

Critical Review

TOXICOLOGY REWRITES ITS HISTORY AND RETHINKS ITS FUTURE: GIVING EQUAL FOCUS TO BOTH HARMFUL AND BENEFICIAL EFFECTS

EDWARD J. CALABRESE*

School of Public Health and Health Sciences, Department of Public Health, University of Massachusetts, Amherst, Massachusetts, USA

(Submitted 2 May 2011; Returned for Revision 17 August 2011; Accepted 25 August 2011)

Abstract—This paper assesses how medicine adopted the threshold dose–response to evaluate health effects of drugs and chemicals throughout the 20th century to the present. Homeopathy first adopted the biphasic dose–response, making it an explanatory principle. Medicine used its influence to discredit the biphasic dose–response model to harm homeopathy and to promote its alternative, the threshold dose–response. However, it failed to validate the capacity of its model to make accurate predictions in the low-dose zone. Recent attempts to validate the threshold dose–response indicate that it poorly predicts responses below the threshold. The long marginalized biphasic/hormetic dose–response model made accurate predictions in these validation studies. The failure to accept the possibility of the hormetic-biphasic dose–response during toxicology’s dose–response concept formative period, while adopting the threshold model, and later the linear no-threshold model for carcinogens, led toxicology to adopt a hazard assessment process that involved testing only a few very high doses. This created the framework that toxicology was a discipline that only studied harmful responses, ignoring the possibility of benefit at low doses by the induction of adaptive mechanisms. Toxicology needs to assess the entire dose–response continuum, incorporating both harmful and beneficial effects into the risk assessment process. *Environ. Toxicol. Chem.* 2011;30:2658–2673. © 2011 SETAC

Keywords—Hormesis Hormetic Biphasic Dose–response Risk assessment

INTRODUCTION

To claim that the scientific discipline of medicine got the dose–response wrong and with this error damaged our health, environment, and economy sounds wrong, irresponsible, and unfair to such a dignified and life-serving profession. This accusation seems right off the pages of an attention-grabbing national tabloid one may often see while waiting in line to pay for groceries, rather than an academic appraisal. The problem is that the accusation has a compelling and detailed historical record, a record that has taken more than two decades to unravel and reconstruct. This story is a historical detective adventure, initially set in northern Germany in the modestly sized academic city of Greifswald during the latter decades of the 19th century, later transforming into an international affair, reaching the highest levels of government in the United States and in multiple European countries, with documentation of its continuing international presence in the control of legislation, major governmental programs, university curricula, as well as the lives of citizens and their health judgments [1–3].

Although many articles and book chapters have been written on the history of toxicology and pharmacology, the perspective offered here is unique. The present paper contends that the most fundamental principle of toxicology and pharmacology, the dose–response relationship, arose out of a dispute between two intense professional rival organizations—traditional medicine and homeopathy. It will be shown that homeopathy was the first of the two organizations to claim that what they believed was the most fundamental nature of the dose–response: asserting it to be the biphasic dose–response, later to be called hormesis. Having lost the benefit of first discovery or procla-

mation, traditional medicine led a powerful, prolonged, and unrelenting attack on this dose–response concept and its formulator, Hugo Schulz (1853–1932), all in an effort to discredit the model (which Schulz called a law) and their medical opponent, homeopathy. In time, leaders within traditional medicine then proposed their dose–response alternative (the threshold dose–response model) and used their influence to get it established at all levels of scientific society, including government regulatory agencies, academic institutions, the chemical and pharmaceutical industries, and professional societies. As will be seen, they were amazingly successful, but at a very high cost, for in their quest for victory over homeopathy, they failed to validate their dose–response model in the critical low-dose zone, where most human exposures to drugs and chemicals occur.

The claim

Modern medicine is the parent of pharmacology, which begat toxicology, which then produced risk assessment and risk communication, all products of the 20th century. The fields of pharmacology and its then nascent offspring, toxicology, adopted the threshold dose–response model in the 1930s, convincing governmental regulatory agencies to incorporate it into all subsequent regulations for chemicals and drugs. The leaders of these disciplines, both in and out of government, never attempted to validate this model for accurate predictions, where people live, that is, in the low-dose zone, below the toxicological threshold, nor were they asked to do so by generations of legislative leaders and their scientific advisors. At the core of this issue is the denial by these medically dominated fields of the very existence of an alternative dose–response model, called hormesis, from a rival medical concept/organization called homeopathy.

The hormesis concept has long been marginalized in multiple reinforcing ways by the scientific and medical communities.

All Supplemental Data may be found in the online version of this article.
* edwardc@schoolph.umass.edu

Published online 19 September 2011 in Wiley Online Library
(wileyonlinelibrary.com).

This may be inferred by its long exclusion from the leading textbooks of medicine, pharmacology, and toxicology and, as a consequence, its lack of presence in academic curricula and the classroom. Professional societies have likewise denied hormesis a presence at their annual and regional meetings, thereby preventing professional attendees from getting the opportunity to learn about this concept and to observe research presentations on the topic. Of further significance is that the hormesis concept has long been excluded from governmental grant funding processes, a tactic that would ensure an untenured professor academic failure if he or she pursued a possible hormetic area of research interest. The hormesis concept also has been excluded from affecting the development of environmental and medically related legislation, which develop broad-reaching policies, direct the flow of money and resources, and influence a vast array of human behaviors. In effect, modern medicine and its pharmacological and toxicological offspring incorporated an unproven and nonvalidated dose–response model, called the threshold dose–response, into its profession, made it their gold standard (default) model, while purposefully and severely marginalizing its opponent’s model (hormesis), letting it function as a historical artifact of a discredited medical practice and thus giving the hormesis concept the equivalent of a professional death sentence.

These medically related fields passed on the threshold dose–response model and its hazard assessment testing and risk assessment schemes to ever newer generations of pharmacologists, toxicologists, risk assessors, and risk communicators. These scientists, physicians, and social scientists knew little of their history and even less of the dose–response machinations by their medical grandparents. They too were unaware that they had been professionally misled and manipulated with the firm expectation that they would do the same to their students as was done to them. This concept orchestration and data censoring by traditional medicine and its disciplinary descendants would occur, and continue to occur, in ostensibly free societies, where people, including the scientific community, were led to believe that they were in control of what to accept or not. They were, however, unaware that political, economic, and pervasive institutional (medicine, academic, governmental regulatory agencies) forces converged to suppress the hormesis concept while promoting an alternative or rival model, the threshold dose–response. Now let us examine what this claim is based on.

The medical rivalry—its unintended consequences

It is now the second decade of the 21st century. While the mid-19th century seems quite distant from today, history has a long reach. Consider race relations in the United States. The United States is still feeling the effects of a Constitution in which Blacks were considered three-fifths of a person for census information and Congressional representation. The U.S. Civil War from 1861 to 1865 was a national trauma beyond comprehension. Even a century later, the country was still trying to figure out how to ensure that African Americans would have equal access to education, jobs, health care, housing, restaurants, and other social and professional venues. History is important; it has deep and entangling roots that often impact the present and future in ways that may not be obvious, but nonetheless are real, powerful, and controlling.

Pharmacology and toxicology had their roots in what we now call traditional medicine. Traditional medicine emerged from what is generally referred to as the era of heroic medicine, when physicians of the 18th and 19th centuries often treated patients rather harshly with blood drawings, bloodsucking

leaches, and highly toxic agents such as arsenic and mercury, while often conning patients into purchasing and ingesting elixirs that could have caused harm. As many may know, the death of George Washington in 1799 was probably accelerated by repeated and extensive blood drawings to cure him of what is now thought to have been a bacterial infection (www.eyewitnesstohistory.com/washington.htm). Furthermore, the only daughter of John Adams, the second president of the fledgling United States, underwent surgery for breast cancer without an anesthetic (www.shsu.edu/~pin_www/T@S/2002/NabbyAdamsEssay.html). It was pretty clear that in the days of heroic medicine, one was fortunate to endure and survive a host of what could only be called barbaric treatments. This statement does not even consider the effects on family members who watched and had to deal with suffering directly related to such heroic medicine practices. Thoughtful physicians of the day often recognized that they were far from being healers but often were the equivalent of an unwilling but necessary torturer. One of these conscience-riddled physicians of the late 18th century who just could not “take it anymore” was Samuel Hahnemann (1755–1843), a very bright German, who sought and created an alternative to the practice of torturous heroic medicine [4–6].

Hahnemann challenged the establishment by creating the medical practice of homeopathy, an action somewhat akin to the earlier machinations of Martin Luther when he posted his Ninety-Five Theses in 1517, which eventually led to the creation of the first Protestant church. Let us put it this way: Martin Luther’s challenge did not go over well with the Pope and his bishops; the actions of Hahnemann created a similar set of enemies, only his were in medicine, not theology. In both cases, the stakes were high. On the plane of idealism, Hahnemann’s fight was about life and death in this world, whereas in the case of the church, it was about death and life in the next. In the world of pragmatism, this intense competition was also about power, politics, influence, money and, of course, control. In addition, Hahnemann was not easy to like. Brilliant though he probably was, he offered an equal dose of bitter and arrogant invective that simply fueled the conflict, creating a long line of personal enemies just awaiting their time for payback. Few attempts were made at compromise. Whether homeopathy cured patients really was not the issue, at least at the time of Hahnemann, and in fact, has nothing to do with the premise of the present paper. The scheme worked out by Hahnemann of using extremely dilute doses of plant-derived extracts as homeopathic drugs was at least not very likely to injure his patients. In this dimension, homeopathy was probably superior to heroic medicine. It did not torture its patients or speed them along to an early grave, and it may even have provided a healing boost if only via a placebo effect. Indeed, homeopathy was winning the hearts and minds of many adults in Europe and the United States, and, of course, gaining an ever-greater market share [5,7].

The battle between homeopathy and traditional medicine has been longstanding [5,7], the stuff of bitter internecine hostilities, much like warring political parties, opposing churches, or even the heated family feud, in which ghastly homicides can occur. The battles could be intense. It was about which profession was going to win in this most important aspect of life. As with the U.S. civil war, the traditional medicine–homeopathy conflict also has had a long reach that is as intriguing as it is important. It will now be shown how its outcome has profoundly affected the development of toxicology and pharmacology, the testing and safety of drugs and chemicals, the risk assessment process, the risk communication message, as well as

the entire range of environmental health exposure standards for all environmental (e.g., air, water, food, soil, consumer products) media.

Schulz's mistake: Biphase dose–response to homeopathy

In February 1884, the physician and pharmacologist Hugo Schulz first presented evidence of the biphasic dose–response (to be called *hormesis* in 1943 by Southam and Ehrlich [8]), based on experiments assessing the effects of disinfectants on yeast metabolism, at a meeting of the Greifswald Medical Association (<http://www.Medizin.uni-greifswald.de/medverein/Geschichte.htm>) and subsequently published his findings [9,10]. The low-dose stimulatory response was a surprising observation of which Schulz was initially skeptical. However, repeated successful replication experiments led him to be confident that the observed biphasic dose–response was highly reproducible and extended to a wide range of chemical disinfectants ([11]; translation of the Schulz, 1923 autobiography, [12]).

Schulz used these findings to explain a striking series of clinical observations by Bloedau in 1884 ([13]; cited in Schulz [14]), in which a homeopathic preparation (veratrine) was used to successfully treat gastroenteritis. Schulz was so intrigued with these clinical findings that he tested whether such a preparation would directly kill the recently isolated causal bacterium of this disease. However, his experiments indicated that the veratrine was unable to do so, regardless of the dose applied [14]. Although these experimental findings could have led Schulz to conclude that the homeopathic preparation was not an effective treatment of gastroenteritis, they did not. In fact, Schulz hypothesized that the homeopathic treatment was effective but that its mechanism was not directly bactericidal but via the induction of an adaptive response in the patient to resist the infection. After conversations with his colleague Rudolph Arndt, Schulz linked this adaptation hypothesis with his biphasic dose–response observations in the yeast. He then proposed that the low-dose stimulation represented an adaptive process and that this was how low but measurable doses (i.e., not an ultra low, extremely high dilution dose below Avogadro's number) of homeopathic preparations worked. At this point, Schulz came to believe that he had discovered the explanatory principle of homeopathy, later naming it the Arndt-Schulz Law.

Schulz's gift to homeopathy directly led it to become the first of the medical professions to stake a claim on the nature of the dose–response, especially in the low-dose zone. When seen through the lens of history, homeopathy scooped traditional medicine on the key issue of the dose–response and its potential for drug development and patient treatment.

The scooping of traditional medicine on the nature of the dose–response was no small accomplishment for the underdog homeopathy, beating it to the punch on the critical pillar of their profession. Whether anyone at the time truly appreciated the significance of this achievement is not clear. However, the fact that Schulz quickly became an object of vicious criticism and professional ridicule by his medical colleagues suggests that the leaders of the traditional medicine movement understood only too well what was at stake [12]. In his autobiography, Schulz [11] recounted in a striking way how he incurred professional ostracism by his traditional medical colleagues, indicating how he was viewed with suspicion because of his research on homeopathy. He also became the object of a derisive writing campaign that referred to him as the Greifswald Homeopath. These changes in professional relationships occurred soon after his 1885 publication [14] that proposed a low-dose adaptive

response mechanism for the homeopathic preparation veratrine, in effect excluding him from their group. So polarized was the relationship between traditional medicine and homeopathy that Schulz remained an outsider and the object of continuing criticism and judgmental actions for his entire nearly 50-year academic career. In fact, on the occasion of Schulz's retirement, Martius-Rostock [15] reflected on his long and conflicted career, lamenting “What law did Hugo Schulz break that makes him deserving of the boycott exercised by his scientific peers?” He additionally criticized those who distorted Schulz's image “through incomprehensible, idle talk.” Schulz's biphasic dose–response and its meaning would soon be challenged and denigrated on multiple levels, and he and his supportive colleagues, such as the eminent August Beir (the father of spinal anesthesia) (1861–1949) [16], along with it.

The dose–response “take back” by the medical profession was led by Alfred J. Clark (1885–1941), a professor of pharmacology at the University of Edinburgh, who occupied the most coveted academic position in pharmacology in Europe. Clark had worked his way up through the ranks, with professorships in South Africa (1918–1920), London (1920–1926), and Edinburgh (1926–1941). Along the way, he established himself as an expert in quantitative pharmacology. He was able to combine excellent mathematical skills with his training in medicine and pharmacology and was clearly smarter and more focused than even a normally gifted contemporary physician and pharmacological researcher. His unique combination of skills and drive gave him a clear edge on his peers and enemies. Clark also was a meticulous researcher with a flair for writing. He applied his prodigious skills to the shaping of the field. By virtue of his highly successful textbooks [17–20], he taught pharmacologists and toxicologists for approximately a half century, well after his untimely death in 1941 [21,22]. In the foreword to the book *Towards Understanding Receptors*, Robinson [23] referred to the 1937 text by Clark as the “now classic monograph on *General Pharmacology*, a book that had great influence on a number of individuals.” Clark not only used these textbooks to teach pharmacology and toxicology, but also as a vehicle to emasculate homeopathy, Hahnemann, Schulz, and his Arndt-Schulz Law, that is, the hormesis concept (Appendix 1).

The new medicine man: The threshold model

The threshold dose–response model became medicine's alternative to the biphasic dose–response of homeopathy and Schulz. This model would become the driver for therapeutic medicine for the rest of the century and beyond. How Clark achieved his historical milestone of dose–response control is now described. Clark was a first-rate scholar. His textbooks were thorough, detailed, and second to none, at least for his era and the next generation. This was no small accomplishment, and he had the admiration of many. He never left a scientific stone unturned, so to speak. This high level of detailed professionalism brought him to the leadership of his field, along with a likeable, highly principled personality. It also made him the ideal scientific candidate to attack and discredit homeopathy and Schulz. Thus, the fact that Clark failed to present, discuss, and at least try to refute the substantial body of research that supported Schulz's dose–response model, especially when it had been broadly reported, in excellent journals and by scientists of high visibility and accomplishment [24], is highly surprising. Not that Clark did not have a sense for the scientific literature or how to obtain and analyze it—his textbooks and other writings amply illustrate that he was among the best when

it came to digging out information from even the most obscure sources and integrating apparently disparate information into plausible biomedical theory.

Evidence to support this conclusion may be seen in Clark's selective use of references to refute Schulz. For example, his influential text entitled *The Mode of Action of Drugs on Cells* [18], citing Dannenberg [25], argued that Schulz's low-dose stimulatory responses in yeast were attributable to background variation/experimental error, not real treatment effects. Based on these findings and interpretations, Clark then dismissed both the plausibility and the biological significance of the Arndt-Schulz Law. What Clark [18] did not mention was much more significant. First, he failed to show that the doses used in the Dannenberg [25] study were far below those known to cause stimulatory responses in the earlier yeast studies; the highest doses were some 10-fold to 20-fold below the lowest doses associated with such stimulation. Also not mentioned was that the Dannenberg [25] study assessed responses at only a single time point, thereby eliminating the possibility of observing an overcompensation stimulatory response. Clark [18] also neglected to cite studies that supported the Schulz findings, including a recent detailed investigation by Branham [26], which was published in the highly visible *Journal of Bacteriology*. This study was specifically designed to replicate and extend the research of Schulz on the effects of chemical disinfectants on yeast metabolism. Branham [26] tested a similar broad spectrum of agents as used by Schulz (not simply three as reported by Dannenberg [25]), incorporated a detailed dose time component, and assessed responses over a very broad dose range that included doses both above and below the toxicity threshold. Her findings strikingly supported the observations of Schulz, clearly documenting the low-dose stimulation and high-dose inhibition, while revealing that the stimulatory responses resulted from an overcompensation to an initial disruption in homeostasis (that is, toxicity). A dose-dependent toxicity occurred at the first time point, followed by the overcompensation stimulatory effect. The stimulatory effects were also 40 to 80% greater than the control group response. The study of Dannenberg [25], which Clark [18] so highly relied on, was therefore clearly not designed to fairly test the Schulz hypothesis, whereas the Branham [26] study was.

My colleague Linda Baldwin and I further documented this type of scientific misrepresentation in five publications, detailing the historical foundations of hormesis in the biological and biomedical literature, occupying an entire issue of the journal *Human and Experimental Toxicology* [27–31]. Although considerable support was shown for the hormesis concept in the early decades of the 20th century in chemical toxicology, pharmacology, and radiation biology, these findings were also neglected by Clark. The hormetic-biphasic findings were broadly generalizable, often substantial research contributions, and readily obtainable, even without electronic databases. The failure of Clark to recognize and address these studies supporting the hormesis/biphasic dose–response perspective was very damaging to Schulz, adversely affecting the acceptance and utility of his body of work, including his dose–response concept, the Arndt-Schulz Law, as well as the development of pharmacology, toxicology, environmental health, and risk assessment.

As a result, many academicians and researchers in the fields of pharmacology and toxicology succumbed to an appeal to authority. For if Clark could not find support for the biphasic dose–response—he of such high regard, the leading professor of pharmacology among a bevy of other outstanding professors,

an influential governmental advisor, the masterful researcher and textbook writer, a person with a reputation for being objective, fair-minded, comprehensive, and totally professional—then the support for Schulz's model could not have been convincing. In effect, the fields of pharmacology and toxicology allowed Clark to do their thinking on this critical issue.

Clark was a man of great accomplishment and continues to be held in high regard. The University of Edinburgh has a distinguished chair in his name, and graduate fellowships in his honor are awarded by the British Pharmacological Society. These recognitions and honors are not given out lightly or often; they are earned. Nonetheless, despite these honors, Clark failed the scientific community on the most critical and far-reaching concept, the nature of the dose–response.

The threshold dose–response: Historical foundations

The threshold model was not unreasonable; it seemed consistent with much published data and was a concept that resonated with personal experience and common sense. In fact, this may have been why Schulz initially doubted his biphasic dose–response observations when first they appeared in his experiments with yeast. The threshold dose–response concept is believed to have been originally put forward by the legendary French biologist Claude Bernard (1813–1878) [32] within the context of the excretion of glucose. Others extended this concept to the excretion of additional pharmacological/physiological agents such as chloride, urea, and other metabolic products [33–38]. The threshold perspective was then placed within a more general context by Cushny (1866–1926) [39], who developed a simple formula-based model to describe the threshold response. Clark [40] also had some experience in the study of dose–responses, observing a biological threshold for acetylcholine that required approximately 20,000 molecules acting via receptors to produce an initial effect on a heart cell (e.g., isotonic contraction). Clark had been a professor working in Cushny's department and twice replaced him as Department Chair (at the University of London, 1920, and then at Edinburgh, after Cushny's death in 1926) [41,42]. Such research on the threshold concept was further extended in the laboratory of the Nobel Prize winner and British pharmacologist Charles Scott Sherrington, by Russell Aitken and his advisor J.G. Priestley [33]. Although the research of Aitken [33], Cushny [39], and others was in the pharmacological domain, support was also offered in the toxicological [43–45], radiation/occupational health [46–48], and immunological [49] areas for the generalizing of the threshold dose–response concept. Thresholds were also widely observed in numerous other scientific domains, ranging from the behavioral to the physical sciences, supporting a broad and integrative general scientific concept [50]. No need was felt to consider the Schulz alternative dose–response concept.

The anatomy of a scientific takeover

Just how did the threshold dose–response concept get established?

Step 1: Challenge the alternative model

Clark's various publications on the quantitative features of pharmacology and toxicology had both a devastating and lasting impact on Schulz and the hormetic dose–response. So too did his attempts to reach the broader biomedical community, in highly respected journals, such as the *British Medical Journal* [51]. In this journal and in other writings, he lumped homeopathy with numerous garden-variety versions of medical witch-

craft. In this process, he was broadly successful in creating a repeated focus on homeopathy, that is, his real target, emphasizing its high dilution aspects, calling it quackery and then associating Schulz and his work with it. Clark's mistake was that Schulz did not adhere to the high dilution school of homeopathy [12,52], the only segment of homeopathy that Clark addressed. The written contemporary record was clear that Schulz did not support high-dilution homeopathy in theory or practice. High-dilution homeopathy refers to homeopathic practices in which the dose of a therapeutic agent is diluted to such an extent that there are likely no molecules of agent in the medical treatment. The fact that Schulz offered an explanatory principle for homeopathy based on his biphasic adaptive dose-response while rejecting homeopathy's high dilution features made him both a leader and controversial figure in the homeopathy domain. These points were clearly documented while he was alive and broadly expressed in his obituaries in the homeopathic and traditional medicine literature. For example, Paul Wels, an eminent radiation biologist (Schulz died on July 13, 1932) presented a remembrance lecture on Schulz at a meeting of the Greifswald Medical Society on November 5, 1932. In this presentation, entitled "The Life Time Work of Hugo Schulz", which was published in 1933 [53] in a leading pharmacology journal, Wels stated that "Schulz gave the dosage question its entitled place while homeopathy made it laughable." By unfairly linking Schulz to high-dilution homeopathy, Clark sought to undercut his credibility so he would not be taken seriously by the scientific community.

So strong was the leadership of Clark in the domain of preserving the integrity of traditional medicine via his attacks on its opponents like homeopathy and Schulz that this very point was emphasized in eulogies after his own death and were summarized by his physician-psychiatrist son, David Clark (1920–2009), in an insightful biography. That is, he was remembered positively by even Nobel Prize recipients (e.g., Sir Henry Dale, 1936 recipient of the Nobel Prize for biology and medicine) for his key role in the rapid downturn of homeopathy and other forms of quackery [41].

Step 2: Propose your model, get it accepted

The period of concept consolidation involved extensive support for the threshold model in the numerous publications by Clark and his contemporaries. This served as a basis to establish the necessary peer-review-based credibility for the threshold model. Given his mathematical background, Clark recognized the value of integrating his dose-response concept within a biostatistical framework using the newly developed probit dose-response model of two of his colleagues in the early to mid-1930s [54,55]. This mathematical model was derived independently by John Henry Gaddum (1900–1965), another extremely gifted quantitative pharmacologist, and Chester Bliss (1899–1979), an itinerant, yet highly productive, biostatistician who was befriended by Clark at a critical time in his career. As we shall see, Bliss repaid Clark many times over for his professional and personal support. The probit dose-response model then received a major endorsement and a key provision, both provided by the world-renowned and later to be knighted Ronald A. Fisher (1890–1962). Fisher added a procedure to the model called the maximum likelihood estimate as an appendix to the key 1935 paper of Bliss [54]. The monotonic features of the probit model were employed to constrain predicted responses to asymptotically approach the control value at low doses while never being permitted to transition below the control as would occur in the hormetic-biphasic dose-

response model of Schulz. The significance of such a biostatistical manipulation was that it denied the existence of the hormetic dose-response. Dips in the response below the control surely did occur. However, this was assumed to be only attributable to response variability. It was not real in the sense that the findings were reproducible; thus, the best estimate could never be below the control group. The hormesis idea was not only marginalized; it was not considered to have biological credibility. This was the unmistakable take-home message from the intellectual biomedical leadership of that critical concept-defining era.

Step 3: Make your model the standard procedure

This intellectual fusion of the best, brightest, and the most influential biomedical leaders of the day successfully consolidated the threshold dose-response concept into the mainstream of pharmacology and toxicology. Because Clark was also part of the broader elite that created the British Pharmacological Society in 1929, this provided him with ready access to and acceptance by essentially all professors of pharmacology within the entire UK system and their extended pharmacological families in other countries, including the United States, and their journal publication vehicles. This was especially true for the United States, because most serious graduate students viewed a European graduate education experience during this period as a key to obtaining an excellent education, establishing a broad network of professional contacts, and a subsequent position at a leading U.S. academic institution [56]. This was commonly the case for many disciplines, including chemistry, microbiology, botany, and, of course, pharmacology.

The highly positive view of British pharmacological academic elites made acceptance of the Clark perspective on the dose-response even more efficient internationally. Clark's goal of establishing his model yielded rapid success. He had the concept, the textbooks, and the network with its coordinated activities and influence, with no competition or credible opposition.

Success with his pharmacological colleagues, as important as that was, however, was not enough. Numerous other biological subdisciplines had to be educated on the nature of the dose-response, so that medicine could assert dominance of its model in all of biology and the biomedical sciences. This next phase of concept integration was under the per-view of Clark's colleague, Chester Bliss, who wrote a series of publications for the most prominent journals of biological disciplines concerned with dose-response relationships, such as microbiology, entomology, food science, radiation biology, and others [57–62]. These publications, which describe the nature of the dose-response at low dose as well as how to quantitatively assess, interpret, and apply such findings, further ensured the broad acceptance of Clark's perspective. In effect, Bliss sent different versions of the same conceptual paper to many biological subdisciplines, but tailored to the readership of each area. In his writings, Bliss established the term *threshold dose*, defined it, provided various means to estimate it, and integrated it into the mass-action formula used by Clark [18,57], leaving no room for confusion, debate, or compromise. Bliss was absolutely tireless in getting the dose-response message out, ensuring that it influenced the educational process of most students being professionally trained in the biological sciences, later to become the leaders of academia and governmental agencies. Bliss was inspired to continue this educational process during the years before and after the death of Clark. Although Bliss has never been seen as a major player in the course of mid-20th century

science, he nonetheless may well have had the greatest impact for enhancing the implementation of Clark's dose-response perspectives (at the expense of hormesis) into graduate training during the last half of the 20th century. Without Chester Bliss, the potency slope for the acceptance of Clark's dose-response concept would have been much flatter.

Lasting legacy

The biostatistical constraining of the dose-response to approach the control value only asymptotically became a major factor in the regulation of exposures to carcinogens. In the early 1940s, the U.S. National Cancer Institute showed how this would be implemented for a chemical carcinogen. In a study in which the data demonstrated support for an hormetic dose-response relationship for a carcinogenic polycyclic aromatic hydrocarbon, the investigators determined that this agent could not possibly display a risk of disease below that of the control group [63]. They then followed the constraining monotonicity of the probit model, eliminating the possibility of a hormetic dose-response. This constraining concept was later to become policy in the Food and Drug Administration (FDA) and the U.S. Environmental Protection Agency (U.S. EPA) for carcinogen risk assessment and remains so today.

In the early 1950s, one of the famous and most endearing forefathers of U.S. regulatory toxicology, the late Arnold Lehmann (1900–1979), created the framework for the use of safety factors, also called uncertainty factors. This safety factor procedure was easily understood, implemented, and copied by all regulatory agencies in the United States and worldwide.

What Lehmann accomplished was significant. He gave the regulatory world the safety factor. It was based on the assumption of a threshold dose-response model, compliments of A.J. Clark, on whose textbooks Lehmann had been reared academically. In retrospect, Lehman and his colleagues, following the example of Clark, failed to validate the threshold dose-response model at this most critical juncture of regulatory science history, thereby contributing to toxicology's failure of due diligence.

Linearity wins a seat at the risk assessment table

As much as the threshold model dominated the second half of the 20th century, it had its detractors in the biomedical community. These detractors were not those considered on the fringe, the so-called quacks from the long since defeated and marginalized homeopathy community. They were a fledgling group of brilliant geneticists focused on mutations, led by an insightful, but sometimes hard to fathom, academic and later Nobel Prize winner by the name of Hermann Muller (1890–1967). After training with the future Nobel Prize winner Thomas Hunt Morgan (1866–1945) at Columbia University in the 1910s, life took on added intensity for Muller in the mid-late 1920s in his quest to establish that ionizing radiation could cause mutations. Through a series of intriguing experiments at the University of Texas at Austin, Muller uncovered one of the geneticists' holy grails: X-rays could cause mutations in the gonads of male fruit flies [64]. The surprising aspect of Muller's publication was that it contained no data. In fact, the publication was an oddity, amounting to a detailed discussion of unreported data. The article in *Science* aroused debate and confusion, leading former advisor Thomas Hunt Morgan to proclaim, "Now he's done it. He's hung himself" [65]. Lacking methods and data made Muller's *Science* paper suspect. The suspense ended when Muller presented the actual data behind the paper at the Fifth International Genetics Conference in Berlin later that

year and subsequently published in the conference proceedings [66]. It took 19 years for Muller's dream of a Nobel Prize to come true. However, by 1946 the world was frightened by the bomb, its unique capacity for massive devastation, and its potential to cause all sorts of genetic diseases in future generations. Thus, on December 12, 1946 Muller ([67]; http://www.nobelprize.org/nobel_prizes/medicine/laureates/1946/muller-lecture.html) was awarded the Nobel Prize and in his acceptance speech gave further credibility to the field of genetics, while rallying his geneticist colleagues to the belief that only they could save the world from the harmful effects of ionizing radiation [68].

The National Academy of Sciences

The leaders of the genetics community feared that radioactive fallout from atmospheric atomic bomb testing had the potential to threaten the health of future generations of humans, in the United States and elsewhere, by causing mutations in reproductive cells. Working through various national and international committees, they tried to convince the more medically trained committee members that there was no safe dose of ionizing radiation. They argued that radiation would act in a linear fashion and that the risks were inescapable no matter how low the exposure. The geneticists were united in arguing that radiation-induced birth defects would increase significantly because of the atmospheric testing of atomic bombs. While they were on the losing side of a number of key national and international advisory committee recommendations, they won the big one [68]. That is, things changed in 1956 during the deliberations of the BEAR I (Biological Effects of Atomic Radiation) committee of the U.S. National Academy of Sciences. With the committee finally selected in their favor, the geneticists pushed their agenda through with a recommendation that a linearity at low dose assumption be adopted for estimating reproductive risks in humans from exposure to ionizing radiation. Their argument was not based on data that could adequately address this question, far from it, but rather on an unproven dose-response hypothesis set within a context of societal fear and perceived global responsibility. This unified group of geneticists believed that only they had the necessary insights into the mutational effects of radiation; it was their solemn responsibility to protect future generations. They therefore pushed this agenda forward based on a protectionist philosophy, such as today's precautionary principle concept that their ideological offspring promote, even though it lacked the scientific basis to make their case.

Muller had a strong interest in the nature of the dose-response for radiation-induced mutation. Soon after he had observed that X-rays could cause mutations in reproductive cells, he directed studies to determine the shape of the radiation-induced dose-response curve. Muller guided two researchers, Clarence P. Oliver (1898–1991), later to become a professor of genetics at the University of Minnesota and later still at the University of Texas at Austin and Fred B. Hanson (1886–1945), later to become the associate director of the Natural Sciences Division of The Rockefeller Foundation, in his laboratory, to better address whether the mutational dose-response was linear. Similar research was also initiated by a number of other investigators in the immediate aftermath of Muller's seminal findings. However, the results that emerged did not experimentally resolve the issue of the shape of the mutagenicity dose-response [48]. In fact, the lowest dose tested was still strikingly high, being 275 rads, a truly massive dose to the fruit fly's gonads. This dose was comparable to receiving well over

1,000 chest X-rays in 3.5 min! In addition, most of the published attempts to demonstrate linearity during this time period failed to do so, giving further support to the threshold dose–response concept. Despite very limited data, a lack of overall consistent findings, and the fact that low doses were never even remotely assessed, Muller nevertheless inexplicably developed a very firm, although incorrect, public conviction that mutation frequency is directly proportional to the dosage absorbed, with no evidence of a threshold dosage below which the treatment is too dilute to work [65].

Based on the findings of Oliver and Hansen, which supported an X-ray–induced linearity interpretation even though their exposures were grossly excessive, Muller soon showed his inclination to extrapolate X-ray–induced mutation findings in a linear manner. In what may well have been the very first such effort in foreshadowing the future field of risk assessment, he tried to estimate the background spontaneous mutation rate in fruit flies from ionizing radiation using the linearity method. When his predictions were wrong by approximately 1,300-fold [69], Muller was forced to reassess the significance of background radiation, yet his flirtation with the linearity at a low-dose relationship would remain.

According to Carlson [65], Muller displayed this same belief nearly a decade later in his report to the Medical Research Council of Great Britain. In this report Muller [70] suggested that no exposure to ionizing radiation existed below which mutations could not occur. Therefore, regardless of how much the dose may be attenuated as a result of its dilution in environmental media, ionizing radiation posed a mutagenic risk.

That Muller continued to strongly adhere to this public belief in the linearity hypothesis may be seen in his acceptance speech for the Nobel Prize in December of 1946, which affirmed that Oliver, Hanson, and Temofeeff had shown that the frequency of gene mutations is definitely proportional to dose, despite the extremely high cumulative doses and dose rates used. If one doubted his definitive position on linearity, he then cited the research of former student Ray-Chaudhuri [71,72] which, according to Muller [67], leaves “no escape from the conclusion that there is no threshold.” Muller’s unequivocal Nobel Prize Lecture conclusion apparently was not shared by Ernst Caspari. In a letter to Curt Stern, Caspari stated that the difficulty with the Ray-Chaudhuri [73] data involved confusion over the appropriateness of the control group and that the experimental error was quite large. The Ray-Chaudhuri study was of very modest size, failed to include numerous important methodological details, failed to include critical data on lethal clusters, the sterility and fertility of the females, sex ratios, and the age of the males, among others. Of further note is that he changed to a different fruit fly strain halfway through his study without explanation. This new strain had a control group mutation rate of only one third of the previous strain, yet the data of both strains were combined with the author claiming there were no differences between the strains. Despite these and other misgivings, the Ray-Chaudhuri research lowered the dose rate to 0.01 rad/min for a continuous exposure of 43,200 min (30 d). The result was a cumulative dose to the fruit flies of 400 rads, an exposure that was approximately one fifth that which demonstrated approximately five times as much damage. Such findings supported the linearity interpretation. However, the 400-rads/30-day exposure to the flies would exceed human background rates (cosmic radiation and local gamma radiation) by many thousandfold. Less than two years after Muller asserted the no escape phrase, data from Caspari and Stern [74] suggested that

linearity was not observed in the fruit flies at the lowest dose rate yet tested.

That the Caspari and Stern [74] findings directly challenged the linearity assertions was especially interesting because Muller was a paid consultant to Curt Stern on this project, even supplying the fruit flies [68]. In fact, Stern had sent a draft of the Caspari and Stern manuscript to Muller. In a letter to Stern, dated November 12, 1946 (a month before his December 12, 1946 Nobel Prize Lecture), Muller acknowledged the data challenging his linearity perspective, their potential significance, and the urgent need to replicate the study [75]. Despite his knowledge of these data, which were far more substantial, much better documented, and used approximately one sixth the dose rate of the Ray-Chaudhuri experiments, Muller delivered his linearity pronouncement as if it were unassailable; his real message should have been that more study was needed to resolve this issue [76].

A key question then was how the findings of Caspari and Stern [74] could be marginalized without adversely affecting the careers and reputations of these two well-known geneticists. This would also have to be done within a framework that did not expose Muller’s Stockholm deception. This was achieved in a two-step process [77].

The first involved making the discussion of the Caspari and Stern manuscript somehow disavow their findings without finding fault with the data. Caspari and Stern [74], with the encouragement of Muller [75], achieved this goal by arguing that their threshold supporting data could not be accepted until it was determined why their findings differed from that reported in Spencer and Stern [78]. This study also assessed the effects of ionizing radiation on the frequency of sex-linked recessive lethal mutations in the germ cells of fruit flies. However, it was a study with numerous important differences from that of Caspari and Stern [74]. For example, the Spencer study treated the flies with X-rays, not gamma rays, gave their cumulative dose (50 rads) acutely, that is, over only 2 min, whereas the same cumulative dose required constant exposure for 21 d for Caspari. The diets used by the two studies were totally different, markedly affecting the percent sterility and other reproductive parameters. In all, at least 20 significant differences between the studies made them impossible to compare directly, making the demands of their discussion unrealistic and foolish even to propose. Yet Caspari and Stern [74] demanded that the scientific community not accept their findings until they determined why the two studies reported differing mutation rates. They did not apply that constraint to the Spencer and Stern study. In his January 14, 1947 letter to Stern [79], Muller indicated that it would be acceptable to publish the Caspari paper because now so many qualifying statements appeared (i.e., “cautions”); this was most likely because it would not hurt the linearity case [75]. Because Stern was then the editor-in-chief of *Genetics*, their manuscript would get published even with its inappropriate and misdirected discussion.

The next step was to complete the replication study. This chore was given to a new master’s student by the name of Delta Emma Uphoff. The problem was that Uphoff was new to *Drosophila* research, lacking the experience and expertise of Spencer and Caspari, both of whom were exceptionally talented and experienced, being the equal of Stern himself. In her replication of Caspari, the control group was aberrantly low. This resulted in Uphoff and Stern rejecting their findings, saying they were uninterpretable [80]. In fact, in a very unusual course of action, Stern apparently forced Uphoff to note in the discussion that part of the problem may have been bias on behalf of

the “experimenter” (Delta Uphoff, presumably). A second experiment by Uphoff also displayed an aberrantly low control group, again making the findings useless. The third and final experiment seemed to work, because they reported that a dose of radiation that was double that used by Caspari had induced a significant increase in the germ cell mutation rate. This led to the summarized publication of all the findings, including those of Spencer and Caspari in a brief (slightly more than one page) technical note in *Science*. In this paper, Uphoff and Stern [80] concluded that no dose existed below which radiation could not induce a mutation, a major conclusion, in a major journal, by an eminent geneticist, the editor of the most influential genetics journal. This conclusion would carry considerable weight.

Of particular concern is that these authors failed to point out that their significant mutational findings in their third and final experiment gave evidence of being aberrantly high, being nearly threefold greater than would have been predicted even by a linear model. Of considerable importance is that Uphoff and Stern [80] promised to provide the scientific community with the documentation to support their conclusions but they never did. Thus, Uphoff and Stern [80] provided three new experiments, each with aberrant findings and none of the promised documentation. Despite these critical flaws, acceptance of the Uphoff and Stern [80] perspective was rapid and widespread, as reflected in historical perspectives by leaders in the mutagenicity field [81,82]. The conclusions of the Stern research team were especially highlighted in the profoundly influential publication by future Nobel laureate E.B. Lewis [83] when he made his case for ionizing radiation linearity to be extended to cancer induction as well. According to Neel [81], this linearity conclusion even landed Stern a term (1950–1953) on the Advisory Committee to the Division of Biology and Medicine of the Atomic Energy Commission during a critical period in which health policy relating to radiation research was being formulated, all setting the stage for the BEAR I Committee.

Although the genetics community accepted Uphoff and Stern’s [80] undocumented conclusion of linearity for ionizing radiation-induced germ cell mutation in the fruit fly, they were also focused on newly undertaken research at the University of Rochester by Donald Charles and at the Oak Ridge National Laboratory by William Russell with the mouse model. The case of the University of Rochester mouse studies had important problems, mostly centering around Charles, who frustrated his colleagues by providing only unofficial draft assessments that he continued to revise but failed to finalize [68]. The best that Charles did was to publish a 3.5-page partial summary, lacking any presentation of research methods, of his extensive radiation mouse studies in the journal *Radiology* in 1950 [84]. Unfortunately, Charles died of leukemia in 1955, never publishing any further account during the critical lead up to the BEAR I Committee activities. In 1961 some former colleagues attempted to summarize the study results, but only in a very limited fashion [85].

In the case of the mega-mouse type studies of Russell, they too were unable to provide significant scientific insight during the years leading up to the recommendations of the BEAR I committee. During this pre-1956 period, Russell’s experimental studies dealt with high doses, with the lowest dose being 300 rads. In fact, the Atomic Energy Commission Special Ad Hoc Advisory Committee on Genetics had recommended that he lower the dose to 150 and even 75 rads. According to Jolly [68], Russell was determined not to go lower than 150 rads because of

the insensitivity of his model. The committee further stressed the need to push in the direction of lower doses, perhaps with other more sensitive models, to determine whether the nature of the radiation-induced mutation dose–response was linear or threshold. According to Jolly [68], even though the genetics community knew that no convincing direct evidence existed of linearity at low dose for radiation-induced mutation at this time, they nevertheless were committed to the perspective that the mutagenic effects of ionizing radiation were linear, cumulative, and deleterious. This was the genetics community mind-set as led by Muller and his colleagues as the BEAR I committee began their historic deliberations (November 1955 to June 1956). It was also a mind-set that, according to James Crow [86], a member of the BEAR I Committee, was guided by principles that were “mostly from *Drosophila* research.”

What Muller really believed on the dose–response issue may be gleaned from a 1949 letter to Robley Evans (1907–1995), an MIT professor critical of the low-dose linearity hypothesis. Muller stated that “many of the quantities are only very roughly known even for *Drosophila*, and we are admittedly extrapolating, it is all we can do in our present state of ignorance and we must meanwhile remain on the safe side” [65].

While the genetics community was consolidating its belief in linearity at low dose for the radiation-induced mutation concept, dissenters arose strong enough to resist the group perspective. For example, Willard Ralph Singleton (1900–1982) (trained by L. J. Stadler [1896–1954] at the University of Missouri), at the Brookhaven National Laboratory, was an outspoken critic of the linearity at low-dose mutagenicity hypothesis. His research revealed a nonlinear relationship between mutation rate and dose rate, with disproportionate increases in mutations occurring as the dose increased and nonmutagenic responses at lower doses, thereby challenging the belief in linearity at low dose [87–89]. During the deliberation period of the BEAR I committee, *The New York Times* (April 17, 1955) published an article that provided the opportunity to challenge the emerging linearity at low dose consensus for germ cell mutagenicity. In that article Singleton stated: “there is probably a safe level of radiation below which no genetic changes occur.” Jolly [68] noted that even though Singleton was a well-accomplished genetics researcher, his findings and interpretations were, for the most part, ignored because they were in conflict with the emerging and soon-to-be-dominant linearity paradigm.

The significant uncertainty of the nature of the mutagenicity dose–response in the low dose zone and the uncertainty of extrapolating results from fruit flies to humans had little impact on Muller and his geneticist colleagues. Their linear dose–response recommendation received the authority of the National Academy of Sciences, and the United States was on its way to rejecting the threshold dose–response in favor of the linearity at low dose model for the assessment of radiation-induced reproductive damage. Within approximately a year’s time, the focus shifted to somatic effects and the National Committee for Radiation Protection, following the recommendation of E.B. Lewis (1918–2004), one of Muller’s geneticist colleagues and former student of C.P. Oliver, pushed through the first-ever formal recommendation that radiation-induced cancer also be assumed to act via linearity at low doses. Again, the data did not support the case for a cancer linearity argument. In fact, the case that Lewis [83] had made was considered laughable (see Table 2, page 212 of Calabrese [3]), not even requiring a response by opposing leaders in the field [90]. However, which person/idea wins is often determined by who is in power, as seen in the garnering of influential editorial support from Grahame P.

DuShane (1910–1963), the editor-in-chief of *Science* [91], along with major favorable stories in *Life* magazine (June 10, 1957) and other powerful outlets. Within a few years, multiple national and international advisory committees copied the lead of the National Committee for Radiation Protection, and their low-dose linearity recommendation soon became national policy and remains so even today [48].

The action of BEAR I was a major moment in U.S. regulatory history that was quietly achieved, yet with stupendous consequences. It became the official dogma of U.S. regulatory agencies.

The issue of linearity at low dose for radiation-induced cancer was occurring during the later part of the 1950s. Ironically, the Delaney Amendment to the 1958 Food Additives Amendment in the U.S. became law on April 26, 1958 (http://en.wikipedia.org/wiki/Delaney_clause). It stated “no additive shall be deemed to be safe if it is found to induce cancer when ingested by man or animal, or it is found, after tests which are appropriate for the evaluation of the safety of food additives, to induce cancer in man or animal.” This Delaney clause was later inserted into the Color Additives Amendment of 1960, following the Thanksgiving cranberry crisis of 1959 due to the presence of the herbicide aminotriazole, an animal carcinogen, in cranberries [92,93]. Despite their parallelism in time, there was no apparent interaction between the development of a linearity at low dose methodology for radiation-induced cancer and the science and political framework employed to support the decision to prevent adding carcinogens to food.

During this period, James Delaney (1901–1987), a member of the House of Representatives from the state of New York, began to interact with Dr. Wilhelm Hueper, a National Cancer Institute scientist, and a leading expert on environmental and industrial carcinogens. Hueper offered a very strong protectionist philosophy to Delaney along with powerful credentials, thereby allowing Delaney to proceed. Because scientists were unable to define what a safe level of exposure to carcinogens may be, along with not understanding their mechanisms of action, Delaney asserted that no risk was worth taking with respect to chemical carcinogens, and that chemicals did not have rights.

In the case of radiation, a different concept of risk evolved that related to permissible risk that could be estimated with the linear model. The Delaney amendment, inspired by the strong views of Hueper, was to lead to the prevention of possible exposures. The Food and Drug Administration would later modify the Delaney amendment to address the concept of a de minimis risk, so that carcinogens could be added to the food supply if they were estimated to have a risk less than a certain value (for example, one in a million/lifetime), following a linearity at low dose model. Thus, in time, the radiation and food additive risk perspectives converged. Committee 17 of the Environmental Mutagen Society attempted to have the Delaney Amendment generalized to include chemical mutagens in the early 1970s, but failed to achieve this goal, falling back to the earlier guidance of the 1956 BEAR I committee that assessed genetic risks within the context of a doubling dose framework that was still consistent with the linearity at low dose model [94,95].

National Academy of Sciences—Safe Drinking Water Committee

Nearly 20 years later, the first National Academy of Sciences (NAS) Safe Drinking Water Committee (SDWC) [96] adopted the linear at low dose risk assessment compromise of the late 1950s, somewhat updated, and applied it to chemical carcino-

gens. Their actions make as little sense today as they should have in 1977. The NAS SDWC failed to provide an adequate evaluation and set of recommendations on this risk assessment issue. In their document, eight guiding principles for the support of low-dose linearity for cancer risk assessment were discerned. Within approximately two decades, six were shown to be untenable, another impossible to study practically, and the eighth had yet to be demonstrated [48] (see Table 4, page 217 in Calabrese [48]) (Appendix 2). Their highly precautionary oriented recommendation was presented to the U.S. EPA in the 1977 publication entitled *Drinking Water and Health*. Their eight principles (ie, assumptions) had emerged from a new generation of unified geneticists of the early 1970s as summarized in published comments during a high-level genotoxicity conference chaired by Alexander Hollaender (1898–1986) with the proceedings in *Environmental Health Perspectives* (see the detailed set of comments after the Ernst Freese (1925–1990) paper [97]; these comments present the views of multiple genetic toxicologists concerning mutation and dose–response).

Further confounding this faux pas was another apparent miss by the NAS SDWC. While the Committee was drafting their rubber stamp-like statement on chemical carcinogens and linearity at low doses, other researchers had published a major new finding that could have redirected the committee on the dose–response issue for carcinogens. A March 1977 paper in *Nature* by Samson and Cairns [98] revealed for the first time that a low dose of a chemical mutagen induced an adaptive response that led to protection against a subsequent and more massive exposure to that same mutagen. The paper ushered in the field of adaptive response and its widespread generality. If this paper had been read by the committee or its staff, it might have changed the course of cancer risk assessment for chemical carcinogens. However, the *Nature* paper was never cited in *Drinking Water and Health*.

Not surprisingly, the U.S. EPA accepted the linearity at low dose recommendation of the Committee and applied it to trihalomethanes (chloroform and related agents) by 1979 and then to a subsequent very long list of other chemical carcinogens for the rest of the century to the present. Life in the world of cancer risk assessment has not been the same since. Thus, even though a plethora of new studies have been published challenging the linearity at low dose perspective for cancer, the die was cast based upon an anemic, at best, assessment of the NAS SDWC, an assessment that missed one of the more significant and relevant new findings.

The NAS brings together some of the best and the most experienced professionals in a no conflict of interest manner in which biases are attempted to be balanced. This is the way it is described in writing, giving assurances of objectivity and scientific integrity to Congress, the scientific community, the media, and the public. However, two NAS committees failed on the most critical questions of the past half century. The first failure (the BEAR I committee) came from the geneticist community that introduced ideology into risk assessment. What this committee achieved was a dramatic failure of process, demonstrating that its ends justified the means, setting an unacceptable precedent. In the case of the next generation of NAS experts, the Safe Drinking Water Committee simply became enveloped by the ideologically oriented perspectives of their geneticist and biostatistical colleagues. It was a committee that accepted a series of assumptions without proposing how to validate the dose–response model they promoted into long-standing regulatory influence.

All we are saying... is give validation a chance

While the threshold model was being restricted to the assessment of only noncancer endpoints, the FDA decided to determine the nature of the dose–response in the low dose zone for genotoxic carcinogens. To achieve this goal, the FDA undertook the largest rodent study ever, using some 24,000 female mice (BALB/c strain), the so-called mega-mouse/ED01 study. A single carcinogen was tested, 2 acetylaminofluorene (2-AAF), known to be a mutagen and to cause tumors at multiple sites in different animal models, but especially in the bladder and liver of females of this mouse strain.

It was thought that much would hinge on the findings of this study, including whether the United States would base its carcinogen risk assessment methods on using a linear at low dose model or an alternative. So substantial were the scientific and societal implications that the U.S. Society of Toxicology created a 14-member expert panel to provide an assessment. Their analysis led to the publication of nearly an entire issue of the Society's journal, *Fundamental and Applied Toxicology* [99]. What they found surprised everyone. The dose–time–response was strikingly hormetic for the bladder cancer endpoint, occurring in each of the six different rooms housing the animals. In effect all six replications of the study agreed. The formal writeup by the expert Society of Toxicology panel strongly emphasized the J-shaped dose–response with beneficial effects at low doses as seen in their quoted comments from page 77: "The most striking aspect... is the reduction in probability of bladder cancer from control to doses 30, 35, and 45 ppm. This reduction occurs in all six rooms and is statistically significant... the ED01 study provides more than evidence of a 'threshold.' It provides statistically significant evidence how low doses of a carcinogen are beneficial" [99]. Thus, in the largest rodent cancer study ever undertaken, the data revealed a hormetic response. Despite these striking findings, the U.S. regulatory agencies failed to modify their approaches to carcinogen risk assessment policy and practice.

The foray of U.S. regulatory agencies into dose–response model validation for carcinogen responses using a standard rodent experimental model did not confirm the linearity at low dose hypothesis for bladder cancer nor even the threshold dose–response model. Huge amounts of money were spent, expectations were high, and in the end the federal agencies would not follow the data. The process was expensive, prolonged, and traumatic, all factors that would probably prevent any similar bureaucratic risk taking in the foreseeable future.

Failure of the threshold dose–response model

In the course of developing a methodology to assess the possible validation or limitations of the hormesis model, the question arose as to whether the threshold dose–response model had ever been validated. The general assumption was that it must have been validated, because it was now nearly 70 years since this model had been accepted and integrated into the lexicon of mainstream pharmacology and toxicology. Search we did, using every conceivable database and spectrum of relevant search terms and their combinations, along with the assistance of science librarians trained to uncover difficult-to-find entities. We simply could not find any attempt that had ever been published to assess the capacity of the threshold dose–response to make accurate predictions in the low dose zone, that is, below the threshold. Having reached the proverbial dead-end, we undertook our own attempt to validate the capacity of

the threshold dose–response model to make accurate predictions in the low dose zone.

The validation study of the threshold dose–response was initially undertaken from a data set created from the pharmacological and toxicological literature using rigorous a priori entry and evaluative criteria. These criteria were applied to all the published studies in three journals (*Environmental Pollution*, *Bulletin of Environmental Contamination and Toxicology*, and *Life Sciences*) from the time of their inception in the mid 1960s to the present. The threshold dose–response model predicts that responses below the threshold should vary or bounce in a quantitatively similar fashion on either side of the control, just like random noise in a system. If the threshold model is dominant, then the ratio of responses above and below the control value should be very close to 1. The surprise was that the ratio did not approach 1; it exceeded this value by approximately 250%, a frequency that is far beyond any reasonable probability. The threshold model was not able to account for the findings in the below threshold zone. However, the responses below the threshold did display a consistent pattern, very closely paralleling the hormesis model [100,101].

This study had two major conclusions. The first indicated that the threshold dose–response model failed to make accurate predictions in the below threshold zone. The data set was intentionally very general, including data from plants, microbes, invertebrates and vertebrates, and from a wide range of biological endpoints and chemical agents, thereby enhancing the significance of the findings. Second, the validation study strongly supported an hormetic interpretation, findings that were consistent with the thousands of hormetic dose–responses that had been previously assessed based on a priori evaluative criteria.

This new study was especially important because the rules of the game applied equally to the threshold and hormetic models. The hormetic model could no longer be ignored. These were findings as important as they were unexpected.

Publicity follows hormesis

The validation data were important because they challenged a 70-year-old dose–response tradition. Our manuscript [100] was submitted to the journal *Toxicological Sciences*, the main journal of the U.S. Society of Toxicology, and made it successfully through their typically thorough peer-review process. Soon the toxicological world would learn that the threshold model was not as good as they had long been taught, whereas the hormetic model had performed far better than they might have imagined. In fact, we thought that this paper was one that had the potential to alter the toxicological landscape.

At approximately the time that *Toxicological Sciences* accepted our manuscript, I received a letter from the editor-in-chief of *Nature* with an invitation to write an article on hormesis. This publication in *Nature* provided a significant boost for the hormesis concept [102]. The key decision was its placement in the journal's media package, a position of high visibility. Soon we were inundated with calls for interviews from leading publications and other media outlets all over the world. Articles on hormesis quickly appeared in *The Wall Street Journal* [103], *Forbes* [104], *Fortune* [105], *Discover* [106], *Scientific American* [107], *Science News* [108], *Insight* [109], *Reason Online* [110], *U.S. News and World Report* [111], and in major stories in large daily newspapers like *The London Times Online* [112], *St. Louis Post-Dispatch* [113], *Boston Globe* [114], *The Baltimore Sun* [115], and others. Of further note

was that *Science* [116] published a four-page story on hormesis. Hormesis had finally arrived and with a big splash.

As a result of the *Toxicological Sciences* and *Nature* papers, interest in the hormesis concept was on a marked upswing. Round two of the validation tests would also take place in the journal *Toxicological Sciences*, using a large National Cancer Institute public database (57,000 dose–responses) that assessed the effects of nearly 2,200 potential anti-tumor drugs on 13 different strains of yeast, 12 of which had a different genetic error similar to some found in human cancers. Following the same basic plan as before, the dose–responses were put through a rigorous set of a priori entry and evaluative criteria and multiple statistical evaluations [117]. Regardless of the statistical analysis strategy, the findings once again strongly supported a hormetic interpretation whereas the threshold dose–response model performed extremely poorly, in effect, failing the test. Round three in the validation series was a study using yet a different approach; this time an *Escherichia coli* strain was tested in over 2,100 different potential antibiotics within an experimental replication framework. The results were the same, good job of predicting below threshold results for the hormetic dose–response model, but a poor job for the threshold model [118].

The threshold dose–response had now failed a third major challenge, strongly undercutting the scientific status of the U.S. EPA and FDA's default model used to establish most health standards. How many times would the threshold model have to be shown to be inadequate before regulatory agencies would reconsider their continuing acceptance of it as their gold standard, that is, the default model?

Linearity model performance: Noncancer endpoints

Of potential importance in the head-to-head comparisons between the threshold and hormetic dose–response models was that the predictive capacity of the linear at low dose model was also tested. It was not the main focus of the studies, because the endpoints were noncancer. However, in each of the extensive validation studies, the linear at low dose model was a failure, just like the threshold dose–response model. These observations challenge the general predictive utility of a linear at low dose model. Despite these findings, a recent NAS committee has proposed [119] generalizing the linearity at low dose model to all endpoints. Although this position was principally hypothetical, our data demonstrate that it fails to predict accurately in the low dose zone, even more than the threshold model.

Big pharma-enhancing biological performance and hormesis

In contrast to governmental regulatory agencies, which can be affected by ideological perspectives, businesses follow data that lead to profit. The early decades of the 20th century witnessed pharmaceutical companies racing to discover agents that would destroy major disease-causing microbes. The late 1940s likewise revealed the birth of cancer chemotherapy with its massive expansion in the following decades. Although killing cancer cells and harmful organisms has been a major pharmaceutical preoccupation, the 1970s ushered in a new initiative for this industry, one that concerns enhancing biological performance, all kinds of performance. These include, but are not limited to, improving memory, strengthening bone, enhancing sexual performance, growing more hair, faster and stronger wound healing, reducing anxiety, and reducing the risks of seizures. In each of these cases, the increased performance was attributable to the hormetic dose–response, all with copious supportive pharmacological studies [120]. The phar-

maceutical industry and their regulatory oversight agency, the FDA, are at the core of these discoveries and their implementation within society. Not once, however, has the industry or the FDA credited or linked such successes to hormesis.

Pharmaceutically oriented scientists have called these performance-enhancing drug-induced dose–responses by a wide range of names, including biphasic, diphasic, parabolic, bitonic, bell-shaped, U-shaped, J-shaped, inverted U, low dose stimulation, pre-conditioning and several others, but not hormesis. In this process, the research community has failed to recognize that these biphasic dose–responses, which affect so many different types of biological endpoints, display the same quantitative features. These responses are not a haphazard grouping of dose–response entities that exhibit such remarkable similarity by chance. In fact, the quantitative features of the hormetic dose–response are the same regardless of the biological system studied, whether at the cell, organ, or individual level, the endpoint measured, or the chemical inducing the effect. It displays remarkable generality. Of considerable potential importance is that the hormetic dose–response likely provides a quantitative index of the limits of biological plasticity for each of these drug-induced performance enhancing effects [121]. In so doing, the hormesis concept reveals the magnitude of drug-induced responses that pharmaceutical companies can expect in human populations, and whether developing their product further would be profitable. This knowledge can be a key determinant in designing preclinical and clinical studies and could have a major impact on the assessment of clinical efficacy. The industry has unfortunately failed to adequately appreciate that hormesis is a broadly integrative and central biological principle that can revolutionize the drug development process. Nonetheless, in its own way, this industry has embraced the hormesis concept, in principle, in practice, and in fact. They have yet to embrace it by name. The failure of the pharmaceutical industry to use the term *hormesis* reflects its origin in traditional medicine and the long-standing conflict with homeopathy. This conflict has now come full circle. The dose–response explanatory principle of Schulz, so strongly rejected by the medical community nearly a century ago, underlies much of the success of the modern pharmaceutical industry in an ironic twist of scientific fate and promises to be even more significant in the future.

Final perspectives

The nature of the dose–response and its underlying mechanisms will remain toxicology's *raison d'être*. The central issue of toxicology has quickly transformed into that of a low dose paradigm to reflect the societal concerns in which most people live. The capacity to investigate low doses has been revolutionized with respect to profound advances in chemical analysis, which has been directly linked to experimental systems for *in vitro* studies in which large numbers of concentrations of chemicals can be studied. In fact, the resurgence of interest in hormesis is being driven by such technical improvements, with a focus on assessing the biological effects of chemical and physical agents at low doses. Although high dose toxicology is not yet a historical remnant, and may never fully be so, the present and future of toxicology are in the low dose domain. This powerful development will drive the field for the foreseeable future and places hormesis directly at the forefront.

During the entire decade of the 1980s, only 10 to 15 citations per year could be found of the terms *hormesis* or *hormetic* in the vast Web of Science database. In 2010 alone, the number of citations exceeded 3,200, a sign of growing acceptance and

progressive integration within the scientific community and its research foundations. Its success and influence emerge from its broad generality across biological models, endpoints measured, and chemical and physical stressor agents along with its potential biomedical significance and reproducibility.

The hormesis concept has become integrated into a growing number of highly influential textbooks [122–124], and its usage by biomedical scientists has expanded worldwide. Five monographs on hormesis have also been published within the past few years [125–129]. Hormesis has also become a central concept in the areas of aging and biogerontology as well as becoming a foundation for pharmaceutical agents designed to improve biological performance, especially in the areas of anxiety reduction, memory enhancement, stroke damage prevention, bone strengthening, wound healing, skin care, the numerous domains of pre- and post-conditioning and other areas. These developments point to expansive growth, enhanced clinical significance, and biomedical centrality.

While this brief summary strongly suggests that the future will be hormesis-oriented and central to the biomedical community, several identifiable institutional factors preserve the dose–response status quo. Each is regulatory-agency oriented and reflects specific manifestations of the historical dose–response strategy of traditional medicine during the mid decades of the 20th century, as detailed earlier. These dose–response impediments include the following: research funding by governmental regulatory agencies is likely to ignore hormesis, the fear that hormesis will weaken environmental exposure standards, the default model in risk assessment is a vehicle for conservative risk estimates, and the U.S. EPA's definition of a risk assessment denies the capacity to incorporate health benefits.

Research funding by governmental regulatory agencies

Federal funding agencies control the direction of much research in the United States and elsewhere, including who and what gets funded, as well as the language and culture of research. This process directly affects what ideas and data get published, read, and believed. Broadly diversified funding sources both within and outside of government are an important means to broaden research goals and directions as well as to enhance the likelihood that valid research ideas, including those dealing with hormetic hypotheses, are not minimized or excluded because of historical, structural, or ideological biases. For example, because the U.S. EPA definition of a risk assessment explicitly excludes the concept of a beneficial or adaptive response, it suggests that the Agency would fail to prioritize and therefore be less likely or even fail to fund hormetic hypotheses relating to such beneficial responses (See *Responses excluded in U.S. EPA risk assessment*). A certain proportion of regulatory agency research funding should be made independent of regulatory agency control via the use of external panels to enhance the objectivity of the granting process, from the development of the research priorities to the awarding of the specific grants.

Fear that hormesis will weaken exposure standards

Although the hormetic stimulatory response has the capacity to induce both beneficial and adverse health effects depending on the specific biological context [130], various authors have incorrectly truncated this definition to only include a beneficial effect. This has led to the further position that the hormesis concept will undercut many of the environmental gains that have been made over the past four decades, eventually resulting

in weakened environmental health standards. In fact, the opposite is likely to be the case. An hormesis-guided risk assessment process provides decision-makers with the most complete dose–response information on the biological/toxicological effects of the agent tested, especially in the low dose zone. By long ignoring the hormetic/biphasic dose–response, the biomedical community and regulatory agencies failed to discern the occurrence of endocrine disruption effects at doses below the traditional toxicological threshold [120]. In a similar fashion, oncologists have also long missed the capacity of numerous anti-tumor drugs to enhance the proliferation of tumor cells in patients because of their longstanding belief in threshold dose–response model predictions [131]. Thus, by incorporating the hormesis concept into the risk assessment framework, including the design of the bioassay, risk estimates would be more confidently based regardless of whether the hormetic hypothesis was supported by the data, and whether low-dose stimulatory responses were harmful or beneficial.

The default model in risk assessment

An important issue for regulatory agencies such as the U.S. EPA is what dose–response model is selected as the default in risk assessment. A default dose–response model is usually selected when insufficient data are available in the standard bioassay to convincingly identify the best-fitting dose–response model. This situation occurs nearly all the time in practice. In these cases, the agency will routinely default to the most conservative model. This situation creates a self-fulfilling prophecy of model selection because the poorly designed standard chronic bioassay (that is, too few and too high doses) does not permit one to distinguish in a statistical manner amongst the dose–response models. As noted, several large-scale dose–response validation studies have indicated that the threshold and linear at low dose models do not make reliably accurate predictions in the low dose zone, whereas the hormetic model does. Also, in many thousands of other studies the dose–responses fail to reflect threshold or linear responses while conforming to the hormetic model. Yet these poorly performing dose–response models are guaranteed to win the default model contest for regulated chemicals because of the design limitations of the standard chronic bioassay. The U.S. EPA therefore uses a chronic bioassay that inevitably leads to the selection of a default model that fails in validation studies! Therein lies the self-fulfilling highly conservative risk assessment prophecy of the status quo, which regulatory agencies have failed to correct.

Responses excluded in U.S. EPA risk assessment

The U.S. EPA [132] risk assessment goal is to prevent pollutant-induced harm while not considering possible health benefits (i.e. “as the purpose of a risk assessment is to identify risk [harm, adverse effects, etc.], effects that appear to be adaptive, non-adverse or beneficial may not be mentioned”). This goal creates a framework in which the hormetic dose–response could be ignored. For example, this suggests that the EPA would consider possible harm related to an hormetic/biphasic dose–response, but not if benefits occurred. The U.S. EPA could accept data showing a low dose hormetic stimulation leading to adverse health effects (e.g., increased prostate gland size [120] or significant acceleration in a developmental process such as the onset of puberty [133,134]). In contrast, the U.S. EPA would ignore the hormetic response when it resulted in a reduction in a population-based risk (reduction in tumor incidence).

This discussion illustrates that the U.S. EPA concept of risk is too limited, because two types of risks are present (risk of harm and risk of losing a benefit). Both need to be considered when providing an integrated public health assessment. This position was strongly supported in a recent survey of the membership of the U.S. Society of Toxicology and the Society of Risk Analysis in which 68% advocated for the incorporation of health benefits into the risk assessment process [135].

The U.S. EPA risk assessment policy statement is also inconsistent with community-based programs for fluoride. Fluoride risk assessments have historically centered on preventing harm at high doses while being flexible enough to ensure the existence of community-based drinking water fluoridation programs at lower doses, that is, a beneficial response.

Another conflict of this policy occurs when an agent displays a beneficial effect for one segment of the population at a dose that would be harmful to another population-based subgroup (for example, a high-risk group). In similar situations, the same agent may display a beneficial effect with the high-risk group at a low dose, while having no measurable biological effect on the normal segment of the population at this dose. This general set of conditions/possibilities would be expected to be a common occurrence.

The above series of policy-based inconsistencies indicate that the U.S. EPA risk assessment policy guidance document is problematic by failing to consider interindividual variability (different stakeholder groups) in a hormesis-based risk assessment process. Therefore, the benefits and risks to each of these groups would have to be considered, quantified, made explicit in the assessment process, and then integrated within a comprehensive negotiation or risk management decision. Although this more comprehensive methodology presents new challenges for agencies such as the U.S. EPA and individual states, it also reflects emerging biological realities that must be dealt with and managed, all with the goal of estimating an optimized population-based response.

The 2004 U.S. EPA risk assessment guidance document [132] fails to integrate the health assessment needs of society. Its limited definition of a risk assessment creates a significant gap in the assessment of human health, leaving an institutional blind spot. This flawed U.S. EPA strategy will lead to inadequate population-based health assessments and wasteful allocation of resources. I suggest a revision to the U.S. EPA definition of a risk assessment to one that estimates the net population-based toxicity incidence at each level of incremental exposure.

How will this dose–response debate turn out? The area of the biological effects of low dose exposures will proceed at an expanding pace independent of the regulatory agencies. This will have a transforming impact on understanding of the nature of the dose–response in the low dose zone and its underlying mechanisms as well as a plethora of new public health and biomedical implications. The real issue is to what extent regulatory agencies will embrace such research and incorporate its findings within their risk assessment paradigms. Unless such changes occur within regulatory agencies, society might see an interesting, though troubling, dichotomy in which the world of pharmaceuticals and health care products become based to an ever greater extent on the hormesis concept, improving the quality of our lives, while environmental regulatory agencies cling to an anachronistic belief/policy that their mission is only to prevent harm and not to optimize public health.

Acknowledgement—This effort was sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number FA9550-07-1-0248. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsement, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government. The detailed comments of the multiple peer-reviewers are greatly appreciated.

REFERENCES

- Calabrese EJ. 2008. Hormesis: Why it is important to toxicology and toxicologists. *Environ Toxicol Chem* 27:1451–1474.
- Calabrese EJ. 2005. Historical blunders: How toxicology got the dose-response relationship half right. *Cell Mol Biol* 51:643–654.
- Calabrese EJ. 2010. Hormesis is central to toxicology, pharmacology and risk assessment. *Hum Exper Toxicol* 29:249–261.
- Hahnemann S. 1962. *The Organon of Medicine*, 6th Ed. Translated with Preface by William Boericke, and Introduction by James Krauss. Notes to Sections 88–90. Roysingh and Company, Calcutta, India.
- Coulter HL. 1982. *BT Divided Legacy: The Conflict Between Homoeopathy and the American Medical Association*. North Atlantic Books, Richmond, CA, USA.
- Coulter HL. 1972. *Homoeopathic Medicine*. Formur, St. Louis, MO, USA.
- Nicholls PA. 1988. *Homeopathy and the Medical Profession*. Croom Helm, New York, NY, USA.
- Southam CM, Ehrlich J. 1943. Effects of extracts of western red-cedar heartwood on certain wood-decaying fungi in culture. *Phytopathology* 33:517–524.
- Schulz H. 1887. Zur Lehre von der Arzneiwirkung. *Virchows Arch Pathol Anat Physiol Klin Med* 108:423–445.
- Schulz H. 1888. Über Hefegifte. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere* 42:517–541.
- Crump T. 2003. Contemporary medicine as presented by its practitioners themselves. *Leipzig*. 1923:217–250, Hugo Schulz. NIH Library Translation (NIH-98-134). *Nonlinear Biol Toxicol Med* 1:295–318.
- Bohme H. 1986. Hugo Schulz (8/6/1853-7/13/1932). His Life and Work, PhD thesis, Freien University of Berlin. Berlin, Germany.
- Bloedau CV. 1884. 1884. *General Medical Central* 93:1362.
- Schulz H. 1885. About the treatment of cholera with veratrine. *German Medical Weekly Paper* 11:99.
- Martius-Rostock F. The Arndt-Schulz Axiom. *The Munich Medical Weekly News*, August 2, 1923. 70:1–4.
- Goerig M, Agarwal K, Schulte am Esch J. 2000. The versatile August Bier (1861–1949), father of spinal anesthesia. *J Clin Anesthesia* 12:561–569.
- Clark AJ. 1927. *Applied Pharmacology*. P. Blakiston's Sons, Philadelphia, PA, USA.
- Clark AJ. 1933. *The Mode of Action of Drugs on Cells*. Arnold, London, UK.
- Clark AJ. 1937. General pharmacology. In Heffner AJ, *Handbuch der Experimentellen Pharmakologie*, 4. J. Springer, Berlin, Germany.
- Clark AJ. 1938. *Patent Medicines*. FACT Series 14. London, UK, p. 88.
- Verney EB, Barcroft J. 1941. A.J. Clark (Obituary). *Obituary Notices of Fellows of the Royal Society* 3:969.
- Gaddum JH. 1962. The pharmacologists of Edinburgh. *Ann Rev* 2:1–10.
- Robinson GA. 1981. Forward. In Lambie JW, ed, *Towards Understanding Receptors*. Elsevier, New York, NY, USA, p. 234.
- Calabrese EJ 2009. Getting the dose-response wrong Why hormesis became marginalized and the threshold model accepted *Arch Toxicol* 83:227–247.
- Dannenberg H., 1930. Yeast fermentation: On the question of the validity of the Arndt-Schulz rule. *Naunyn-Schmiedeberg's Archiv für Experimentelle Pathologie und Pharmakologie* 154:211–221.
- Branham SE. 1929. The effects of certain chemical compounds upon the course of gas production by Baker's yeast. *J Bacteriol* 18:247–264.
- Calabrese EJ, Baldwin LA. 2000. Chemical hormesis: Its historical foundations as a biological hypothesis. *Hum Exp Toxicol* 19:2–31.
- Calabrese EJ, Baldwin LA. 2000. The marginalization of hormesis. *Hum Exp Toxicol* 19:32–40.
- Calabrese EJ, Baldwin LA. 2000. Radiation hormesis: Its historical foundations as a biological hypothesis. *Hum Exp Toxicol* 19: 41–75.
- Calabrese EJ, Baldwin LA. 2000. Radiation hormesis: The demise of a legitimate hypothesis. *Hum Exp Toxicol* 19:75–85.

31. Calabrese EJ, Baldwin LA. 2000. Tales of two similar hypotheses: The rise and fall of chemical and radiation hormesis. *Hum Exp Toxicol* 19:85–97.
32. Bernard C. 1877. *Lecons sur le Diabete et la Glycogenese Animale*. Librairie, J.B. Bailliere et Fils, Paris, France, p 132.
33. Aitken RS. 1930. On the renal threshold for chloride in man. *J Physiol Lond* 67:199–210.
34. Ambard L. 1920. *Physiologie normale et pathologique des reins*. 2nd ed. Masson, Paris, France, p. 145.
35. Ambard L, Weill A. 1912. Les lois numeriques de la secretion renale de l'uree et der chlorure de sodium. *J Gen Physiol Pathol* 14:753–765.
36. Haden RL. 1925. The relation of the chloride metabolism to the toxemia of lobar pneumonia: Preliminary report. *J Lab Clin* 10:0337–0342.
37. Haldane JS, Priestley JG. 1916. The regulation of excretion of water by the kidneys. *J Physiol Lond* 50:296–303.
38. Magnus R. 1900. On diuresis. II Announcement: Comparison of the diuretic effectiveness of isotonic saline solutions. *Arch Exp Pathol Pharmacol* 44:396–433.
39. Cushny AR. 1926. *The Secretion of the Urine*, Chapter II. Longmans, London, UK.
40. Clark AJ. 1926. The reaction between acetyl choline and muscle cells. *J Physiol* 61:530–546.
41. Clark DH. 1985. *Alfred Joseph Clark (1885–1941): A memoir*. C & J Clark Archives, Glastonbury, UK.
42. Maehle A-H. 2009. A binding question: the evolution of the receptor concept. *Endeavour* 33:134–139.
43. Shackell LF, Williamson W, Deitchman MM, Katzman GM, Kleinman BS. (1924–1925). The relation of dosage to effect. *J Pharmacol Exp Ther* 23/24:53–56.
44. Shackell LF. 1925. The relation of dosage to effect. II. *J Pharmacol Exp Ther* 25:275–288.
45. Shackell LF. 1923. Studies in protoplasm poisoning. I. Phenols. *J Gen Physiol* 5:783–805.
46. Mutscheller A. 1925. Physical standards of protection against roentgen ray dangers. *Am J Roentgenol Rad Ther* 13:65–70.
47. Mutscheller A. 1928. Safety standards of protection against X-ray danger. *Radiology* 10:468–476.
48. Calabrese EJ. 2009. The road to linearity: Why linearity at low doses became the basis for carcinogen risk assessment. *Arch Toxicol* 83:203–225.
49. Francis T Jr. 1939. Quantitative relationships between the immunizing dose of epidemic influenza virus and the resultant immunity. *J Exp Med* 69:283–300.
50. Dinman DB. 1972. Non-concept of no threshold chemicals in environment. *Science* 4021:495–497.
51. Clark AJ. 1927. The historical aspect of quackery. *Br Med J* 1:589–590.
52. Helmstadter A. 2007. Is there a tonic in the toxin? The Arndt-Schulz law as an explanation for non-linear dose-response relationships. Workshop on *Precarious Matters*, Max Planck Institute for the History of Sciences, Berlin, Germany, March 22–24, pp 29–38.
53. Wels P. 1933. The life time work of Hugo Schulz. *Naumyn-Schmiedebergs Archiv fur Experimentelle Pathologie und Pharmakologie* 170:744–757.
54. Bliss CI. 1935. The calculation of the dosage-mortality curve. *Ann Appl Biol* 22:134–167.
55. Gaddum JH., 1933. Methods of biological assay depending on the quantal response. Medical Research Council. Sp Rep Ser 183. Her Majesty's Stationery Office London, UK
56. Servos JW. 1996. *Physical Chemistry from Ostwald to Pauling: The Making of a Science in America*. Princeton University Press, Princeton, NJ, USA.
57. Bliss CI, Cattell M. 1943. Biological assay. *Annu Rev Physiol* 5:479–539.
58. Bliss CI. 1935. Estimating the dosage-mortality curve. *J Econ Entomol* 25:646–647.
59. Bliss CI. 1940. The relation between exposure time, concentration and toxicity in experiments on insecticides. *Ann Entomol Soc Am* 33:721–766.
60. Bliss CI. 1941. Biometry in the service of biological assay. *Ind Eng Chem* 13:84–88.
61. Bliss CI. 1956. The calculation of microbial assays. *Bact Rev* 20:243–258.
62. Bliss CI, Packard C. 1941. Stability of the standard dosage-effect curve for radiation. *Am J Roentgenol Rad Ther* 46:400–404.
63. Bryan WR, Shimkin MD. 1943. Quantitative analysis of dose-response data obtained with three carcinogenic hydrocarbons in strain C3H male mice. *J Natl Cancer Inst* 3:503–531.
64. Muller HJ. 1927. Artificial transmutation of the gene. *Science* 66:84–87.
65. Carlson EA. 1981. *Genes, Radiation and Society: The Life and Work of H.J. Muller*. Cornell University Press, Ithaca, NY, USA.
66. Muller HJ. 1928. The problem of genic modification. Verhandlungen des V. Internationalen Kongresses fur Vererbungswissenschaft (Berlin, 1927) in Zeitschrift fur inductive abstammungs- und Vererbungslehre. *Suppl Band* 1:234–260.
67. Muller HJ. 1946. *The production of mutations. Nobel Lectures*. Elsevier, Amsterdam.
68. Jolly C. 2003. Thresholds of Uncertainty: Radiation and Responsibility in the Fallout Controversy. PhD thesis. Oregon State University, Corvallis, OR, USA.
69. Muller HJ, Mott-Smith L. 1930. Evidence that natural radioactivity is inadequate to explain the frequency of “natural” mutations. *PNAS* 16:277–285.
70. Muller HJ. 1939. Report of investigation with radium. *Medical Research Council. Sp Rep Ser* 236. London, UK.
71. Ray-Chaudhuri SP. 1939. The validity of the Bunsen-Roscoe law in the production of mutations by radiation of extremely low intensity. *Proceedings, 7th International Congress on Genetics*, Edinburgh, Scotland, UK, 246.
72. Ray-Chaudhuri SP. 1944. The validity of the Bunsen-roscoe law in the production of mutations by radiation of extremely low intensity. *Proc R Soc Edinb* 62:66–672.
73. American Philosophical Society. 1947. Caspari papers, Stern file, November 7. Philadelphia, PA.
74. Caspari E, Stern C. 1948. The influence of chronic irradiation with gamma-rays at low dosages on the mutation rate in *Drosophila melanogaster*. *Genetics* 33:75–95.
75. American Philosophical Society. 1946. Stern papers, Muller file, November 12. Philadelphia, PA.
76. Calabrese EJ. 2011. Muller's Nobel lecture on dose-response for ionizing radiation: Ideology or sciences? *Arch Toxicol*, DOI 10.1007/s00204-011-0728-8 (in press).
77. Calabrese EJ. 2011. Key studies to support cancer risk assessment questioned. *Environ Mol Mutagen* 52:595–606.
78. Spencer WP, Stern C. 1948. Experiments to test the validity of the linear R-dose mutations frequency relation in *Drosophila* at low dosage. *Genetics* 33:43–74.
79. Muller HJ. 1947. Letter to Stern, American Philosophical Society, Curt Stern papers – H.J. Muller file. American Philosophical Society, Philadelphia, PA.
80. Uphoff DE, Stern C. 1949. The genetic effects of low intensity irradiation. *Science* 109:609–610.
81. Neel JV., 1987. Curt Stern (1902–1981). *Biographical Memoire*. National Academy of Sciences, NAS Press, Washington, DC, USA pp 443–447.
82. Lipshitz HD. 2005. From fruit flies to fallout: Ed Lewis and his science. *Dev Dynam* 232:529–546.
83. Lewis EB. 1957. Leukemia and ionizing radiation. *Science* 125:965–972.
84. Charles DR. 1950. Radiation induced mutations in mammals. *Radiology* 55:579–581.
85. Charles DR, Tihen JA, Otis EM, Grobman A. 1961. Genetic effects of chronic X[-irradiation exposure in mice. *Genetics* 46:1.
86. Crow JF. 1995. Quarreling geneticists and a diplomat. *Genetics* 140:421–426.
87. Singleton WR. 1954. The effect of chronic gamma radiation on endosperm mutations in maize. *Genetics* 39:587–603.
88. Singleton WR. 1954. Radiation effects on living systems. *J Hered* 45:58–64.
89. Richter A, Singleton WR. 1955. The effect of chronic gamma radiation on the production of somatic mutations in carnations. *PNAS* 41:295–300.
90. Mole RH. 1958. The dose-response relationship in radiation carcinogenesis. *Br Med Bull* 14:184–189.
91. DuShane G. 1957. Loaded dice. *Science* 125:964.
92. White SR. 1994. Chemistry and controversy: regulating the use of chemicals in foods: 1883–1959. PhD thesis. Emory University, Atlanta, GA, USA.
93. Henahan JF. 1977. Whatever happened to the Cranberry Crisis. *The Atlantic Monthly* 239:26–36.

94. Drake JW, Abrahamson S, Crow JF, Hollaender A, Lederberg S, Legator MS, Neel JV, Shaw MW, Sutton EE, von Borstel RC, Zimmering S, de Serres FJ. 1975. Environmental mutagenic hazards. *Science* 187:503–514.
95. Drake JW., 1978. Some guidelines for determining maximum permissible levels of chemical mutagens. In Flamm WG, Mehlman MA, eds, *Advances in Modern Toxicology*, Vol 5—Mutagenesis. Hemisphere, New York, NY, USA p. 926.
96. National Academy of Sciences. 1977. *Drinking Water and Health*. Safe Drinking Water Committee, Washington, DC.
97. Freese E 1973. Thresholds in toxic, teratogenic, mutagenic, and carcinogenic effects. *Environ Health Perspect* 6:171–178
98. Samson L, Cairns J., 1977. New pathway for DNA-repair in *Eschericia-coli*. *Nature* 267:281–283.
99. Bruce RD, Carlton WW, Ferber KH, Hughes DH, Quast JF, Salsburg DS, Smith JM. (Members of the Society of Toxicology ED₀₁ Task Force). 1981. Re-examination of the ED₀₁ study: Adjusting for time on study. *Fundam Appl Toxicol* 1:67–80.
100. Calabrese EJ, Baldwin LA. 2003. The hormetic dose–response model is more common than the threshold model in toxicology. *Toxicol Sci* 71:246–250.
101. Calabrese EJ, Baldwin LA. 2001. The frequency of U-shaped dose-responses in the toxicological literature. *Toxicol Sci* 62:330–338.
102. Calabrese EJ, Baldwin LA. 2003. Toxicology rethinks its central belief: Hormesis demands a reappraisal of the way risks are assessed. *Nature* 42:691–692.
103. Begley S. 2003. Scientists revisit idea that a little poison could be beneficial. *The Wall Street Journal*, December 19, Vol CCXLII (121).
104. Lambert E. 2003. A pinch of poison. *Forbes Life Online*, December 22.
105. Stipp D. 2003. A little poison can be good for you. *Fortune* 147: 54–55.
106. Hively W. 2002. Is radiation good for you? *Discover* December, 74–80.
107. Renner R. 2003. Hormesis: Nietzsche’s toxicology. *Sci Am* 28–30.
108. Raloff J. 2007. Counterintuitive toxicity. *Science News Online* January 20, 171:40.
109. Pike J. 2004. Can toxins lead to healthier lives? *Insight Magazine* January 6.
110. Bailey R. 2003. The dose makes the cure. *Reason Online* March 12.
111. Boyce N. 2004. Is there a tonic in the toxin? *U.S. News and World Report* October 18.
112. Ahuja A. 2003. Science: Some toxins may be good for us. *London Times Online* October 30.
113. Roberts R. 2003. Here’s to your health. *St. Louis Post Dispatch* January 9.
114. Cook G. 2003. A scientist finds benefit in small doses of toxins. *Boston Globe* December 12:A16–A18.
115. Bell J. 2004. Can a low dose of poison be a good thing? *Baltimore Sun* March 15.
116. Kaiser J. 2003. Sipping from a poisoned chalice. *Science* 302:376–379.
117. Calabrese EJ, Stanek EJ III, Nascarella MA, Hoffmann GR. 2008. Hormesis predicts low-dose–responses better than threshold models. *Int J Toxicol* 27:369–378.
118. Calabrese EJ, Hoffmann GR, Stanek EJ III, Nascarella MA. 2010. Hormesis in high-throughput screening of antibacterial compounds in *E. coli*. *Hum Exp Toxicol* 29:667–677.
119. National Academy of Sciences. 2009. Science and decisions: Advancing risk assessment. Washington, DC, USA.
120. Calabrese EJ. 2008. Hormesis and medicine. *Br J Clin Pharmacol* 66:594–617.
121. Calabrese EJ, Mattson MP. 2011. Hormesis provides a generalized quantitative estimate of biological plasticity. *J Cell Commun Signal* 5:25–38.
122. Hayes AW. 2008. *Principles and Practices of Toxicology*. Taylor and Francis, Philadelphia, PA, USA.
123. Klaassen C. 2008. *Casarett’s and Doull’s Essentials of Toxicology*, 7th ed. McGraw Hill, New York, NY, USA.
124. Hacker M, Bachmann K, Messer W. 2009. *Pharmacology: Principles and Practice*. Elsevier, New York, NY, USA.
125. Lebourg E, Rattan S. 2009. *Mild Stress in Healthy Aging: Applying Hormesis in Aging Research and Interventions*. Springer, New York, NY, USA.
126. Mattson M, Calabrese EJ. 2010. *Hormesis: A Revolution in Biology, Toxicology and Medicine*. Springer, New York, NY, USA.
127. Saunders C. 2010. *Radiation Hormesis and the Linear No-threshold Assumption*. Springer, New York, NY, USA.
128. Stebbing A. 2011. *A Cybernetic View of Biological Growth: The Maia Hypothesis*. Cambridge Publishers, London, UK.
129. Elliott KC. 2011. *Is A Little Pollution Good For You? Incorporating Societal Values Into Environmental Research*. Oxford University Press, London, UK.
130. Calabrese EJ, Baldwin LA. 2002. Defining hormesis. *Hum Exp Toxicol* 21:91–97.
131. Calabrese EJ. 2005. Cancer biology and hormesis: Human tumor cell lines commonly display hormetic (biphasic) dose–responses. *Crit Rev Toxicol* 35:463–582.
132. U.S. Environmental Protection Agency. 2004. An examination of EPA risk assessment principles and practices. EPA/100/B-04/001. Washington, DC.
133. Iavicoli I, Carelli G, Stanek EJ, Castellino N, Li Z, Calabrese EJ. 2006. Low doses of dietary lead are associated with a profound reduction in the time to the onset of puberty in female mice. *Reprod Toxicol* 22:586–590.
134. Iavicoli I, Carelli G, Stanek EJ, Castellino N, Calabrese EJ. 2004. Effects of low doses of dietary lead on puberty onset in female mice. *Reprod Toxicol* 19:35–41.
135. Jones A. 2010. *Knowledge and Attitudes Concerning Hormesis by Toxicologists and Risk Assessors*. PhD thesis. University of Massachusetts, Amherst, MA, USA.

Appendix 1A. Quotes relating to homeopathy and the Arndt-Schulz Law [18].

High dilution scheme of Hahnemann lacking credibility – page 24

“Hahnemann, for example, claimed that drugs at the 30th potency produced reliable effects, and actions produced by similar dilutions are still occasionally described (e.g. Konig, 1927). A homeopathic potency means a dilution of hundredfold, and hence the 30th potency corresponds to a concentration of 1 part in 10⁶⁰. This works out at about one molecule in a sphere with a circumference equal to the orbit of Venus. Such results may be either believed or disbelieved, but their acceptance involves discarding the fundamental laws of chemistry and physics.”

“Other results have been published which are almost equally improbable.”

Associates Schulz with homeopathy – page 195

“In 1885 Rudolf Arndt put forward the suggestion that if a weak stimulus excites an organism, then any drug in sufficiently weak dose ought to do this also. This suggestion was developed by Schulz, who had a leaning to homeopathy.”

Challenges the biological significance of the Arndt-Schulz Law – pages 195-196

“...many pharmacologists have pointed out that it (Arndt-Schulz Law) expresses no general truth. It is interesting to note that no trace of evidence in support of such a law can be found in the majority of drugs.”

The Arndt-Schulz Law probably confused with experimental errors – page 196

“As in the case of potential actions, evidence in favor of this law can easily be obtained from experimental errors.”

Appendix 1B. More quotes on homeopathy and the Arndt-Schulz Law [19].*Challenges mechanistic understanding of the Arndt-Schulz Law effects – page 215*

“...laws have been enunciated which merely state that certain phenomena frequently occur, without providing any explanation or their occurrence. The Arndt-Schulz Law... (is) (an) example of this type...”

Arndt-Schulz Law is usually discredited when carefully assessed – page 204

“Arndt-Schulz Law. This law states that any drug which causes stimulation at low concentrations will cause inhibition at high concentrations. This law is in accordance with homeopathic doctrines and hence has maintained a certain popularity. The law is true in so far that nearly all drugs if given in sufficiently high dosage or concentration will produce injury or death in living cells.”

“The chief objection to the law is that it is obviously untrue in the case of most drugs that have been studied carefully.”

“Many of the effects which appear to support this law have found simple explanations...”

Arndt-Schulz Law was related to vitalism – page 30/example 1

“Diphasic actions of drugs on tissues are frequently observed, and their occurrence led to the postulation of the Arndt-Schulz Law, which states that drugs which paralyze at high concentrations stimulate at low concentrations. It is true that such effects are often observed but there is no necessity to postulate any mysterious (emphasis added) property of living tissues because similar effects are frequently observed with enzyme systems.”

Arndt-Schulz Law was related to vitalism – page 30/example 2

“This peculiar effect is mentioned here because it is the simplest example known to the writer of a reaction following the “Arndt-Schulz Law”. In this case a high concentration of oxygen prevents the formation of HbCO but if hemoglobin is exposed to a low concentration of carbon monoxide, then a low concentration of oxygen may increase the formation of HbCO. Hence oxygen may be said to stimulate in low concentrations and to inhibit in high concentrations. This diphasic action can be explained on physico-chemical grounds and although our present knowledge is inadequate to explain most of the diphasic actions met with in more complex systems, yet there seems no reason to consider them as peculiarly mysterious (emphasis added).”

Challenges high dilution proposal of Hahnemann and homeopathy – page 26

“...Hahnemann claimed that drugs produced effects when given in the 30th potency... in the case of a drug with a molecular weight of 100, (this) corresponds to 1 molecule in about 100,000 liters. It is obvious that (when) a sample of a few c.c. of such a mixture is taken, the odds against the presence in the sample of a single molecule of the drug are at least a million to one. Hence the claims of the homeopathist conflict more immediately with the laws of mathematics, physics and chemistry than with the biological sciences. It does not appear necessary for pharmacologists to discuss the evidence adduced by the homeopathists until the latter have succeeded in convincing the physicists that they have demonstrated the existence of a new form of subdivision of matter. It may be mentioned that the existence of such recognized subdivisions of the atom as electrons, etc. does not help the homeopathic claims in a significant manner because, to explain the results of Hahnemann, it is necessary to assume that a molecule can be divided into millions of sub-units.”

Appendix 2. National Academy of Sciences Safe Drinking Water Committee (1977) low dose linearity guiding principles: no longer tenable three decades later [48].

Only one or two changes in a cell could transform it and this could lead to cancer.	Not tenable
Human population heterogeneity was a factor, and some people may be at greater risk. Such heterogeneity leads to the conclusion that there was no population-based threshold.	Impossible to practically study
A transformed cell will be irreversibly propagated.	Not tenable
If the mechanism involved mutation, there would be no threshold; in fact, if there were no information on mechanism and cancer occurred, mutation should be assumed.	Not tenable
It is necessary to assume that a single molecule or a few molecules can cause a mutation. Therefore, linearity at low dose can be assumed.	Not tenable
There is also the assumption that the exposure would be directly additive to background, if acting via the same mechanism. This would also support the linearity conclusion.	Generally not shown
Available mutagenicity data with radiation indicated that it was linear at relatively low doses.	Not tenable
Since chemical carcinogens act like ionizing radiation, low dose linearity should also be assumed to be the case for such chemicals.	Not tenable