
6. Lower 95% confidence bound on Human POD =

   (median human POD)/$10^{[(1.645)(\sigma\text{logHuman POD})]}$ = $85/10^{[(1.645)(0.71)]} = 85/14.7 = 5.8\ \mu g/m^3$
7. Linear extrapolation to risk-specific dose - inflammation of 1 in 10^5 people would be affected:

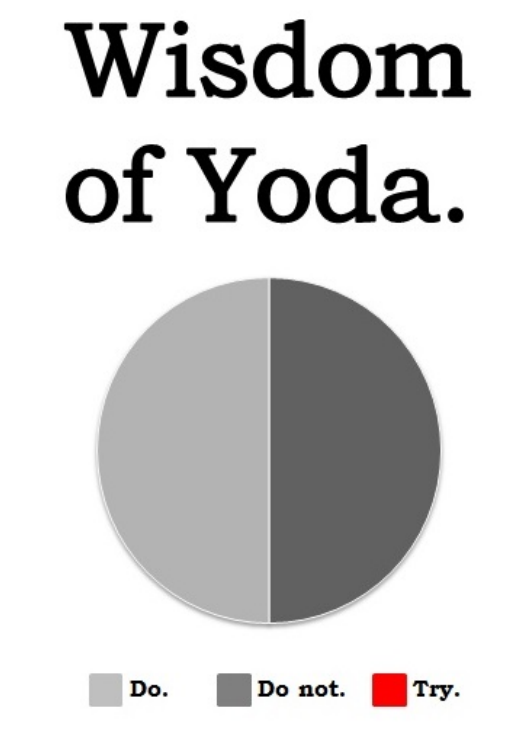
   risk-specific dose = $10^{-5} \times (85/0.1) = 0.0085\ \mu g/m^3$, with lower bound $0.00058\ \mu g/m^3$
8. Estimate risk at different doses: for example, at 0.01 $\mu g/m^3$, three people in 10^5 (median estimate) would be affected.

"Bright lines" are useful for binary conditions (e.g., pregnant/not). They are much less useful when grafted onto continuous variables (e.g., a list of U.S. cities that are "far away from Philadelphia").

A TLV or PEL in effect treats (0.9 * OEL) and (1.1 * OEL) as *completely different*– do we think this is so?

Meanwhile, it treats (1.1 * OEL) and (110 * OEL) as *the same*– and do we think this is so?

[ditto for (OEL/1.1) and (OEL/110)]



Uncertainty *versus* Variability

-- same mathematics and terminology, but…

U

- A property of us
- Sometimes reducible through further study
- Forces decisions about whether to be "better safe than sorry"

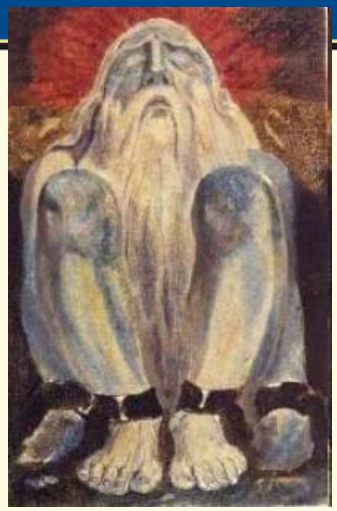
V

- A property of nature
- Irreducible (but understandable)
- Forces decisions about "*who* gets to be safe, who ends up sorry"

Is this door frame a good use of "best estimate" public policy?

Is this ladder built with wasteful "conservatism"?

How Can Hiding Uncertainty Lead The Decision-Maker Astray? A New Typology:



("In every voice, in every ban,
The mind-forged manacles I hear")
-William Blake, *London* (1794)

1. Imposing an "estimanacle," while also mis-estimating it! *In particular, (for right-skewed uncertainties) you cannot estimate the mean accurately without mapping the tail!* (that is, the mean is supposed to make uncertainty analysis unncessary, but you need the uncertainty analysis to estimate the mean)

2. Imposing an unbiased "estimanacle," but when a fully-informed decision-maker might choose a different estimator (in the context of IRIS, a different balance between T1 and T2 errors, or between underestimation and overestimation)

In the spirit of Winston Churchill ("Madam, we've already established that– now we are trying to establish the price"), I offer a syllogism:

1. Human beings differ one to another in their susceptibility to carcinogenesis (a.k.a., their individual risk at a particular exposure);
2. A single number (a cancer potency factor, an ED_{xx}, a risk at an exposure below the POD, an MOE, etc., etc.) will correctly predict individual risk to *someone* within the spectrum of human susceptibility;
3. Therefore, this number will underpredict risk to everyone who is more susceptible than this person, and over-predict risk to everyone who is less susceptible.

"(We've already established that: now by how much…?)"

How many of us have our cancer risks under-estimated by EPA, and by how much, concerns me, because it leads to under-regulation. Others may well be concerned with the converse (over-estimation of individual risk).

Everyone (even the economists) should be concerned with whether EPA's estimates of ***population risk*** ("body counts") are biased low:

Population risk = (mean risk) * (size of population)

Mean risk = Potency * (mean susceptibility) * (mean exposure)

Mean susceptibility > (median susceptibility)

TABLE H-1 Examples of Common Predisposing Factors

Predisposing Factor	Mechanism Influencing Susceptibility to Cancer
A. Temporal Factors[a]	
• Circadian rhythms	
• Changing ingestion and inhalation characteristics during life	
• Depression and stress	
B. Nutritional Factors[b]	
• Vitamin A and iron deficiencies	May increase susceptibility to carcinogenic hydrocarbons
• Dietary-fiber intake	Insufficient intake may increase residence time of carcinogens in contact with epithelium of digestive tract
• Alcohol intake	May affect susceptibility through effect on liver
C. Concurrent Diseases[c]	
• Respiratory tract infections and bronchitis	May predispose lungs to cancer by disturbing pulmonary clearance or promoting scarring
• Viral diseases, e.g., Hepatitis B	May activate proto-oncogenes and cause liver necrosis and regeneration
• Hypertension	May increase the potential for DNA damage in peripheral lymphocytes

[a]Data from Fraumeni, 1975; Borysenko, 1987.
[b]Data from Calabrese, 1978.
[c]Data from Warren and Weinstock, 1987.

TABLE H-2 Examples of Rare Predisposing Factors[a]

Predisposing Factor	Mechanism Influencing Susceptibility to Cancer
• Ataxia-telangiectasia	Chromosomal fragility, causing sensitivity to agents that increase genetic recombination
• Bloom's syndrome	Hypermutability
• Chediak-Higashi syndrome	Depletion of "natural killer" cells that combat incipient malignancies
• Down's syndrome trisomy 21	Tenfold excess leukemia risk
• Duncan's disease	Lymphoma in those infected by Epstein-Barr virus
• Epidermodysplasia verruciformis	Skin carcinoma associated with chronic infection with human papilloma virus
• Familial polyposis coli	Mutation in APC tumor suppressor gene leads to benign colonic growths that are predisposed to malignant transformation
• Fanconi's anemia	Possible deficiency of enzymes that scavenge active oxidizing species
• Glutathione reductase deficiency	Very high excess risk of leukemia
• Hereditary retinoblastoma	Predisposition to retinal cancer due to mutation of one allele of a tumor suppressor gene
• Li-Fraumeni syndrome	Germline mutation in the p53 tumor suppressor gene predisposes to multiple carcinomas and sarcomas
• X-linked agammaglobulinemia	Immune deficiency, predisposing to leukemia
• Xeroderma pigmentosum	Inability to repair some kinds of DNA damage, predisposing to skin cancer caused by ultraviolet radiation

[a]Data from Swift et al., 1991; Orth, 1986; Kinzler et al., 1991; Nishisho et al., 1991; Groden et al., 1991; Cleaver, 1968; Friend et al., 1986; Harris, 1989.

Human Interindividual Variability in Steps along the Pathway to Carcinogenesis

(Hattis and Barlow, *Human and Ecological Risk Assessment*, 1996)

Category	# Data Sets	σ(ln X) (90% c.i.)
Metabolic Activation	22	0.58 (0.30 – 1.1)
Detoxification	19	0.67 (0.2 – 1.6)
DNA Repair	18	0.75 (0.31 – 1.5)
"Complex" (mixed in vivo *measure-ments)*	5	0.95 (0.38 – 1.9)
OVERALL		1.5 (0.61 – 3.1)

From ***Science and Judgment in Risk Assessment*** (NRC 1994):

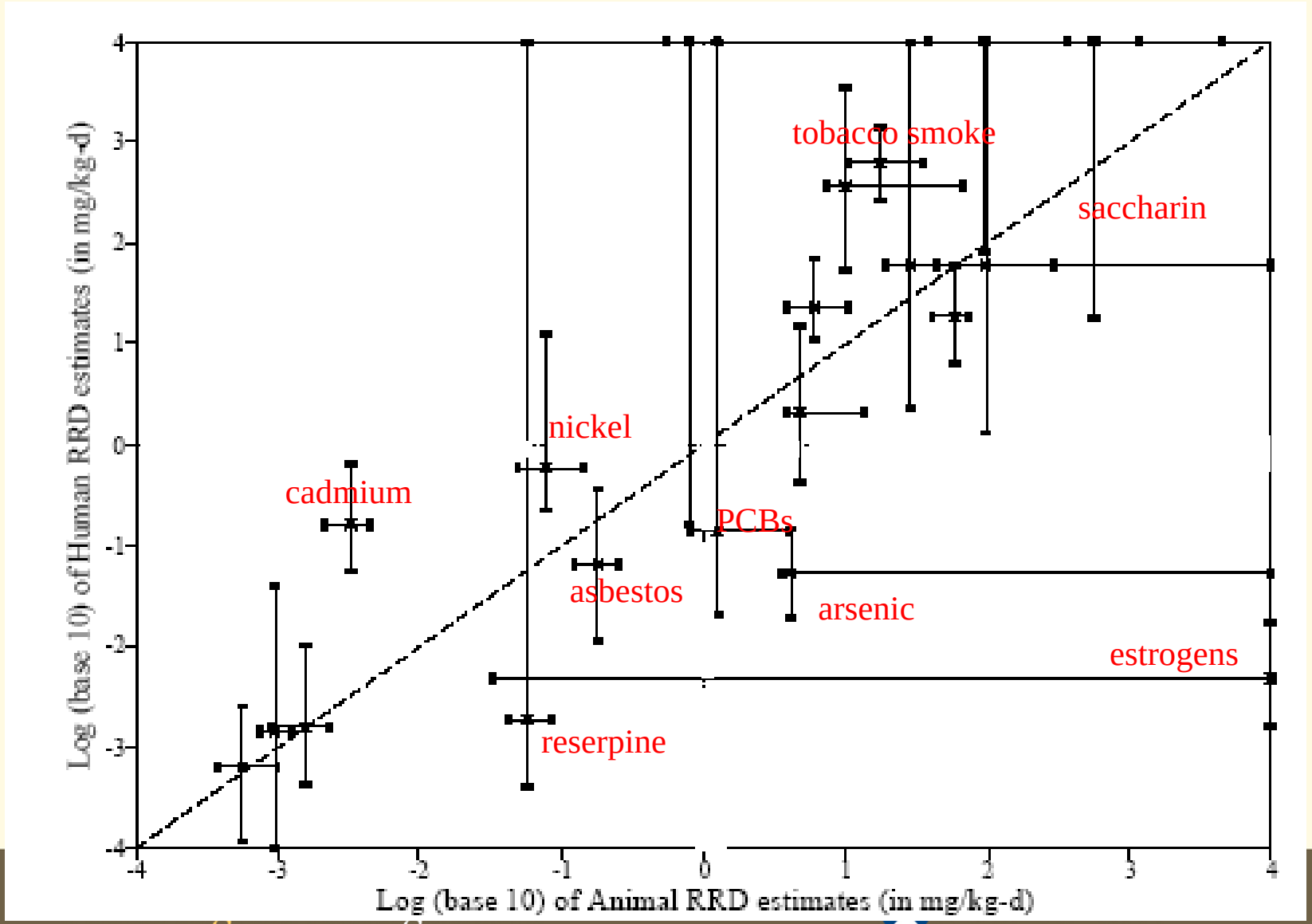
"Recommendation: EPA should adopt a default assumption for susceptibility … EPA could choose to incorporate into its cancer risk estimates for individual risk a "default susceptibility factor" greater than the implicit factor of 1 that results from treating all humans as identical. EPA should explicitly choose a default factor greater than 1 if it interprets the statutory language [in the Clean Air Act Amendments of 1990: "the individual most exposed to emissions"] to apply to an individual with high exposure and above-average susceptibility."

"It is possible that ignoring variations in human susceptibility may cause significant underestimation of population [cancer] risk."

A Colossal *Non Sequitur:*

"The EPA has considered [the NAS recommendation] but has decided not to adopt a quantitative default factor for human differences in susceptibility [to cancer] when a linear extrapolation is used. In general, the EPA believes that the linear extrapolation is sufficiently conservative to protect public health. Linear approaches from animal data are consistent with linear extrapolation on the same agents from human data (Goodman and Wilson, 1991; Hoel and Portier, 1994)"

-- EPA Proposed Guidelines for Carcinogen Risk Assessment (1996)


tobacco smoke
saccharin
nickel
cadmium
PCBs
asbestos
arsenic
estrogens
reserpine
Log (base 10) of Human RRD estimates (in mg/kg-d)
Log (base 10) of Animal RRD estimates (in mg/kg-d)
4
3
2
1
0
-1
-2
-3
-4
-4
-3
-2
-1
0
1
2
3
4

The linear default is thought to generally provide an upper-bound calculation of potential risk at low doses, for example, a 1/100,000 to 1/1,000,000 risk. This upper bound is thought to be public-health protective at low doses for the range of human variation, considering the typical Agency target range for risk management of 1/1,000,000 to 1/10,000, although it may not completely be so (Bois et al., 1995) if pre-existing disease or genetic constitution place a percentage of the population at greater risk from exposure to carcinogens. The question of what may be the actual variation in human susceptibility is one that was discussed in general in the NRC (1994) report, as well as the NRC report on pesticides in children and infants (NRC, 1993b). NRC has recommended research on the question, and EPA and other agencies are conducting such research. Given the current state of knowledge, EPA will assume that the linear default procedure adequately accounts for human variation unless there is case-specific information for a given agent or mode of action that indicates a particularly susceptible subpopulation or lifestage, in which case the special information will be used.

NAS "Science and Decisions, 2009

An assumption that the distribution is lognormal is reasonable, as is an assumption of a difference of a factor of 10 to 50 between the median and upper 95th percentile people… *It is clear that the difference is significantly greater than the factor of 1, the current implicit assumption in cancer risk assessment.* In the absence of further research leading to more accurate distributional values or chemical-specific information, the committee recommends that EPA adopt a default distribution or fixed adjustment value for use in cancer risk assessment. A factor of 25 would be a reasonable default value to assume as a ratio between the median and upper 95th percentile persons' cancer sensitivity … For some chemicals, as in the 4-aminobiphenyl case study below, variability due to interindividual pharmacokinetic differences could be greater.

The suggested default of 25 will have the effect of increasing the population risk (average risk) relative to the median person's risk by a factor of 6.8If the risk to the median human were estimated to be 10^{-6}, and a population of one million persons were exposed, the expected number of cases of cancer would be 6.8 rather than 1.0.

Conclusions on Susceptibility and Defaults:

- Distributions accounting for uncertainty and interindividual variability are preferable to point estimates.
- EPA has stated for 25+ years that its point estimates of cancer risk are "plausible upper bounds, and could be as low as zero": ***the first statement is false, and the second is misleading*** (a linear term in the LMS polynomial of zero is a totally different concept than "zero potency.***")***
- A plausible upper bound would account for the most basic characteristic of human beings (biological individuality); a zero lower bound would require a sensible attitude towards defaults and departures therefrom.

Illogic on Defaults:

(from final EPA Cancer Guidelines)

Rather than viewing default options as the starting point from which departures may be justified by new scientific information, these cancer guidelines view a critical analysis of all of the available information that is relevant to assessing the carcinogenic risk as the starting point from which a default option may be invoked if needed to address uncertainty or the absence of critical information.

EPA's Human Health Research Program is strategically aimed at providing the methods, tools, and data needed to improve risk assessments to protect public health. The primary goal of the program is to reduce reliance on default assumptions and simplified approaches used in risk assessments in the absence of conclusive data.

Anonymous Footnote, Chapter 6 of **Science and Decisions:**

The problem with EPA's new formulation is that a policy of "retreating to the default" if the chemical- or site-specific data are "not usable" ignores the vast quantities of data (interpretable via inferences with a sound theoretical basis) that *already support* most of the defaults EPA has chosen over the past 30 years. In order for a decision to not "invoke" a default to be made fairly, data supporting the inference that a rodent tumor response was irrelevant would have to be weighed against the *data* supporting the default inference that such responses are generally relevant (see, for example, Allen et al 1988), data supporting a possible nonlinearity in cancer dose-response would have to be weighed against the data supporting linearity as a general rule (see, for example, Crawford and Wilson 1996), data on pharmacokinetic parameters would have to be weighed against the data and theory supporting allometric interspecies scaling (see, for example, Clewell et al 2002), and so on.

In short, this Member of the Committee sees most of the common risk assessment defaults *not* as "inferences we retreat to because of the absence of information," but rather as "inferences we generally endorse *on account of the information*."

Therefore, EPA's stated goal of "reducing reliance on defaults" *per se* is problematic; it raises the question of why a scientific-regulatory agency would ever want to reduce its reliance on those inferences that are supported by the most substantial theory and evidence. This member of the Committee certainly endorses the idea of reducing EPA's reliance on *those defaults* that are found to be outmoded, erroneous, or correct in the general case but not in a specific case—but identifying those inferior assumptions is exactly what a system of departures from defaults, as recommended in the Red Book, in *Science and Judgment*, and in this report, is designed to do.

EPA should modify its language to make clear that across-the-board skepticism about defaults is not scientifically appropriate. This member urges EPA to delineate what evidence will determine how it makes these judgments, and how that evidence will be interpreted and questioned—and EPA's current policy (yet again) sidesteps these important tasks.

Many of us (see, e.g., Chapter 5 in the NAS **Science and Decisions** report) believe that the "divide by 100 and pray" method of setting non-cancer exposure limits is insufficiently protective. For those substances where humans are truly 10x more sensitive than test animals, and for those humans who are truly 10x more susceptible than the median person, their risk at the NOAEL/100 (the RfC) will be the SAME as the animals' risk at the NOAEL– which is to say, perhaps 5-10 chances **per 100**.

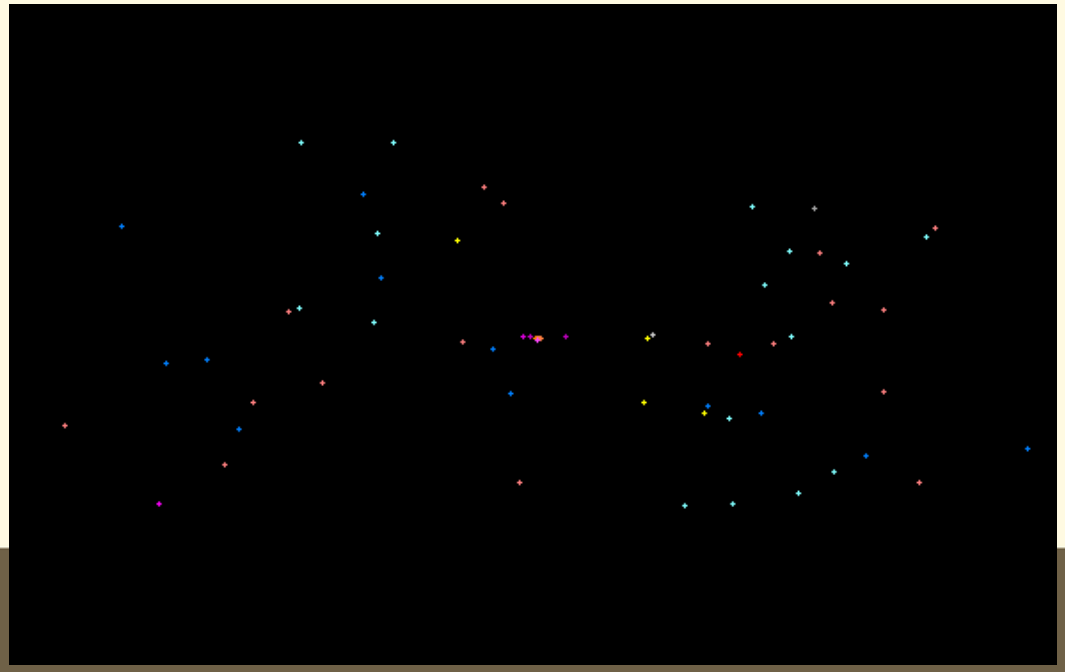
Therefore, exposures 10, 100, 1000 times HIGHER than the RfC may be barbaric.

Scientific Reaction to HSIA Research: 3

"This interpretation of mRNA distribution is profoundly in error and contradicts some of the most well established and fundamental principles of molecular biology… Finding mRNA in the nucleus is unsurprising and uninformative about the eventual location of the protein products."

--Dr. Lorenz Rhomberg, Gradient Corp./Harvard School of Public Health (62 FR 1526)

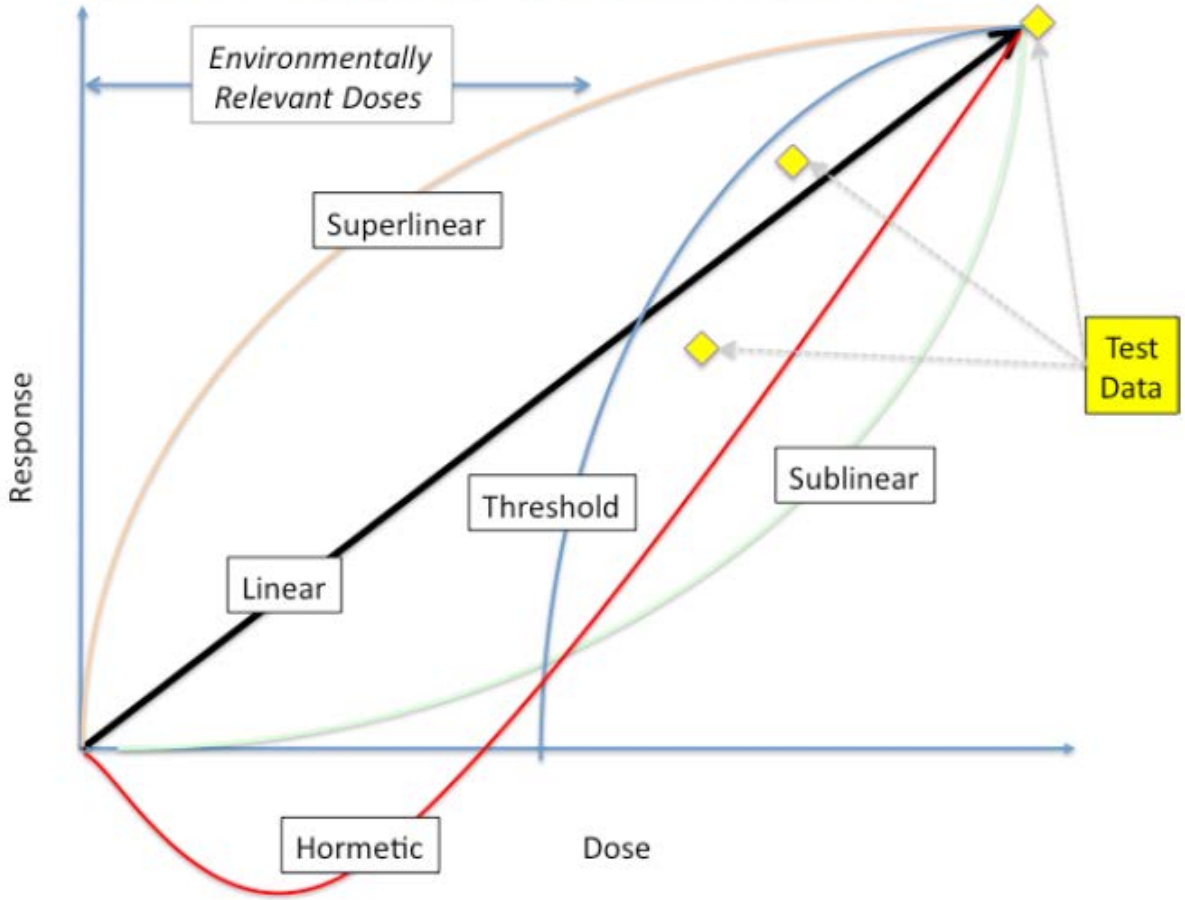
Galileo's "epidemiology" was an instrument powerful enough to discover 4 moons orbiting Jupiter.



The 63 OTHER Jovian moons were not visible to Galileo, but they were always there…

MERCATUS CENTER AT GEORGE MASON UNIVERSITY

FIGURE A: ALTERNATIVE WAYS TO EXTRAPOLATE FROM HIGH TO LOW DOSES



Source: Author's calculations.

The observed data are VASTLY closer to the "regulatory windows" than they are to any "near-zero" region (the previous figure is useful but visually misleading):