Annu. Rev. Public Health. 2001. 22:15–33
Copyright © 2001 by Annual Reviews. All rights reserved

U-Shaped Dose-Responses in Biology, Toxicology, and Public Health*

Edward J Calabrese and Linda A Baldwin
School of Public Health & Health Sciences, Environmental Health Sciences Department, University of Massachusetts, Amherst, Massachusetts 01003-5712;
e-mail: edwardc@schoolph.umass.edu, lbaldwin@schoolph.umass.edu

Key Words hormesis, biphasic, optimization, adaptation

■ **Abstract** The occurrence of U-shaped dose-response relationships (often termed hormesis) has been documented in numerous biological, toxicological, and pharmacological investigations. Many of the endpoints studied are of considerable significance to public health (e.g. body weight, cholesterol levels, ethanol consumption, longevity, cancer incidence, etc). Despite the fact that U-shaped dose-responses are widely and independently observed, little attempt has been made to assess this phenomenon in an integrative manner. This review provides an overview of the historical foundations of hormesis and a discussion of its definition within a mechanistic framework. The occurrence, generalizability, and biological significance of U-shaped dose-response relationships along with the concept of biological optimality are addressed.

INTRODUCTION

In 1943, researchers at the University of Idaho reported that extracts from the red cedar tree enhanced fungal metabolism at low doses and inhibited at higher doses. The low-dose stimulation of subinhibitory doses was referred to as "hormesis," based on the Greek word meaning "to excite" (72). Over the past six decades, the term hormesis has become progressively more widespread, especially in the field of environmental toxicology and radiation biology when describing a dose-response continuum involving a low-dose stimulation/high-dose inhibition.

Despite this 60-year history of use of the term hormesis, the initial recognition that the dose-response relationship may be biphasic in nature was actually contemporary with the start of the modern experimental investigations, as seen with the work of Pasteur and Koch in the 1880s. In fact, credit is usually given to Hugo Schulz, a pharmacologist at the University of Greifswald in northern Germany,

*See related article on page 63: "Hormesis: Implications for Public Health Policy Regarding Toxicants" by L Lave.

who reported in 1887 (65) that numerous toxins enhance fungal metabolism at low but inhibit at higher concentrations.

The findings of Schulz were not controversial nor even greatly appreciated when first reported. In fact, the principal debate over the nature of the dose-response during these early years centered around the public health issue of disinfection and whether the dose-response was based on a unimolecular mechanism leading to a linear response, or on a more threshold-like response due to interindividual variability in response to the toxin exposure (17). Nonetheless, the mundane work of Schulz became highly controversial, sparking more than a century of debate on the nature of the dose-response relationship, that is, whether the fundamental shape of the dose-response is linear, threshold, or biphasic (e.g. U/J shaped).

Although this issue is ostensibly scientific in nature, it soon became politicized by Schulz, who saw his findings as providing the underlying scientific framework for the medical practice of homeopathy. Following his microbiological findings of a low-dose stimulation/high-dose inhibition, Schulz became interested in an 1884 report (see 35) indicating that highly dilute solutions of the drug veratrine were used successfully to treat salmonellosis. Because the bacteria causing this disease had just been identified, Schulz and a colleague assessed whether the veratrine would destroy the microbe in culture. However, even at very high doses, the drug had no effect on the disease-causing bacteria. In addition, the veratrine at high doses caused symptoms of gastroenteritis in humans, symptoms similar to salmonellosis. Based on these observations, Schulz concluded that the drug must be acting via a nontoxic mechanism that restores the adaptive capacity of the body. He also concluded that the findings were consistent with the homeopathic "law of similars." When this work became known by homeopathic physicians, Schulz became the object of intense attention. In a visit by Rudolph Arndt, a homeopathic physician, Schulz linked his original microbial investigations to the veratrine study and drew the conclusion that the fundamental dose-response was biphasic and that this could account for the homeopathic treatment response. This phenomenon became known as the Arndt-Schulz law, thereby providing a basis for the controversial assessment of hormesis.

Nearly 60 years after Schulz's initial discovery, Southam & Ehrlich (72), unaware of the work of Schulz and of the Arndt-Schulz law, made a similar observation and called it hormesis. Thus, hormesis has been the stepchild of controversy as well as the object of one of the most important toxicological and public health debates: defining the nature of the dose-response relationship.

Despite the controversial relationship of the Arndt-Schulz law to the medical practices of homeopathy, a long history of legitimate research in the area of chemical and radiation toxicology provides support for the phenomenon of low-dose stimulation/high-dose inhibition. In the chemical domain, considerable research in the early decades of the twentieth century established support for this perspective using plants, fungi, and bacteria. The relationship of plant research to the Arndt-Schulz law originates within a wide range of field and laboratory investigations. The principal areas where support for this perspective emerged included the

following: (*a*) the estimation of toxicity thresholds similar to what is currently defined within the context of a hazard assessment; (*b*) the evaluation of how plants respond to physical and chemical stressor agents of a limited nature; (*c*) the ability to differentiate essential nutrient functions from the capacity of nonnutritive agents to enhance growth and other metabolic functions; and (*d*) fungicidal and insecticidal treatments with a direct stimulatory impact on plant growth separate from their pesticidal actions. The concept that low doses of toxins may be stimulatory to bacteria was first reported in 1896 by Ferdinand Hueppe (44), a protégé of Robert Koch. In fact, as a result of his very strong reputation in the field of bacteriology and his promotion of the concept that low doses of toxic substances stimulate biological processes, this phenomenon became known as Hueppe's rule (as well as the Arndt-Schulz Law). Numerous papers (23, 26, 27, 43, 46, 55, 87, 88) and dissertations (41, 42) in bacteriology were published that were generally highly supportive of the hormetic hypothesis and the Arndt-Schulz law. By the 1930s, numerous leading botany and microbiology textbooks in the United States devoted considerable attention to this phenomenon and recognized it as a legitimate and central aspect of their respective fields. Similar activities were also occurring in the field of radiation biology, especially with respect to plant biology, insect responses, and immune stimulation. The principal point is that the concept of biphasic dose-responses was broad based, often reported in diverse biological models with similar quantitative descriptive characteristics.

Despite this impressive start to the development and early maturing of hormesis as a widely recognized phenomenon, it soon became the object of considerable criticism in both the chemical and radiation domains. In the chemical area, the Arndt-Schulz law was attacked by the prominent pharmacologist AJ Clark in his legendary text, *Handbook of Pharmacology* (19). His attack appeared principally motivated by the close linkage of the Arndt-Schulz law to homeopathy and the unjustified generalizations offered by Schulz and his followers. This was highly unfortunate because the legitimate criticism of Clark was entangled with an inadequate evaluation of the impressive database supporting the concept of hormesis. During the 1930s, similar criticism was directed at radiation hormesis by leading experts, again in response to extravagant agricultural and medical claims. Like Clark's, this criticism amounted to throwing the baby out with the bath water and a rapid marginalization of the concept ensued, leading to its exclusion from major toxicological and pharmacological texts and from major governmental funding programs. [For a detailed documenting of the history of chemical and radiation hormesis and why they became marginalized, see Calabrese & Baldwin (12–16).]

Even though hormesis was marginalized from the mainstream of toxicological and pharmacological research, evidence supporting it has continued to be published in highly regarded journals. Despite such findings, the area of hormesis has not received the type of critical scientific evaluation needed to assess it as a possible legitimate area of inquiry. In fact, it has become an area often frequented by scientists with ideological perspectives, thereby impeding the proper evaluation of hormesis.

DEFINITION OF HORMESIS

Dose-response relationships that display an inverted U- or U-shaped curve depend on the endpoint measured and have typically been viewed as examples of hormesis (Figure 1). As seen in Figure 1*a*, the U-shaped response would reflect a decrease in the incidence of an adverse response such as disease/injury at low dose and an increase at higher dose. In Figure 1*b*, the inverted U-shaped response may enhance growth, longevity, etc, at low but reduce it at higher doses. Thus, the occurrence of the U- or inverted U-shaped response is a reflection of the endpoint measured within the context of a single concept. Because U- and inverted U-shaped dose-response curves display similar descriptive features with respect to maximum stimulatory response, range of stimulation response, and relationship of maximum stimulation to the toxic threshold, it has been assumed that the underlying biological features of such responses are likely similar. Although this logic has been useful in developing categorical descriptions of dose-response phenomena, it is not mechanistically meaningful.

The principal historical issue affecting the acceptance of the Arndt-Schulz law (i.e. low-dose stimulation/high-dose inhibition) was whether it represented a direct stimulatory response or an overcompensation stimulatory response following an initial, and usually minor, injury (i.e. disruption in homeostasis). In fact, some argued that if the stimulation was not a direct one, it should be seen as a rejection of the Arndt-Schulz law. For example, although Manfried Fraenkel (see 34) argued that small doses can stimulate by a direct biopositive action, G Holzknecht and F Pordes denied the possibility of a direct stimulatory response without simultaneous damage (see 34). This confusion over whether the stimulatory response of the Arndt-Schulz law was a direct one or only a response to damage became an important issue that several decades later was still highly visible (68, 69). Such disputes remained active and continued to affect acceptance of the Arndt-Schulz law because it was not clear, even to experts and advocates, exactly what constituted an Arndt-Schulz law stimulation (i.e. direct or indirect).

During the 1940s, the lack of understanding of the biology inherent in the Arndt-Schulz law continued to be a critical factor in its rejection in the field of radiation health. For example, the highly prestigious Shields Warren (84) continued to promote the concept of Holzknecht and Pordes (see 34) by stating "the assumption that small doses of X-ray or radium radiation are stimulatory (the Arndt-Schulz 'law') is invalid. The slight evidences of proliferative activity offered as evidence by the proponents of this hypothesis are in fact only reparative responses to the injury that has been done." Recognition of reparative overcompensation due to radiation-induced damage was proposed in 1920 by Hektoen (38), head of Pathology at the University of Chicago, with respect to antibody production, and by Pohle (61), Koga (51), Teneff & Stoppani (79), and Schurer (66) for enhancement of reticuloendothelial activity. The key element in this assessment was the incorporation of an adequate temporal component in the study design. For example, in the case of Schurer (66), phagocytosis was inhibited during the initial 4 h after

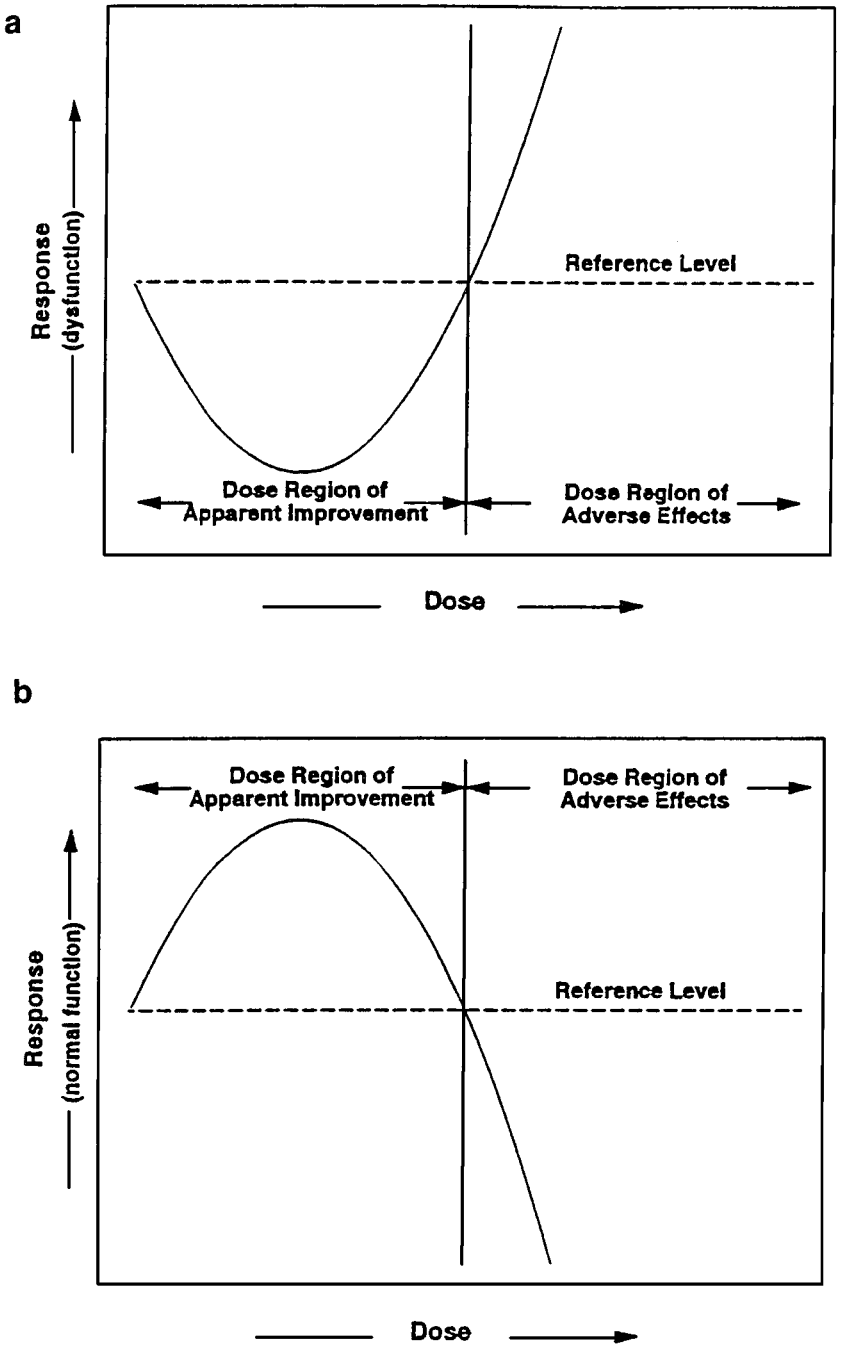

Figure 1 U-shaped dose-response curves illustrating apparent (*a*) reduced dysfunction and (*b*) enhanced function. (From Reference 22.)

exposure to X-rays; however, by 8 h after treatment, this condition had yielded to one of enhanced phagocytic activity. These findings, indicating an overcompensation response to an initial toxic insult, have been supported in later reports by Bloom & Jacobson (5), Dunlop (24), and Taliaferro & Taliaferro (78). Radiation-induced reparative responses were also comparable to the responses reported by Smith (68) for ultraviolet-induced fungal mycelium growth. That is, enhanced growth was observed only after damage and it was necessary to include a repeat measures design to properly describe this phenomenon.

It is ironic that nearly 60 years later, the definition of hormesis most prominently articulated is that of an overcompensation stimulation response following a disruption in homeostasis (77). This is the very concept that was recognized as being most consistent with the available data in the 1920s and 1940s and yet was dismissed because it was not a "direct" stimulation. It thus appears that Warren and others who rejected the Arndt-Schulz law derived a proper scientific concept but marginalized its role to the point of irrelevancy.

The rejection of the Arndt-Schulz law by prominent individuals over the observation that the stimulatory response was merely a response to damage rather than a direct stimulatory effect was perhaps the critical judgmental factor in marginalizing the hormesis concept. In fact, these dismissing individuals neglected to hypothesize that the process they were marginalizing was a basic feature of the toxicologic dose-response curve observed in plant and animal models without regard to whether the damage was induced by chemicals or radiation. The fact that the "stimulation" (i.e. overcompensation) was modest, consistently distanced (i.e. three- to sixfold) from the traditional No Observed Adverse Effect Level (NOAEL) (i.e. toxic threshold), and with a modest overall range of about one order of magnitude supported the fact that this response was likely due to a limited induction of damage. Rather than offering a refinement of a hypothesis (i.e. the Arndt-Schulz law) to incorporate an appropriate temporal experimental feature in the study design and to recognize the possible or likely role of an overcompensation reparative response to account for the quantitative aspects of the low-dose stimulatory response, the astonishing collective conclusion was to reject the Arndt-Schulz law and the hormesis concept. It is important to note that in 1929, Branham replicated the original Schulz findings with an added temporal component (7). These findings not only replicated the original low-dose stimulation/high-dose inhibition, more important they showed that the stimulation recovered only after an initial inhibitory response. If Warren (84) had been aware of the Branham (7) paper, it is possible the Arndt-Schulz law could have readily incorporated a revised overcompensation stimulation response concept.

Based on the above discussion, therefore, hormesis represents an overcompensation to a disruption in homeostasis. This concept is consistent with the limited nature of the magnitude and range of the stimulatory response and their relationship to the traditional toxic threshold. This definition indicates that in order for hormesis to be properly assessed, it is necessary to define the NOAEL/Lowest Observed Adverse Effect Level (LOAEL), to have adequate numbers of properly spaced doses below the NOAEL, and to incorporate a proper temporal framework

within the study design. In practice, however, it is unusual for toxicological studies to have an adequate number of doses, proper dose spacing, a temporal component, and appropriate endpoint selection. In fact, in the overwhelming number of cases where the so-called hormetic curve is present, the data reflect the response of only a single time point. In such cases, researchers often assume that the hormesis hypothesis is satisfied.

Even though the overcompensation hypothesis has compelling theoretical features and sufficient data to support it (12, 77), the vast majority of data used to support the hormetic hypothesis, as noted above, lacks the critical temporal information. Thus, it is not certain whether studies lacking a temporal component were derived from a direct stimulation or from an overcompensation stimulation response that has permeated and sustained the original controversy.

Despite the long-standing definitional problem of hormesis, considerable receptor-based pharmacological data accumulated over the past two decades have indicated that endogenous and synthetic agonists can directly stimulate biological responses over one dose range while inhibiting in another dose range. In these cases, there is convincing evidence that hormetic-appearing curves consistent with the direct stimulation hypothesis exist.

Although this would support the position that both direct and overcompensation stimulation can occur, it is critical to recognize that the overriding tendency of physiological systems is to reestablish homeostasis. Regardless of whether the mechanism of direct or overcompensation stimulation occurs, the response must ultimately fit into the reestablishment of homeostasis context. In the case of the overcompensation stimulation, it reflects the response to the initial disruption in homeostasis via an inhibitory response. In contrast, the direct stimulation could reflect the hormetic curve characteristics at two temporal points and in opposite directions. That is, the initial direct stimulation/inhibition may occur and represent an early disruption in homeostasis. It would be expected that homeostasis would be reestablished so that an overcompensation stimulation would subsequently occur (Figure 2). It is interesting to note that in the study of direct stimulation–appearing hormetic-like dose-responses, there are few studies where temporal features have been employed. It is ironic that investigations of both direct and overcompensation stimulation responses focus on a temporally driven dose-response of interest and generally ignore the critical temporal perspective that is necessary to study the hormetic response.

We believe that hormesis is an overcompensation to a disruption in homeostasis and should be bound by that temporally based restrictive definition. Hormesis is, therefore, a subtype of a broad array of biphasic dose-response relationships that, although displaying some similar descriptive as a direct stimulatory response with hormetic features, operate on a different mechanistic plane and are not part of the same biological process.

This restrictive overcompensation stimulation definition of hormesis offers numerous advantages over the lumping together of vast numbers of dose-response curves with similar features. It (*a*) defines its role in the biological system, (*b*) offers an experimental strategy for evaluation, (*c*) links the hormetic response to the

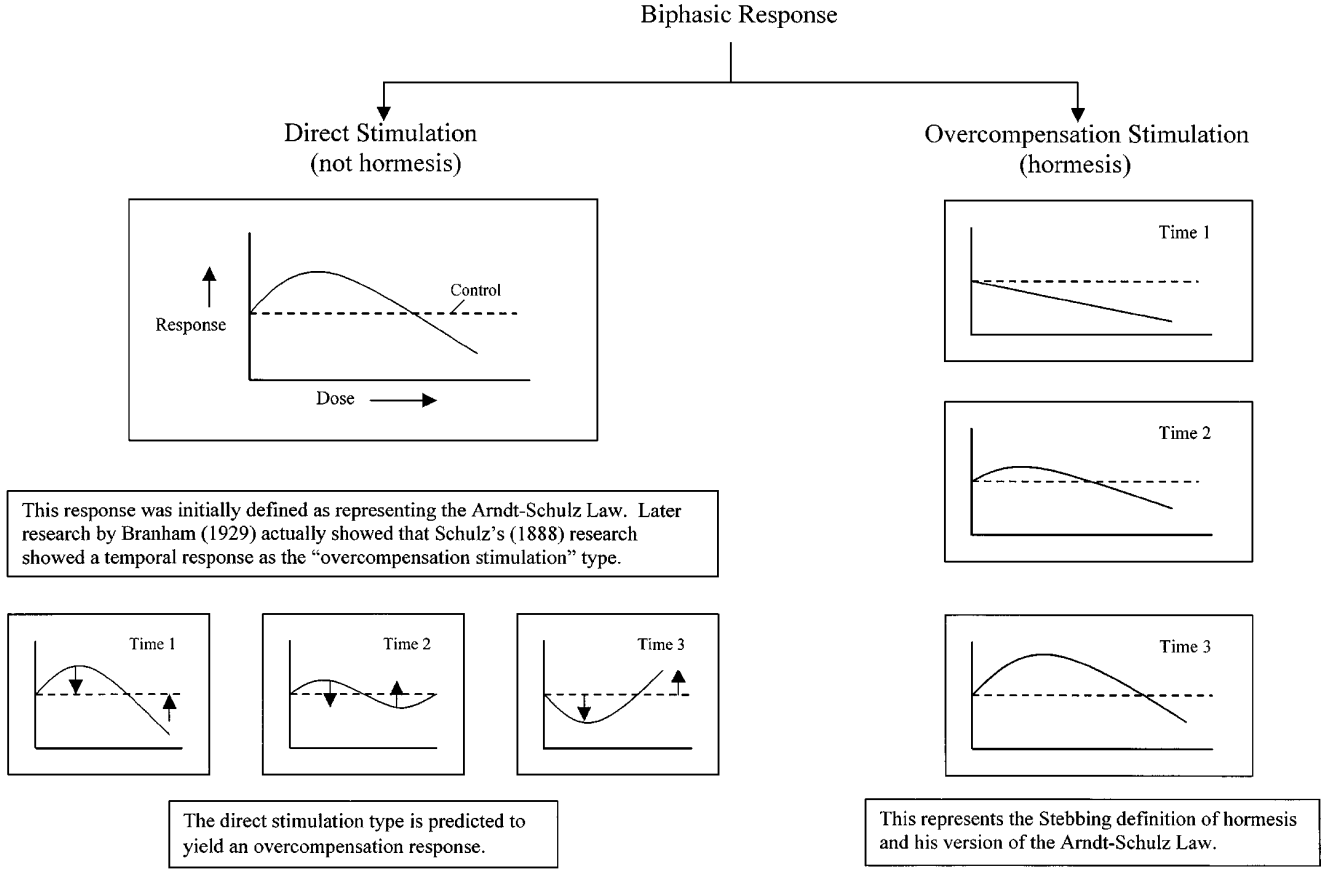

Figure 2 Comparison of direct stimulation and overcompensation stimulation.

broad toxicologically based dose-response continuum, enabling it to have direct application to the risk assessment paradigm, and (*d*) is readily evaluated within a homeostatic adaptive and evolutionary context. Within this context, the actual hormetic mechanism would relate to the specific biological system(s) undergoing the disruption in homeostasis, including which agonist receptor systems and their message systems were affected. Despite this restrictive definition of hormesis and its exclusion of the hormesis-like direct stimulation phenomenon, we recognize the existence of the direct stimulatory phenomenon and its widespread occurrence, especially in the pharmacology literature. However, the direct stimulatory response with biphasic features similar to the current definition of hormesis needs to be viewed as a different phenomenon.

The biological and public health implications of whether the low-dose stimulation is either a direct one or the result of an overcompensation response to a disruption in homeostasis remains to be explored. However, there are abundant examples in the toxicological and pharmacological literature indicating that overcompensation stimulation, which is observed during multidose and time interval study designs, is generally of a modest nature, usually maximally seen at 30%–60% greater than the control, with rare exceptions approaching twofold (11, 17). In addition, the range of the stimulatory response would usually be limited to a factor of 10, or 20 at the most, and always contiguous with a toxic response at higher doses. With respect to a direct stimulatory response, it is possible that the magnitude and range of the response could be much greater than that seen with the overcompensation stimulation response. There is also no requirement that the direct stimulatory response phenomenon be related to or contiguous with a toxic response at higher doses. In addition, the latter point is significant because it lacks the capacity to relate low-dose stimulatory responses to the NOAEL, which is a critical judgment for many risk assessment activities traditionally dealing with noncarcinogens. Consequently, it is possible that significant biological and public health implications could exist depending on whether the phenomenon represented a direct or an overcompensation response. For these reasons, it becomes critical for researchers to incorporate adequate dosing, proper spacing of doses, and temporal components within study designs in order to differentiate between direct and overcompensation stimulation mechanisms. Clearly, further research will be necessary to clarify these possible distinctions.

PUBLICATIONS WITH U-SHAPED DOSE-RESPONSES

During the past decade there has been increased recognition of U-shaped dose-responses (often called hormetic responses) in numerous biological systems. Although several significant reviews have documented the occurrence of U-shaped dose-responses in biological and toxicological systems (9, 10, 17, 18, 22), such responses have also been commonly seen in numerous pharmacological investigations (8) as well as for endpoints of considerable public health significance, including body weight and longevity (1–3, 20, 30, 36, 40, 48, 62, 63, 80–83, 89), cholesterol and longevity (28, 47), alcohol consumption and cardiovascular disease

incidence (33, 49, 50, 52, 54, 56), and exercise intensity/duration and immune responsive disease resistance (6, 29, 39, 53, 57, 71). The occurrences of U-shaped dose-responses are thus widely and independently observed phenomena. Yet, despite the widespread nature of their occurrence, little attempt has been made to assess U-shaped dose-responses as integrative phenomena. Instead they are regarded as a string of apparently reproducible, but biologically unrelated, responses. Because U-shaped dose-responses have been reported as being independent of biological model, endpoint measured, and chemical class/physical stressor (9, 10, 17), they may be legitimately considered a broadly generalizable phenomenon. However, this suggests that the widespread occurrence of these U-shaped dose-responses are in fact examples of biological optimization processes with a causal linkage to the broad concept of hormesis.

HORMESIS DATABASE

A comprehensive effort has been undertaken to identify articles demonstrating chemical hormesis (9, 10, 17). The definition of hormesis used in this research has been low-dose stimulation followed by higher-dose inhibition, although as noted above, this definition cannot differentiate direct from overcompensation stimulation. Nearly 6000 potentially relevant articles were retrieved from computer searches utilizing various key word descriptors and extensive cross-referencing. Evidence of chemical hormesis was judged to have occurred in approximately 1000 of the 6000 studies evaluated based on a priori evaluation criteria that included study design features (e.g. number of doses below NOAEL, dose response, statistical analysis, and reproducibility of results) that were integrated into a quantitative weighting scheme for numerical evaluation.

In general, the hormetic dose-response range is usually within a ten- to twenty-fold range. Stimulatory effects, however, have been reported over dosage ranges of two or more orders of magnitude, as well as over a more narrow range of dosages depending on the agent, endpoint, and model assessed (10). The magnitude of stimulatory responses has been observed as high as severalfold, but the majority of low-dose stimulations are 30%–60% greater than the controls. The distance from the maximum stimulatory response to the NOAEL is difficult to discern because it is a function of the number of doses employed, their variability in response, and the estimated value of the NOAEL. Nonetheless, the distance between the maximum stimulatory response and the estimated NOAEL is typically observed in the three- to sixfold range (i.e. the NOAEL is about three- to sixfold greater than the maximum stimulatory response).

Dose-response relationships conforming to the hormetic curve are affected by (*a*) the magnitude of the low-dose stimulatory response, (*b*) the number of doses establishing the reliability to the hormetic curve, (*c*) statistical power, and (*d*) the reproducibility of the findings. The capacity to evaluate conformity to the hormetic curve ideally requires the establishment of an endpoint-specific LOAEL and NOAEL with multiple doses within two orders of magnitude immediately

below the NOAEL. This suggests that in order to be an appropriate study for the evaluation of hormesis, an experiment would be expected to have equal to or greater than four doses distributed in a highly specific manner below the NOAEL. Therefore, highly restrictive study design requirements must be satisfied in order to adequately assess hormesis.

Evidence of hormesis was assessed by comparing the summation of point values to point ranges established for six evidence categories: high, moderate-high, moderate, low-moderate, low, and no-low. Because of the emphasis on rigorous study design and statistical analysis criteria, ∼50% of the experiments received total scores within the low evidence category, whereas 15%–20% were ranked in the high and moderate-high categories. These results reflect the scheme's strengths in that it rewards studies with (*a*) statistically significant data, (*b*) multiple doses within the hormetic zone (i.e. below the high NOAEL), and (*c*) a high magnitude of stimulatory response. Overall, the current findings are consistent with the more qualitative judgment that chemical hormesis appears to be a widely occurring phenomenon with respect to biological model, endpoint, and chemical class.

GENERALIZABILITY OF HORMESIS

Although the occurrence of hormetic responses is firmly established, the principal question is how generalizable it is. The issue of generalizability, developed in detail elsewhere (8), is based on several complementary lines of evidence.

1. Hormesis is directly linked to the concept of homeostasis. This suggests possible involvement in a vast array of biological responses. To the extent that biological systems display a limited, temporally based overcompensation after a disruption in homeostasis, the phenomenon of hormesis would occur. Numerous examples indicating overcompensation stimulation involve a wide range of biological responses and further support the broad generalizability of the hormesis concept (8). In addition, the magnitude of the overcompensation response in these examples closely follows the 30%–60% maximum range reported for examples comprising the hormesis database.
2. All receptor systems assessed to date (adenosine, adrenergic, bradykinin, cholescytokinin, corticosterone, dopamine, endothelin, estrogen, excitory amino acids, nitric oxide, numerous neuropeptides, opiates, prostaglandin, serotonin, testosterone) display biphasic dose-response relationships with identifiable mechanisms regulated by agonist concentration gradients. This integrates the hormetic response with universal implementable mechanisms to account for U-shaped responses for essentially all types of responses and represents a strong argument that most, if not all, biological endpoints are likely to display hormetic responses under specific

physiologic conditions. These findings are also consistent with the broad range of overcompensation responses.

3. Although the prediction of near universality of hormetic responses is not achieved in practice (i.e. it is not observed in all or even the majority of cases), it is nevertheless seen frequently and is reproducible. Preliminary findings with over 700 dose-response relationships satisfying a priori study design criteria for assessing the possible occurrence of hormesis reveal that the hormetic response reliably occurred in nearly 50% of the studies. The reasons it is not seen more often has been previously discussed (17) but may result from inadequate study design, lack of temporal features, and selection of endpoints that do not lend themselves to hormetic assessment (e.g. use of endpoints where the disease incidence is negligible). It is also possible, and even likely, that some biological systems will not display an overcompensation response due to other biological reasons, although this possibility requires further evaluation. Although these represent research questions, we believe that the overcompensation stimulation response is broadly generalizable, is probably universal, and represents an evolutionary strategy to select biological optimization responses.

Within the framework of natural selection, hormesis would be a particular manifestation of the concept of biological adaptation to environmental stressors. However, there was no effort to place the concept of hormesis within an evolutionary context until the work by Stebbing (73–76) and, more recently, Parsons (60) and Hart & Frame (37). According to Stebbing (75), hormesis evolved from the well-founded and highly generalizable framework of Selye (67) concerning stress and how biological systems respond to stress, including the well-known "general adaptation syndrome."

By linking the work of Selye to hormesis, Stebbing (75) theorized that natural selection processes favor generalized, rather than specific, responses to successfully deal with environmental changes and stressor conditions. Within this evolutionary paradigm, highly diverse exogenous stressors would need to be counteracted in a general way rather than by the development of a strategy of specific responses to offset the toxic effects of each stressor agent. Such a generalized counteractive responsiveness would result in organisms preadapted to respond effectively to any stressor condition (e.g. change in temperature or salinity) and/or toxic substance (e.g. heavy metals, organic contaminants, toxic metabolites). Stebbing (75) then applied the concept of regulation of nonspecific adaptive responses to environmental stressor agents to the concept of individual and ecological (e.g. population based) homeostasis. This application of nonspecific adaptive responses to assist in the maintenance of homeostasis links this to the most basic of biological processes (e.g. thermoregulation, osmoregulation, tissue repair, reproductive processes, etc.) via negative-feedback control responses. Furthermore, such control systems respond to alterations in homeostatic processes via actions that do not discriminate

between the different agents that initiated or caused the changes/disruptions. Such control systems would display nonspecific responses, counteracting the actions of a broad range of exogenous stressor conditions, involving exposure to toxic substances or changes in environmental conditions.

Relating homeostasis to stress within a toxicological context was the key conceptual framework for developing an understanding of dose-response relationships. The term load was used to describe the effects of exogenous stressor agents on the control of homeostatic processes. Imposition of any load forces the system to work harder and adds a metabolic cost to counteract the load and therefore to maintain health. Toxicity, within this conceptual model, occurs when the homeostatic capacity is unable to adequately respond because of an overload on the system. By the early 1960s, Frenster (31, 32) proposed the concept of load tolerance of numerous homeostatically controlled processes as a quantitative measure of an individual's health. The principal feature of Frenster's model is that health is concerned not only with being able to successfully deal with normal loads, but also with being able to counteract the effects of stressors that would impose an alteration or disruption to the homeostatic process. Thus, Stebbing (75) seized on the concept of counteractive capacity as an index of health and related it to the concept of dose-response in toxicology.

Using this concept of counteractive capacity with dose-time response assessment, Stebbing noted that low levels of various stressor agents (e.g. copper), inducing a minimal disruption in homeostasis in hydra colony growth, resulted in an overcorrection (i.e. increase in colony size). Such an overcorrection will eventually be reduced with the goal of reestablishing colony homeostasis. However, during the period of overcorrection, such changes would be observed as an enhancement in performance, which Stebbing (75) argued would be similar to the Arndt-Schulz law and to the definition of hormesis by Southam & Ehrlich (72).

They key conceptual development of Parsons (60) and Stebbing (75) is that low levels of stressor agents that induce a disruption of homeostatic processes will effect, in the terminology of Stebbing (75), a temporally limited overstimulation or, in the terminology of Parsons (60), a "subsidy," which both term as hormesis.

Despite the compelling rationale for an evolutionarily based explanation of "hormetic" responses, it is unfortunate that the Arndt-Schulz law evolved in close association with the medical practice of homeopathy (19) and isolated pharmacological observations (85) rather than as a necessary component of evolutionary theory with the concept of adaptation as its underlying biological mechanism. The role of hormesis should be considered within the context of an evolutionary natural selection process involving toxicological mechanisms as part of a strategy to enhance survival to low levels of stressor agents. Within this evolutionarily based framework, hormesis becomes a critical determinant affecting the nature of dose-response relationships, especially in the area of low or modest disruption of

homeostatic processes (i.e. toxicity), and therefore it is a central feature in low-dose risk assessment processes.

During recent decades, underlying mechanisms have emerged concerning how organisms adapt to environmental insults and/or stressor agents. The roles of xenobiotic metabolism, synthesis of stress proteins, acute-phase proteins and antioxidant enzymes, DNA-repair strategies, tissue-repair mechanisms, induction of antioxidant/radical scavenging molecules, and other tissue adaptive strategies in assisting species in coping with environmentally induced perturbations in their homeostatic processes have been markedly changed (37). The recognition of such highly generalizable mechanisms are part of species-specific strategies for counteracting a wide range of environmental stressors of a physical, chemical, or microbial nature.

The induction of the above types of adaptive change is consistent with the concept of Stebbing (75) with respect to generalized responses to a wide range of stressor agents. For example, the acute-phase protein response is qualitatively similar to the broad spectrum of proteins synthesized in mammalian systems regardless of the type of damage and causative agent (45). This generalized phenomenon has also been emphasized for stress proteins, where so many structurally divergent toxic stimuli strikingly lead to an increased expression of the same group of proteins (45, 86). In fact, it is likely that such responses are linked in higher organisms via activation of the brain-hypothalamic-pituitary-adrenal axis (59), in which immune-derived factors, released in response to damage or inflammation, signal hypothalamic corticotropin-releasing hormone neurons to stimulate the pituitary-adrenal axis. This series of events affects the orchestration of a highly complex, yet coordinated, series of responses to counteract the effects of inflammation leading to the reestablishment of homeostasis. As a result of the integration of neurochemical, immune, and behavioral systems, changes in any particular system will initiate adjustments in others. Such an integrative response reveals an overall adaptive strategy to develop a response to exposure to homeostatically disrupting stressor agents and provides a mechanistic foundation to account for hormetic responses in complex and single-celled organisms. This provides a general mechanistically based theory to explain hormesis as credible within the context of evolutionarily selected biosynthetic control systems where there is a strong tendency for such systems to overcorrect for low levels of inhibitor (i.e. toxicant) challenge via a nonspecific response to homeostatic disruption.

BIOLOGICAL OPTIMIZATION

U-shaped curves have been reported for various biological endpoints, including body weight, cholesterol, exercise, ethanol consumption and cardiovascular disease incidence, stress, and numerous other examples. What, if anything, do these have in common? In each case, the situation describes a type of biological optimization zone of responsiveness, not unlike that seen with dose-response

relationships for vitamins and minerals. In all cases, either too much or too little would not achieve the optimized performance.

The U-shaped dose-responses for these endpoints of public health significance reflect genetically based patterns of behaviors/responses that enhance survival (i.e. an optimized nature-nurture interface). The concept of optimality is not only biological/medical, it is also considerably broader and woven into the general fabric of all biological processes. Seen within this context, it follows that physiological systems and accompanying adaptive responses serve to integrate biological optimization systems. Thus, hormesis, as a limited overcompensation to a disruption in homeostasis, may be seen as a type of optimization response sequence that ensures recovery is achieved in an efficient manner, incorporating a modest "overcorrection" that not only achieves the needed repair but also provides sufficient biological insurance that protection is available to prevent damage from a more-massive subsequent exposure for a limited time (several days). This latter biological response would be a manifestation of the widely demonstrated adaptive response seen in chemical and radiation toxicology, where a low prior dose protects against lethality from a more-massive subsequent exposure (4).

The overcompensation stimulation to limited injury phenomenon (i.e. hormesis) is no less a component of biological optimization than the above-mentioned U-shaped dose-responses. Although this process is generally related to immediate damage and its repair, other biological optimization responses—although not responding to immediate damage, as in the case of hormesis—recognize either physiological stress or other signals that predispose the system to achieve the optimal zone within a beneficial zone over a prolonged period, making the hormesis concept applicable over a broad spectrum of temporal frameworks.

Optimality is now widely seen in biological and social systems (64). Optimality theory is widely discussed as a component of evolutionary biology (21, 25). In medicine/public health, the U-shaped dose-response phenomenon has led to questioning of long-standing treatment goals that equate to minimization rather than optimization. This has been highlighted in a recent report by Okumiya et al (58) and an accompanying editorial (70), which discuss the observation that older adults with systolic blood pressure less than 125 mmHg had a twofold higher risk of all-cause mortality compared with subjects with systolic blood pressure of 125–134 mmHg. These findings were not seen as quirks of study design and are consistent with the growing body of similar findings.

However, the optimization phenomenon extends to other fields, such as psychology, cybernetics, ergonomic, economic, and even perhaps physical systems as well. When seen in this context, hormesis represents a subset of a much broader concept.

As in the case of medicine, which has been guided by principles of minimization to the possible detriment of some members of the population, the field of environmental risk assessment may need to reflect on how its regulatory policies and standard-setting procedures of minimization correspond to the emerging concept of optimization as seen through the hormetic concept.

Visit the Annual Reviews home page at www.AnnualReviews.org

LITERATURE CITED

1. Allison DB, Faith MS, Heo M, Kotler DP. 1997. Hypothesis concerning the U-shaped relation between body mass index and mortality. *Am. J. Epidemiol.* 146:339–49
2. Allison DB, Gallagher D, Heo M, Pi-Sunyer FX, Heymsfield SB. 1997. Body mass index and all-cause mortality among people age 70 and over: the longitudinal study of aging. *Int. J. Obes.* 21:424–31
3. Andres R. 1985. Mortality and obesity: the rationale for age-specific height-weight tables. In *Principles of Geriatric Medicine*, ed. R Andres, EL Bierman, WR Hazzard, pp. 311–18. New York: McGraw-Hill
4. BELLE Newsletter. 1999. Adaptive response induced by low levels of radiation. *BELLE Newsl.* 7(3):1–31
5. Bloom W, Jacobson LO. 1935. Some hematologic effects of irradiation. *Blood* 3:586–92
6. Brahmi Z, Thomas JE, Park M, Dowdeswell JRG. 1985. The effect of acute exercise on natural killer-cell activity of trained and sedentary human subjects. *J. Clin. Immunol.* 5:321–28
7. Branham SE. 1929. The effects of certain chemical compounds upon the course of gas production by baker's yeast. *J. Bacteriol.* 18:247–64
8. Calabrese EJ. 2001. Scientific foundations of chemical and radiation hormesis. *Crit. Rev. Toxicol.* In press
9. Calabrese EJ, Baldwin LA. 1997. The dose determines the stimulation (and poison): development of a chemical hormesis database. *Int. J. Toxicol.* 16:545–59
10. Calabrese EJ, Baldwin LA. 1997. A quantitatively-based methodology for the evaluation of chemical hormesis. *Hum. Ecol. Risk Assess.* 3:545–54
11. Calabrese EJ, Baldwin LA. 1998. A general classification of U-shaped dose-response relationships in toxicology and their mechanistic foundations. *Hum. Exp. Toxicol.* 17:353–64
12. Calabrese EJ, Baldwin LA. 2000. Chemical hormesis: its historical foundations as a biological hypothesis. *Hum. Exp. Toxicol.* 19:2–31
13. Calabrese EJ, Baldwin LA. 2000. The marginalization of hormesis. *Hum. Exp. Toxicol.* 19:32–40
14. Calabrese EJ, Baldwin LA. 2000. Radiation hormesis: its historical foundations as a biological hypothesis. *Hum. Exp. Toxicol.* 19:41–75
15. Calabrese EJ, Baldwin LA. 2000. Radiation hormesis: the demise of a legitimate hypothesis. *Hum. Exp. Toxicol.* 19:76–84
16. 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
17. Calabrese EJ, Baldwin LA, Holland CD. 1999. Hormesis: a highly generalizable and reproducible phenomenon with important implications for risk assessment. *Risk Anal.* 19:261–81
18. Calabrese EJ, McCarthy M, Kenyon E. 1987. The occurrence of chemical hormesis. *Health Phys.* 57:531–41
19. Clark AJ. 1937. *Handbook of Pharmacology*. Berlin: Springer
20. Cornoni-Huntley JC, Harris TB, Everett DF, Albanes D, Micozzi MS, et al. 1991. An overview of body weight of older persons, including the impact on mortality. *J. Clin. Epidemiol.* 44:743–53
21. Crow JF. 1991. Commentary: some optimality principles in evolution. *Behav. Brain Sci.* 14:218–19
22. Davis JM, Svendsgaard DJ. 1990. U-shaped dose-response curves: their occurrence and implications for risk assessment. *J. Toxicol. Environ. Health* 30:71–83

23. Delepine AS, Greenwood A. 1914. The action of some metals upon certain water and other bacteria. *R. Sanit. Inst. J.* 35:317–27
24. Dunlop CE. 1942. Effects of radiation on the blood and the hemopoetic tissues, including the spleen, the thymus and the lymph nodes. *Arch. Pathol.* 34:562–608
25. Emlen JM, Freeman DC, Mills A, Graham JH. 1998. How organisms do the right thing: the attractor hypothesis. *Chaos* 8:717–26
26. Fabian FW, Bryan CS. 1933. The influence of cations on aerobic sporogenesis in a liquid medium. *J. Bacteriol.* 26:543–58
27. Fabian FW, Winslow C-EA. 1929. The influence upon bacterial viability of various anions in combination with sodium. *J. Bacteriol.* 18:265–91
28. Fagot-Campagna A, Hanson RL, Narayan KMV, Sievers ML, Pettitt DJ, et al. 1997. Serum cholesterol and mortality rates in a Native American population with low cholesterol concentrations. A U-shaped association. *Circulation* 96:1408–15
29. Ferry A, Picard F, Duvallet A, Weill B, Rieu M. 1990. Changes in blood leucocyte populations induced by acute maximal and chronic submaximal exercise. *Eur. J. Appl. Physiol.* 59:435–42
30. Folsom AR, Kaye SA, Sellers TA, Hong C, Cerhan JR, et al. 1993. Body fat distribution and 5-year risk of death in older women. *JAMA* 269:483–87
31. Frenster JH. 1962. The magnitude of disease as measured by tolerance. *J. Theoret. Biol.* 2:159–64
32. Frenster JH. 1962. Load tolerance as a quantitative estimate of health. *Ann. Intern. Med.* 57:788–94
33. Friedman LA, Kimball AW. 1986. Coronary heart disease mortality and alcohol consumption in Framingham. *Am. J. Epidemiol.* 124:481–89
34. Gordon MB. 1930. The stimulative effect of roentgen rays upon the glands of internal secretion. *Endocrinology* 14:411–37
35. Grote LR. 1923. Hugo Schulz. In *Die Medizin der Gegenwart*, pp. 217–50. Leipzig, Ger.: Meiner
36. Hanson RL, McCance DR, Jacobsson LTH, Narayan KM, Nelson RG, et al. 1995. The U-shaped association between body mass index and mortality: relationship with weight gain in a Native American population. *J. Clin. Epidemiol.* 48:903–16
37. Hart RW, Frame LT. 1996. Toxicological defense mechanisms and how they may affect the nature of dose-response relationships. *BELLE Newsl.* 5:1–16
38. Hektoen L. 1920. Further observations on the effects of roentgenization and splenectomy on antibody production. *J. Infect. Dis.* 27:23–30
39. Hines MT, Schott HC II, Bayly WM, Leroux AJ. 1996. Exercise and immunity: a review with emphasis on the horse. *J. Vet. Intern. Med.* 10:280–89
40. Ho SC, Woo J, Sham A. 1994. Risk factor change in older persons, a perspective from Hong Kong: weight change and mortality. *J. Gerontol.* 49:M269–72
41. Hofmann P. 1922. *Ueber die Gultigkeit des Arndt-Schulzschen biologischen Grundgesetzes bei der Wirkung von Bakteriengiften*. Inaugural dissertation. Tierarztl. Fakultat. Munchen. 13 pp.
42. Hotchkiss M. 1923. *The influence of various salts upon the growth of Bacterium communis*. PhD thesis. Yale Univ., New Haven, CT. 113 pp.
43. Hotchkiss M. 1923. Studies on salt action. VI. The stimulating and inhibitive effect of certain cations upon bacteria growth. *J. Bacteriol.* 8:141–62
44. Hueppe F. 1896. *Principles of Bacteriology*. Chicago: Open Court
45. Huggett RJ, Kimerle RA, Mehrle PM Jr, Bergman HL. 1992. Biomarkers: biochemical physiological, and histological markers of anthropogenic stress. *Proc. 8th Pellston Workshop*, July 23–28, 1989. Chelsea, MI: Lewis

46. Hune. 1909. Die begunstigende Reizwirkung kleinster Mengen von Bakteriengiften auf die Bakterienvermehrung. *Centralb. Bakteriol. I. Abt. Orig.* 48:135
47. Jacobs D, Blackburn H, Higgins M, Reed D, Iso H, et al. 1992. Report of the conference on low blood cholesterol: mortality associations. *Circulation* 86:1046–60
48. Jarrett RJ, Shipley MJ, Rose G. 1982. Weight and mortality in the Whitehall study. *Br. Med. J.* 285:535–37
49. Klasky AL, Armstrong MA, Fiedman GD. 1990. Risk of cardiovascular mortality in alcohol drinkers, ex-drinkers and nondrinkers. *Am. J. Cardiol.* 66:1237–47
50. Klasky AL, Friedman GD, Sieglaub AB. 1981. A ten-year Kaiser-Permanente experience. *Ann. Intern. Med.* 95:139–45
51. Koga Y. 1933. Uber die Wechselbeziehungen zwischen den Veranderungen des Farbstoffspeicherungsvermogens des retikuloendothelial-Systems, der hamobakterizidie und des mineral stoffgehaltes der Gewebe bei bestrahlten Kaninchen. *Strahlentherapie* 47:201–32
52. Maclure M. 1993. Demonstration of deductive meta-analysis: ethanol intake and risk of myocardial infarction. *Epidemiol. Rev.* 15:325–51
53. MacNeil B, Hoffman-Goetz L, Kendall A, Houston M, Arumugam Y. 1991. Lymphocyte proliferation responses after exercise in men: fitness, intensity, and duration effects. *J. Appl. Physiol.* 70:179–85
54. Manttari M, Tenkanen L, Alikoski T, Manninen V. 1997. Alcohol and coronary heart disease: the roles of HDL-cholesterol and smoking. *J. Intern. Med.* 241:157–63
55. Marshall MS, Hrenoff AK. 1937. Bacteriostasis. *J. Infect. Dis.* 61:42–54
56. Moore RM, Pearson TA. 1986. Moderate alcohol consumption and coronary artery disease. A review. *Medicine* 65:242–67
57. Nieman D, Johanssen L, Lee J. 1989. Infectious episodes in runners before and after a road race. *J. Sports Med. Phys. Fit.* 29:289–96
58. Okumiya K, Matsubayashi K, Wada T, Fujisawa M, Osaki Y, et al. 1999. A U-shaped association between home systolic blood pressure and 4-year mortality in community-dwelling older men. *J. Am. Geriatr. Soc.* 47:1415–21
59. Pagliacci MC, Migliorati G, Smacchia M, Grignani F, Riccardi C, et al. 1993. Cellular stress and glucocortiocoid hormones protect L929 mouse fibroblasts from tumor necrosis factor alpha cytotoxicity. *J. Endrocrinol. Invest.* 16:591–99
60. Parsons PA. 1992. Evolutionary adaptation and stress: the fitness gradient. *Evol. Biol.* 26:191–233
61. Pohle EA. 1929. Effect of roentgen rays on the reticuloendothelial system. *Am. J. Roentgenol.* 22:439–47
62. Rissanen A, Heliovaara M, Knekt P, Aaromaa A, Reunanen A, et al. 1989. Weight and mortality in Finnish men. *J. Clin. Epidemiol.* 42:781–89
63. Rissanen A, Knekt P, Heliovaara M, Aaromaa A, Reunanen A, et al. 1991. Weight and mortality in Finnish women. *J. Clin. Epidemiol.* 44:787–95
64. Schoemaker PJH. 1991. The quest for optimality: a positive heuristic of science. *Behav. Brain Sci.* 14:205–45
65. Schulz H. 1887. Zur Lehre von der Arzneiwirdung. *Arch. Pathol. Anat. Physiol. Klin. Med.* 108:423–45
66. Schurer F. 1928. *Wien. Klin. Wchnschr.* 41:1581
67. Selye H. 1975. *From Dream to Discovery.* New York: Arno
68. Smith EC. 1935. Effects of ultra-violet radiation and temperature on Fusarium. II. Stimulation. *Bull. Torrey Bot. Club* 62:151–64
69. Smith EC. 1936. The effects of radiation on fungi. In *Biological Effects of Radiation*, ed. BM Duggar, 2:889–918. New York: McGraw-Hill
70. Somes GW, Shorr RI, Pahor M. 1999. A new twist on the J-shaped curve. *J. Am. Geriatr. Soc.* 47:1477–78

71. Soppi E, Varjo P, Eskola J, Laitinen LA. 1982. Effects of strenuous physical stress on circulatory lymphocyte number and function before and after training. *J. Clin. Lab. Immunol.* 8:43–46
72. Southam CM, Ehrlich J. 1943. Effects of extracts of western red-cedar heartwood on certain wood-decaying fungi in culture. *Phytopathology* 33:517–24
73. Stebbing ARD. 1979. An experimental approach to the determinants of biological water quality. *Philos. Trans. R. Soc. London Ser. B* 286:465–81
74. Stebbing ARD. 1981. The kinetics of growth control in a colonial hydroid. *J. Mar. Biol. Assoc. UK* 61:35–63
75. Stebbing ARD. 1981. Stress, health and homeostasis. *Mar. Pollut. Bull.* 12:326–29
76. Stebbing ARD. 1982. Hormesis—the stimulation of growth by low levels of inhibitors. *Sci. Total Environ.* 22:213–34
77. Stebbing ARD. 1997. A theory for growth hormesis. *BELLE Newsl.* 6:1–11
78. Taliaferro WH, Taliaferro LG. 1951. Effect on X-rays on immunity: a review. *J. Immunol.* 66:181–212
79. Teneff S, Stoppani F. 1935. L'influenza delle irradiazioni sulle linfoghiandole e sulla circolazione linfatica. *Radiol. Med.* 22:768–87
80. Troiano RP, Frongillo EA Jr, Sobal J, Levitsky DA. 1996. The relationship between body weight and mortality: a quantitative analysis of combined information from existing studies. *Int. J. Obes.* 20:63–75
81. Vandenbroucke JP, Mauritz J, de Bruin A, Verheesen JHH, van der Heide-Wessel C, et al. 1984. Weight, smoking, and mortality. *JAMA* 252:2859–60
82. Waaler H. 1984. Height, weight, and mortality: the Norwegian experience. *Acta Med. Scand.* S679:2–56
83. Wannamethee G, Shaper AG. 1989. Body weight and mortality in middle aged British men: impact of smoking. *Br. Med. J.* 299:1497–502
84. Warren S. 1942. The histopathology of radiation lesions. *Phys. Rev.* 25:225–38
85. Webb JL. 1966. *Enzyme and Metabolic Inhibitors*, Vols. 2, 3. New York: Academic
86. Welch WJ. 1993. How cells respond to stress. *Sci. Am.* 268:56–64
87. Winslow C-EA, Dolloff AF. 1928. Relative importance of additive and antagonistic effects of cations upon bacterial viability. *J. Bacteriol.* 15:67–92
88. Winslow C-EA, Haywood ET. 1931. The specific potency of certain cations with reference to their effect on bacterial viability. *J. Bacteriol.* 22:49–69
89. Yao C-H, Slattery ML, Jacobs DR Jr, Folsom AR, Nelson ET. 1991. Anthropometric predictors of coronary heart disease and total mortality: findings from the US railroad study. *Am. J. Epidemiol.* 134: 1278–89

Copyright of Annual Review of Public Health is the property of Annual Reviews Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.