Relating Hormesis to Ethics and Policy
Conceptual Issues and Scientific Uncertainty

George R. Hoffmann

CONTENTS
1. Models for the Effects of Chemicals and Radiation at Low Doses .................................................. 308
   1.1 Thresholds and Linear Nonthreshold (LNT) as Standard Models ..................................................... 308
   1.2 Hormesis: A Biphasic Dose-Response Relationship ........................................................................ 309
   1.3 Hormesis as a Challenge to Threshold and LNT Models ................................................................. 310
2. The Nature of Hormesis ..................................................................................................................... 311
   2.1 Difficulty of Detecting and Measuring Hormesis .............................................................................. 311
   2.2 Evidence Supports Hormesis as Real ............................................................................................... 312
   2.3 Apparent Hormesis Arising as an Artifact ...................................................................................... 314
3. Biological Stress Responses ............................................................................................................. 314
   3.1 Adaptive Responses and Preconditioning ....................................................................................... 314
   3.2 Parallels between Adaptive Responses and Hormesis .................................................................... 316
4. Mechanisms of Hormesis and Stress Responses ............................................................................. 317
5. Prospects for Hormesis in Risk Assessment .................................................................................... 318
   5.1 Controversy over Assimilating Hormesis into Policy ...................................................................... 318
   5.2 Precautionary Principle and Scientific Reality ............................................................................... 319
   5.3 Hormesis and Biomedical Ethics ..................................................................................................... 321
6. Challenges of Assimilating Hormesis into Risk Assessment ............................................................ 321
   6.1 Default Assumptions for Low-Dose Effects ..................................................................................... 321
   6.2 Disagreement about the Generalizability of Hormesis .................................................................. 322
   6.3 Disagreement about the Prevalence of Hormesis ........................................................................ 323
   6.4 Uncertainties in the Quantification of Hormetic Effects ................................................................ 323
   6.5 Accounting for Heterogeneity in Susceptibility ............................................................................ 324
   6.6 Interactions among Agents ............................................................................................................. 326
   6.7 Hormesis and Concomitant Toxicity ............................................................................................... 326
   6.8 Feasibility of Hormesis-Based Risk Assessment ........................................................................... 327
7. Why an Understanding of Hormesis is Essential ............................................................................ 327
   7.1 Concerns about Hormesis and Risks of Ignoring It ...................................................................... 327
   7.2 Optimizing the Benefits of Mild Stress Responses ....................................................................... 327
   7.3 Avoiding Unforeseen Risks to Public Health ................................................................................ 328
   7.4 Agricultural Productivity and Environmental Quality .................................................................. 329
8. Conclusions ...................................................................................................................................... 330
Acknowledgments ................................................................................................................................. 330
References ............................................................................................................................................... 331
1. MODELS FOR THE EFFECTS OF CHEMICALS AND RADIATION AT LOW DOSES

1.1 Thresholds and Linear Nonthreshold as Standard Models

Humans are routinely exposed to low doses of toxicants and radiation, but it is extremely difficult or impossible to measure biological effects at the low doses that are of interest. Dose-response models are therefore useful not only in assessing risks associated with measured effects but also in shaping our expectations for effects at dosages below which accurate measurements are impossible or impractical. Figure 1 shows the threshold and linear nonthreshold (LNT) models that have dominated thought about low doses in toxicology and radiation biology. The curves show the frequency of an adverse effect plotted against dosage. In the threshold model (A), there is a dosage below which the frequency does not differ from the unexposed control population. This dosage represents a biological threshold, often represented in toxicological studies as a "no observed adverse effect level" (NOAEL). In contrast, the linear nonthreshold model (B), often referred to as LNT, extrapolates to the spontaneous frequency on the ordinate.

The dominant dose-response model in toxicology is a threshold model. A threshold dose-response relationship (Figure 1A) has a slope of zero at low doses, followed by an increasing response, which may be nonlinear, in the zone of toxicity above the threshold. It is often represented in toxicology textbooks as a sigmoid curve that describes the proportion of a defined population showing a quantal characteristic, such as death, with increasing dosage (Eaton and Gilbert 2008). For risk assessment purposes, the aim under the assumption of a threshold model is to ensure that exposures are below the threshold or NOAEL. Although the

![Dose-response curves](image)

**FIGURE 1.** Dose-response relationships: Threshold dose-response model (A) and linear nonthreshold (LNT) dose-response model (B). The response is the frequency of an adverse effect, and the thin horizontal line is its spontaneous frequency in an unexposed control population. A threshold dose is indicated by the NOAEL (no observed adverse effect level). (Adapted from Hoffmann, G.R., *Dose-Response* 7, 1-51, 2009.)
threshold model is the prevailing assumption for biological endpoints other than mutation and cancer, its acceptance is not universal, and the debate has intensified with arguments for (White et al., 2009) and against (Rhomberg et al., 2011; Bukowski et al., 2013) the use of linear extrapolation for noncancer effects.

In contrast to the threshold model, all doses in the LNT model (Figure 1B) are considered to have an effect, and one strives to ensure that exposures are either zero or small enough that the risk is negligible. LNT has been the prevailing assumption in genetic toxicology (Doak et al., 2007; Gocke and Müller, 2009; Lutz and Lutz, 2009; Bryce et al., 2010) and carcinogenesis (EPA, 2005; Brenner and Sachs, 2006; Preston and Hoffmann, 2013). The adoption of LNT for genotoxic agents was based on conceptual, historical and experimental reasons (Hoffmann, 2009). Hit theory supporting an assumption of linearity for the induction of genetic damage was central to radiation biology, and radiation biology provided the historical foundation for interpreting chemical mutagenesis, which was discovered 15 years after the classical demonstration of x-ray mutagenesis by H.J. Muller in 1927. While the early studies concentrated on radiation doses that would be moderate or high by today's standards, linearity appeared to extend to the lowest doses (Brenner et al., 2003). However, resolution of the shape of dose-response curves for mutagenesis in the low-dose zone is extremely difficult, owing to the fact that the low spontaneous frequencies of genetic alterations make it hard to measure small changes in their frequency.

Evidence that accumulated over decades made it clear that mutagenesis is not a unitary interaction between agent and target as envisioned in hit theory. Biological systems show genetic and physiological responses to damage inflicted by radiation and chemicals, and they cannot be thought of as inert elements to which mutagenesis simply happens. Rather, such factors as mutagen uptake and metabolism, complex interactions with DNA, cellular processing of damage through repair and recombination, regulation of cellular proliferation, and factors in mutant expression can all lead to deviations from linearity (Hoffmann, 2009). Mechanistic reasons for nonlinearity are also supported by experimental evidence that dose responses are often sublinear (Lutz and Lutz, 2009; Dobo et al., 2011). While LNT remains the best model for describing some genotoxic effects (Bryce et al., 2010; Spassova et al., 2013), thresholds are now well documented in many other cases (Doak et al., 2007; Gocke and Müller, 2009; Lutz and Lutz, 2009; Bryce et al., 2010; Hoffmann et al., 2013; Thomas et al., 2013).

1.2 Hormesis: A Biphasic Dose-Response Relationship

Hormesis differs fundamentally from the threshold and LNT models, in that the hormesis model proposes biphasic responses to toxicants or radiation (Calabrese and Baldwin, 2001a; Calabrese, 2008a; Hoffmann, 2009). The threshold and LNT curves can be described as monotonic, in that they show either an increase or a decrease in response over the full range of doses that have an effect. In contrast, the hormetic curve is nonmonotonic, meaning that the response changes in more than one direction with a unidirectional change in dose (Davis and Svendsgaard, 1994). Using the term "hormesis" does not refer to a specific mechanism or pathway but, rather, to the shape of the dose-response curve.
Figure 2. Hormetic curves. The J-shaped curve (A) shows hormesis at low doses and an adverse effect at high doses. The inverted U curve (B) shows hormesis for a biological function that is stimulated by low doses and inhibited by high doses. At the zero equivalent point (ZEP) the curve crosses the level of response of the unexposed control. (Adapted from Hoffmann, G.R., *Dose-Response* 7, 1-51, 2009.

Figure 2 shows two manifestations of hormesis. In Figure 2A, sometimes called a J-shaped curve, the response is the induction of an adverse effect at high doses, while low doses lead to a reduction of its frequency below the level in the unexposed control population. The response in Figure 2B is a biological function that is inhibited by high doses but stimulated by low doses. For example, one would expect such a curve, called "an inverted U," if a growth-inhibiting chemical actually stimulated growth in the low-dose zone. A concept befitting the hormesis model is the zero equivalent point (ZEP), defined as the point at which the biphasic curve crosses the level of response of the unexposed control (Calabrese, 2005a). The ZEP is a variation on the NOAEL that is independent of whether effects are adverse or beneficial, and it differs from a no-effect level (NOEL), in that there is a biological effect at dosages above and below this point (Hoffmann, 2009). Thus, the defining feature of hormesis is the biphasic nature of the response, not whether that response happens to be detrimental or beneficial. If hormesis is defined on the basis of opposite responses at high and low doses, it encompasses a broad array of phenomena for which low doses stimulate a process while high doses inhibit it, and vice versa. While one may be inclined to think of the hormetic zone as beneficial and the high-dose zone as harmful, there are circumstances where the reverse is true.

1.3 Hormesis as a Challenge to Threshold and Linear Nonthreshold Models

Hormesis presents a challenge for toxicology at two levels: first, it implies that the dose-response relationships that are a cornerstone of the discipline may often be wrong (Calabrese, 2005a, 2009); and second, its biphasic nature implies that low doses of toxicants and radiation may be beneficial. The most heated disagreement about hormesis concerns the latter, especially as it relates to regulatory policy and protection of public health. A perspective will be
offered in this chapter on the merits of the proposal that hormesis should be incorporated into risk assessment.

Aspects of hormesis apart from toxicological and radiological risk are less controversial, but they are not yet widely recognized. It would be unfortunate if arguments against the application of hormesis in risk assessment impeded the elucidation of hormesis as a biological phenomenon. There is growing evidence that hormesis-like mechanisms contribute to neurological, cardiovascular, skeletal and muscular well-being (Arumugam et al., 2006; Radak et al., 2008) and that mild stress can contribute to healthy aging (Rattan, 2008; Le Bourg, 2009). Hormesis is also relevant to environmental issues, since hormetic mechanisms may figure into how organisms respond to ecosystem disturbance and agricultural practices (Hoffmann, 2009).

2. NATURE OF HORMESIS

2.1 Difficulty of Detecting and Measuring Hormesis

Much evidence supports the view that hormesis is a real biological phenomenon, not merely an artifact of data selection or a consequence of random variation (Davis and Svendsgaard, 1994; Calabrese and Baldwin, 2001a; Calabrese et al., 2006). Yet, it is not readily detected. One must be able to detect change at low doses both above and below the background level of damage, and this may not be possible for effects that are rare or absent in an unexposed population (Calabrese and Baldwin, 2001b). Similarly, hormesis may not be observable when the background level of exposure is already above a toxicity threshold, as has been suggested for lead, or if the agent mimics endogenous substances, such as estrogens, that are themselves risk factors for adverse effects (Welshons et al., 2003; Hoffmann, 2009).

The design of many studies works against detecting hormesis because toxicology and radiation biology emphasize adverse effects. Doses that have little effect in preliminary studies are often not used in follow-up studies. The dosages commonly used are those that elicit measurable responses, that is, the high-dose range. Hormetric effects tend to be small deviations from the control, and there are typically too few doses below a NOAEL to evaluate the shape of the curve in the low-dose zone (Calabrese and Baldwin, 2001b). In addition, there are often too few replicates to generate the statistical power required to measure small changes in the low frequency of events in that zone. These factors can make observations of possible hormesis uncertain, in that the effects can also be attributed to random variation, and the evidence of hormesis is rarely robust enough to exclude other dose-response models (Hoffmann, 2009).

Even studies conducted on a large scale can be consistent with several models. For example, the ED_{01} study of the National Center for Toxicological Research (NCTR) evaluated the carcinogenicity of 2-acetylaminofluorene specifically to evaluate the low-dose zone. Despite using over 24,000 mice to have large sample sizes, the responses were considered to show a threshold for bladder cancer but no threshold for liver cancer (Eaton and Gilbert, 2008). Moreover, there was a debatable claim of hormesis for bladder cancer (Bruce et al., 1983;
Kodell et al., 1983). Speaking on behalf of a Society of Toxicology task force, Bruce et al. (1983) pointed out that “statistical uncertainty makes it impossible to establish the true shape of the dose-response curve.” Scaling up the size of experiments may therefore fail to resolve responses to low doses, and mechanistic evidence on the mode of action is more apt to clarify which carcinogens and carcinogenic effects exhibit thresholds (Andersen et al., 2003; EPA, 2005). Extending from thresholds to hormesis will require expanded efforts to understand hormesis mechanistically.

2.2 Evidence Supports Hormesis as Real

Early claims of hormesis were often based on the observation of curves that appeared to be biphasic. For example, Townsend and Luckey (1960) cited over 100 examples of what they called "hormoligosis" a half-century ago. The identification was based on a biphasic response, which they called a “β-pattern.” Many early reports were essentially anecdotal, in that they relied largely on an accumulation of examples and ad hoc criteria for hormesis (Calabrese et al., 1999). Nonetheless, they implied that hormesis is a common phenomenon. Stronger evidence for hormesis has come from quantification of the frequency of hormetic curves in scientific literature (Davis and Svendsgaard, 1994; Calabrese and Baldwin, 2001a, 2003) and analysis of data from the high-throughput screening of chemicals for reasons other than measuring hormesis (Calabrese et al., 2006, 2010). The reliance on literature surveys and pre-existing databases rather than experiments specifically designed to measure hormesis has been criticized (Kitchin and Drane, 2005; Mushak, 2009; Elliott, 2011), but it is nevertheless true that both scientific and historical factors have hindered a more direct approach.

Crump (2001) succinctly described a problem in many claims of hormesis, pointing out the lack of a valid statistical test for hormesis and the lack of objective criteria for measuring its prevalence. Studies suggesting that hormesis is common typically lacked an appropriate denominator, and the inclusion of studies because they appeared to be hormetic was a source of bias in some evaluations (Crump, 2001). Yet, the fact that some low-dose responses differed significantly from the control in the opposite direction from high-dose responses, while not definitive, gave credibility to the claim of hormesis. Moreover, an analysis of the literature by Davis and Svendsgaard (1994) provided limited but objective evidence of hormesis.

Davis and Svendsgaard (1994) estimated a frequency of biphasic curves from the Environmental Protection Agency Integrated Risk Information System (IRIS) database of reference doses (RfDs). An RfD is an estimate of a daily exposure that is likely to be without an appreciable risk of adverse effects over a lifetime. It is typically derived from a NOAEL by dividing by an uncertainty factor. Criteria for excluding published toxicology papers were delineated, and 147 papers were evaluated by at least two people. They acknowledged that an appropriate statistical test for frequencies of U-shaped curves was lacking, so they applied such ad hoc criteria as a change of 5% or more compared to the control in an initial evaluation and a difference of two standard errors in a follow-up evaluation. These papers contained 780 dose responses, and they judged 12% of them to meet their criteria for U-shaped biphasic responses. An unidentified “independent academic toxicologist” that they asked to review the same articles
judged the frequency of U-shaped curves to be 24%. Thus, there was an estimate of 12-24% prevalence of hormesis.

A survey of peer-reviewed toxicological literature by Calabrese and Baldwin (2001a) led to the interpretation that only a small fraction of papers (195 / 20,285) permitted an evaluation of hormesis. The majority were excluded because of the lack of at least two sub-NOAEL doses on which to make the evaluation, the lack of proper controls, or the lack of a toxic response at high dose (Calabrese and Baldwin, 2001a, 2001b). Hormesis was indicated by either of two criteria: statistical significance or a 10% difference from the control in three or more doses below a NOAEL. Of 668 dose-response relationships in the qualifying papers, 245 were judged to provide evidence of hormesis (Calabrese and Baldwin, 2001b). Slightly over 20% of responses below the NOAEL differed significantly from the control and, among these, 19.5% differed in the direction expected for hormesis, while only 0.6% differed in the direction of toxicity (Calabrese and Baldwin, 2001a). The data correspond to a prevalence of hormesis of 19 to 37%, depending on the stringency of the criteria used for classifying a response as hormetic. In a follow-up study, Calabrese and Baldwin (2003) compared the hormesis model with a threshold model, using 664 dose-response relationships that contained 1800 sub-NOAEL doses. They evaluated the assumption that the sub-NOAEL doses should fall above and below the NOAEL with equal frequency if the data fit a threshold model. The hypothetical 1:1 ratio was not observed. Rather, the ratio was 2.5:1 in the direction predicted by hormesis.

The analysis of databases from the high-throughput screening of chemicals for toxicity complements the analyses of published literature. Both support the reality of hormesis. Calabrese et al. (2006) analyzed data on yeast growth in a National Cancer Institute antitumor drug-screening database for roughly 57,000 dose responses representing over 2000 chemicals. There was a wild-type strain and 12 mutants whose genotypes are relevant to toxicant responses, such as having altered DNA repair genes. A Benchmark Dose (BMD), roughly equivalent to a NOAEL, was calculated as the minimum dose causing toxicity. If the threshold model were correct, one should expect responses below the BMD to be randomly distributed above and below the control. Growth at doses below the BMD was significantly greater than the control in all strains, both for highly toxic compounds and relatively nontoxic compounds. The mean growth for chemicals grouped by toxicity was 103.6% to 106.6% of the control when the yeast were exposed to sub-NOAEL doses, and hormetic response patterns were four times more common than would be expected by chance. Several different modes of analysis supported the occurrence of hormesis in the yeast database (Calabrese et al., 2006, 2008).

A bacterial database provided qualitatively similar evidence of hormetic effects but the magnitude of the low-dose stimulation of growth was smaller than in yeast (Calabrese et al., 2010). The data were measurements of growth of Escherichia coli after exposure to 11 concentrations of 1888 chemicals. Growth at concentrations below the threshold was significantly greater than that in the controls. The determination of the stimulatory effect was not straightforward because of differences between chemicals of different toxicity and between edge rows and interior rows of the 96-well plates used in high-throughput screening. An estimate of the hormetic effect is 1%-4% above the controls (Calabrese et al., 2010).
2.3 Apparent Hormesis Arising as an Artifact

A difficulty that plagues the detection of hormesis is that there are artifacts that can resemble hormetric responses. Controls that happen to be atypical can cause low doses that are not different from the true control value to seem hormetric (Thayer et al., 2005). For example, a higher-than-normal frequency of tumors in a control can make low doses that are consistent with the true or historical control value appear to show a hormetric response for carcinogenesis. Pooling biological endpoints can also give an erroneous appearance of hormesis (Thayer et al., 2005). For example, if a chemical causes a modest decrease in the incidence of tumors at one site but a large increase in tumors at another site at a higher dose, the composite response for total tumors may appear as a biphasic curve. Rather than being hormesis, which is by definition biphasic, such a case is actually the summing of two monotonic curves.

Essentiality is another possible cause of an illusion of hormesis (Kefferd et al., 2008). If a compound is physiologically required or can substitute for an essential nutrient, it may stimulate growth at low doses and then cause a decline in growth in the toxic zone. While the distinction between a xenobiotic and an essential nutrient is usually obvious, it may not always be so. For example, if plants are treated with a chemical mixture containing a mineral nutrient required for growth, the nutrient may stimulate growth at doses below which the toxicant exerts any effect. The resultant inverted-U curve actually represents the summing of the two curves, rather than a hormetric effect of the toxicant at low dose. Vigilance is needed to ensure that noncritical interpretation of biphasic curves does not lead to the misidentification of responses as hormetic.

3. BIOLOGICAL STRESS RESPONSES

3.1 Adaptive Responses and Preconditioning

Sequential exposures to a toxicant or radiation show that the outcome is not only dependent on the agent causing damage but also on the organism's biological response in processing the damage. A first exposure, typically at a low dose, often causes a reduction in susceptibility that manifests itself as a smaller effect of a subsequent, larger exposure. This phenomenon has been described under different names in different disciplines (Calabrese et al., 2007), but it is most commonly called an adaptive response, preconditioning, or a stress response that leads to tolerance.

The first reported adaptive response entailed diminished bacterial mutagenicity of the potent mutagen N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) caused by a small prior exposure of the bacteria to the same agent (Samson and Cairns, 1977). A few years later, studies in human lymphocytes provided the first evidence of an adaptive response to ionizing radiation. Cells chronically given a small exposure to ß particles from tritiated thymidine experienced fewer chromatid aberrations after irradiation with 1.5 Gy x rays than did cells without the
ß-particle exposure (Olivieri et al., 1984). Small priming doses of x rays similarly induced an adaptive response that conferred reduced susceptibility to the induction of chromosomal aberrations by a subsequent higher dose (Shadley and Wolff, 1987). Such responses are sometimes called stress responses, and they are known in organisms throughout the phylogenetic hierarchy. They confer resistance to diverse stressors, including hypoxia, high osmotic strength, oxidants, and various toxicants and metabolites (Samson and Cairns, 1977; Olivieri et al., 1984; Davies et al., 1995; Wiese et al., 1995; Wolff, 1998; Miura, 2004; Calabrese et al., 2007; Hoffmann, 2009; Guan et al., 2012; Morano et al., 2012).

There is overlap between adaptive responses, in that exposure to one agent may confer resistance to others. For example, a small exposure to H$_2$O$_2$ can confer resistance not only to H$_2$O$_2$ but also to other inducers of oxidative stress, such as menadione or paraquat (Temple et al., 2005). Conversely, resistance to H$_2$O$_2$ can be induced by exposures to NaCl (Guan et al., 2012), heat, or lipid peroxidation products (Temple et al., 2005). In some instances an adaptive response to one agent may be accompanied by enhanced susceptibility to another, as in the case of γ-radiation reducing the induction of chromosomal damage in human lymphocytes by bleomycin or mitomycin C but increasing the damage caused by methyl methanesulfonate (Wolff et al., 1988).

While adaptive responses often show cross-resistances (Wolff, 1996; Wheeler and Wong, 2007), these are not always reciprocal, and the patterns suggest that there are several adaptive pathways with overlapping components (Temple et al., 2005; Morano et al., 2012). The dose dependence of the induction of adaptive responses is not well understood. In some cases it occurs within a window of effective dosage, such that smaller doses are insufficient to trigger the response, and higher doses are ineffective in inducing the response or contribute to damage to an extent that swamps the adaptive response. Thus, the induction is biphasic and resembles a hormetic curve. For example, in the cytogenetic studies of Shadley and Wolff (1987), an adaptive response to x rays was observed after priming doses from 0.5 to 20 cGy, but not after exposure at high doses. Similarly, a priming dose of 13 cGy x rays was optimal for the induction of an adaptive response that made the growth of HE22 human embryonic fibroblasts more tolerant of a challenge dose of 2 Gy (Ishii and Watanabe, 1996). Unlike these studies, which suggest a biphasic induction, others found similar adaptive responses over a broad range of priming doses. For example, doses of 1 mGy to 500 mGy γ rays at a low dose rate conferred roughly the same extent of reduced susceptibility to the induction of micronuclei by 4 Gy γ rays in human AG1522 fibroblasts (Broome et al., 2002). Similarly, an adaptive response to the induction of chromosomal inversions by x rays in pKZ1 mice was induced to a comparable extent by a 1000-fold range of priming doses (Day et al., 2007). The explanation for such discrepancies may lie in factors other than dose itself, such as other stressors, dose rate, physiological conditions, and genetic susceptibility (Mitchel, 2010).

The term “adaptive response” is common in genetics, but in other fields similar phenomena are sometimes called “preconditioning” (Murry et al., 1986; Arumugam et al., 2006; Lin et al., 2008). The unifying feature that defines these responses is that cells or organisms that are exposed to a mild stress become tolerant of more severe stress (Arumugam et al., 2006). For example, the severity of myocardial infarction caused by coronary occlusion in dogs...
was reduced if the dogs were preconditioned by exposing them to brief ischemia from mild coronary occlusion (Murry et al., 1986). A similar phenomenon has been reported in humans, in that survival after cardiogenic shock was higher in patients who had preinfarction angina than in those without angina (Le Bourg, 2009). Taken as a whole, adaptive responses, preconditioning, and the overlapping qualities of the various biological responses suggest the existence of a broad family of conserved responses to environmental stressors.

3.2 Parallels between Adaptive Responses and Hormesis

A possible relationship of hormesis to stress responses was noted by an early advocate of hormesis, T.D. Luckey, who suggested that levels of a stressing agent that are too small to be detrimental will be stimulatory to the organism (Luckey, 1968). About 30 years later, a paper coauthored by a large group of scientists proposed that adaptive responses are manifestations of hormesis (Calabrese et al., 2007). A striking parallel between them is that both entail a reduction in damage conferred by a small exposure to a stressor. While the proposed relationship is conceivable, it is not trivial that the conditions under which the two phenomena are observed differ sharply. Hormesis, as traditionally defined, involves a biphasic response to a single exposure. The hormetic response reduces the damage to a level below that of the control. In contrast, adaptive responses depend on sequential exposures, in which the first exposure modifies susceptibility to the second.

Hormesis and adaptive responses share a temporal component that affects their detection and quantitative characteristics. In both cases, one should expect a lag after exposure for the induction of the response, a period of protection, and then a return to the ground state (Calabrese and Baldwin, 2001b; Calabrese et al., 2007; Hoffmann, 2009). This pattern has been observed for adaptive responses in various experimental systems (Shadley and Wolff, 1987; Ishii and Watanabe, 1996; Stecca and Gerber, 1998; Wolff, 1998; Zhang et al., 2009; Guan et al., 2012; Hoffmann et al., 2013). Therefore, measurements made too early or too late may fail to detect an adaptive response or hormesis.

Some studies showing an adaptive response also suggest the occurrence of hormesis, per se. For example, Davies et al. (1995) reported that yeast exposed to low doses of \( \text{H}_2\text{O}_2 \) survived a subsequent higher dose of \( \text{H}_2\text{O}_2 \) and continued to divide at normal rates, whereas yeast that had not been pretreated were arrested by the challenge. The response for single exposures to \( \text{H}_2\text{O}_2 \) suggested hormesis, in that viability increased to 125% of the control value at 0.4 mM and then declined sharply at higher doses (Davies et al., 1995). Adaptive responses are more easily detected and measured than hormesis because an adaptive response entails a substantial reduction in a high level of damage caused by a high dosage. In contrast, the detection of hormesis requires measuring a modest reduction in a low level of spontaneous damage (Figure 2A) or a slightly enhanced biological function (Figure 2B). It is therefore not surprising that most reports of adaptive responses do not provide evidence of whether the adaptive response is associated with hormesis per se.

It is uncertain whether the parallels between hormesis and stress responses are sufficient to consider them manifestations of the same phenomenon. Hormesis may be
separable from adaptive responses for mechanistic reasons. One should not expect to see hormesis if the adaptive response is based on an inducible response that is specific to a lesion that does not contribute significantly to spontaneous damage (Hoffmann et al., 2013). It has been suggested that the proposed linkage is an excessively broad application of the hormesis concept (Jonas, 2010; Elliott, 2011). The finding that the same conditions that induce an adaptive response to H$_2$O$_2$ in a yeast genetic assay show no evidence of hormesis in its original sense is consistent with this view, but it is difficult to exclude other explanations, such as subtle differences in the time course for expression of the responses (Hoffmann et al., 2013). Given the fact that adaptive responses and hormesis are measured under different circumstances and with different levels of difficulty, the hypothesis of a fundamental linkage is difficult to test experimentally. In lieu of further evidence, it seems preferable to maintain a clear distinction between hormesis as originally defined and stress responses that depend on sequential treatments. While it is possible that both are part of a broad family of evolutionarily conserved responses to stress (Calabrese et al., 2007), pooling the two kinds of phenomena may obfuscate the elucidation of the basic properties of hormesis itself.

4. MECHANISMS OF HORMESIS AND STRESS RESPONSES

There is a growing understanding of mechanisms of how adaptive responses prevent damage and enhance repair (Stecca and Gerber, 1998; Wolff, 1998; Miura, 2004; Hoffmann, 2009; Morano et al., 2012). For example, responses to oxidative stress involve cell-cycle alterations and such antioxidant defenses as endogenous scavengers and detoxication enzymes that inactivate reactive oxygen species, reduce their production, and repair the damage that they cause (Benzie, 2000; Miura, 2004; Arumugam et al., 2006; Morano et al., 2012). Such responses involve reorganization of gene expression and metabolism that occurs by the regulation of transcription, translation, and posttranslational processes (Temple et al., 2005; Shenton et al., 2006; Morano et al., 2012).

The relative ease of detection makes adaptive responses more amenable to mechanistic exploration than is hormesis. Some adaptive responses are well characterized at the mechanistic level, such as inducible DNA repair making an organism less susceptible to a second challenge. Such mechanisms could, in principle, lead to hormesis in the traditional sense if a small exposure activates a repair process that also removes spontaneously occurring damage. It has been speculated that hormesis occurs when there is a disruption of homeostasis, and the biological system responds to the stress with overcompensation (Calabrese, 2002; Conolly and Lutz, 2004) as balance is reestablished (Rattan, 2008). Such speculation draws support from parallels with adaptive responses and from those laboratory systems that permit an experimental analysis of hormesis, such as a cell transformation system in which radiation hormesis has been ascribed to inducible DNA repair (Redpath and Elmore, 2007).

Hormesis and adaptive responses can arise by upregulation of genes encoding protective proteins, growth factors, cytokines, and enzymes of signaling pathways (Stecca and
Gerber, 1998; Mattson et al., 2004; Miura, 2004; Arumugam et al., 2006; Hoffmann, 2009). Other hormetic effects have been ascribed to substances interacting with stimulatory and inhibitory receptor subtypes of regulatory systems (Calabrese, 2002; Conolly and Lutz, 2004) and to enhanced immune responses (Conolly and Lutz, 2004). Selective death may also contribute to hormesis if direct killing or apoptosis occurs preferentially in abnormal cells (Bauer, 2007; Portess et al., 2007; Redpath and Elmore, 2007). Other mechanisms include inducible repair processes; interactions among cell proliferation, cell-cycle delay, and apoptosis after DNA damage; and enhanced intercellular communication at low doses (Stecca and Gerber, 1998; Rouse and Jackson, 2002; Conolly and Lutz, 2004; Miura, 2004; Fukushima et al., 2005; Arumugam et al., 2006; Bauer, 2007; Redpath and Elmore, 2007; Calabrese et al., 2007; Portess et al., 2007; Rattan, 2008; Hoffmann, 2009). Thus, there is a multiplicity of mechanisms that may contribute to hormesis. However, their distribution among cell types, organisms, agents and biological endpoints is not yet well understood.

5. PROSPECTS FOR HORMESIS IN RISK ASSESSMENT

5.1 Controversy over Assimilating Hormesis into Policy

It has been proposed that considering hormesis in risk assessment for toxic substances and radiation can confer health benefits that would be lost by adhering to threshold or LNT models (Cook and Calabrese, 2006a, 2006b; Calabrese, 2011). A related argument is that economic resources are being wasted by regulating to levels of exposure that cause no harm and may even be beneficial (Calabrese, 2004a, 2011). These views have been strongly challenged (Axelrod et al., 2004; Thayer et al., 2005, 2006; Mushak, 2007, 2009; Elliott, 2011). Critics of hormesis commonly acknowledge that the phenomenon of hormesis occurs (Thayer et al., 2005), but they question the assertion that it is highly prevalent (Mushak, 2007). A major concern relates to public policy – the fear that acceptance of the viewpoint that beneficial effects of hormesis are likely at low doses can lead to weaker standards for environmental policies and public health protection (Axelrod et al., 2004).

The heated debate over hormesis and policy has led to the suggestion that those who may benefit from broader recognition of hormesis are influenced by conflicts of interest (Axelrod et al., 2004; Shrader-Frechette, 2008; Elliott, 2011). Not surprisingly, contrary arguments have suggested political motivations and bias favoring an exaggeration of risks that can hinder the broader acceptance of hormesis (Calabrese, 2005b; Calabrese, 2008a). Although one cannot cleanly separate scientific analysis from the social and political judgments that enter into scientific policy, the heated debate over hormesis calls for skepticism about unequivocal opinions that overlook or minimize the uncertainty that surrounds complex issues.

Monotonic dose-response models lend themselves to one principal objective – avoiding harm. Thus, the aim in public health policies based on threshold and LNT models is prevention of disease and disability. Proponents of the assimilation of hormesis into policy argue that hormesis offers the prospect of improving public health by harvesting the hormetric benefit (Cook and Calabrese, 2006a, 2006b). In areas such as preconditioning and physiological stress, this may be an achievable goal. For example, mild stress through exercise may improve health
through hormetic mechanisms while conferring negligible risk (Radak et al., 2008). In the case of exposure to toxicants, however, the risk in the toxic zone is appreciable, and attempts to acquire a hormetic benefit would entail exposures much closer to the NOAEL than in current practice (Hoffmann, 2009). Great certainty about the ability to target the hor metric zones would be required, so as not to fall into the toxic zone by error.

The hormetic zone is a relatively small target. A large database of hormetic responses indicates that the hormetic zone begins immediately below the ZEP, and its width in dosage is less than tenfold in roughly half the examples and less than 100-fold in the great majority. The database also suggests that hormetic effects are modest, typically amounting to 30-60% differences from a control value (Calabrese and Blain, 2005). Potential risks and benefits are influenced by asymmetry around the ZEP (Figure 2), in that toxic effects to the right can be large, while hormetic effects to the left would be relatively small (Hoffmann, 2009).

5.2 Precautionary Principle and Scientific Reality

A precautionary principle often underlies regulatory policy, and it may be formulated in various ways. The basic idea is that plausible evidence of likely and significant harm warrants public action to protect individuals, society or the environment even in the absence of scientific certainty about the risk (Vineis, 2005; Elliott, 2011). This view effectively shifts the burden of proof toward demonstrating the absence of risk rather than unequivocally showing its presence (Vineis, 2005). The aim of toxicological risk assessment under a threshold model is to ensure that exposures are in the no-effect zone. The imposition of a safety factor, also called an uncertainty factor, below the NOAEL is a standard means of keeping exposures substantially to the left of a NOAEL (Figure 1) and thereby minimizing risk. In the case of LNT, all doses are treated as though they confer risk, and the aim is to keep exposures low enough that the risk is negligible or acceptable. The precautionary principle in this instance would hold that it is better to overestimate risk than to underestimate it. Excessive caution is widely viewed as preferable to too little caution, and this preference guides conservative risk assessment. Risk estimation may therefore reflect a tension between basing policy on the best science available and basing it on a blend of science and conservative risk-assessment philosophy (Hoffmann, 2009).

If LNT is used in radiation risk assessment, we may be choosing a model that is incorrect in light of evidence of nonlinearity at low doses and mechanisms that can explain this nonlinearity. Nevertheless, the LNT model may be judged prudent for its tendency to overestimate risks. In supporting LNT, the U.S. National Academy of Sciences' Committee on the Biological Effects of Ionizing Radiation (BEIR) directly acknowledged this fact, justifying its use of LNT on the grounds of prudent conservatism (NRC, 1980, 2006), while also acknowledging that current data make it impossible to be confident about actual biological responses at very low doses. The Committee thereby attempted to avoid blurring the boundary between scientific interpretation and policy judgment. Thus, LNT may be the best model for radiation protection, even though it may not offer the most accurate scientific description of low-dose effects (Breckow, 2006).
As the gap between a policy position and scientific evidence becomes larger it provides impetus for finding an alternative risk assessment strategy. On the basis of cancer epidemiology and laboratory studies of carcinogenesis in animals, the French Academy of Sciences and French Academy of Medicine issued a joint report rejecting LNT as a realistic model for low doses of ionizing radiation (Aurengo et al., 2005; Tubiana et al., 2006). The report contends that carcinogenic effects at doses less than roughly 100 mSv are substantially overestimated by LNT. It argues for the existence of thresholds for radiation carcinogenesis and is receptive to hormesis (Aurengo et al., 2005; Tubiana et al., 2006). It has been suggested that about 40% of animal carcinogenicity studies with low-LET radiation give evidence of hormesis (Duport, 2003; Tubiana et al., 2006).

It has been argued that carcinogenic risk is apt to be more strongly overestimated at very low doses (e.g., below 10 mSv) than at doses whose effects are readily measured (Aurengo et al., 2005; Tubiana et al., 2006). This view has been countered by the contention that low-dose risks may actually be supralinear rather than sublinear owing to phenomena not typically included in risk assessment (Brenner and Sachs, 2006). Bystander effects are processes whereby unirradiated cells experience such effects as chromosome aberrations, mutations, morphological transformation, cytotoxicity and apoptosis if they receive signals from irradiated cells by means of diffusible messengers or cellular contact at gap junctions (Morgan, 2003a, 2003b; Zhou et al., 2003; Kadhim et al., 2013). Genomic instability refers to delayed biological effects, including mutations and chromosome aberrations, that continue to occur in the progeny of irradiated cells (Morgan, 2003a, 2003b; Kadhim et al., 2013). Such nontargeted and delayed effects may enlarge the effective target size for ionizing radiation beyond the cells that received the radiation damage (Brenner et al., 2001, 2003; Morgan, 2003a, 2003b, 2006; Zhou et al., 2003; Kadhim et al., 2013). If so, low-dose risks may be greater than previously anticipated (Kadhim et al., 2013). A much better understanding of the balance between the protective and adverse elements of low-dose phenomena will be required to resolve the controversy. Such uncertainties may place LNT in the range of a reasonable middle ground, rather than a gross overestimation (Mossman, 2001).

Conservative risk assessment is sometimes criticized because tighter regulatory practices cause an economic burden. Many would argue, however, that public health risks should take priority over economic interests and that overestimation of risk is benign. The latter point has been challenged, in that a large overestimation of risk may not support public health. Avoidance of valuable diagnostic or therapeutic procedures can be an unintended adverse effect of the exaggeration of radiation risks (Aurengo et al., 2005; Scott and Di Palma, 2006; Tubiana et al., 2006), and proponents of radiation hormesis have argued that the diagnostic procedures themselves (e.g., dental x rays, chest x rays, mammograms, thyroid scans) may confer a hormetic benefit. The argument calls for a balanced perspective on the possibility of harm stemming from the underestimation or overestimation of risks.

Growing evidence of hormesis calls for a better understanding of the phenomenon. It does not, however, necessitate changing risk assessment policies because of the possibility of hormetic effects immediately below the NOAEL. Rather, one might decide not to factor hormesis into risk assessment on the basis of uncertainty (Hoffmann and Stempsey, 2008).
Although it may be in the public interest to disregard the possibility of hormesis in risk assessment, this should not lead us to deny the existence of hormesis or to assume that it is unimportant in other circumstances. For example, hormetric effects benefiting a bacterial pathogen would be detrimental to public health, and assimilating knowledge of hormesis into policies would, in that case, be in the public interest (Hoffmann, 2009). Finding a proper way to consider hormesis with respect to public health relates to the relative weights placed on avoidance of harm versus conferring of benefits.

5.3 Hormesis and Biomedical Ethics

Policies related to toxic substances and radiation are typically based on protecting against harm, rather than accruing benefit. This emphasis is related to the ethical principles of nonmaleficence and beneficence (Beauchamp and Childress, 2001; Hoffmann and Stempsey, 2008). Nonmaleficence is based on avoiding the causation of harm, which coincides with the medical tenet “above all, do no harm.” In contrast, beneficence entails conferring a benefit that is properly balanced against risks and costs (Beauchamp and Childress, 2001). While the Hippocratic origins of medical ethics actually encompass doing good as well as avoiding harm, the latter usually takes precedence (Ross, 1988; Beauchamp and Childress, 2001).

These principles are relevant to hormesis because its biphasic responses (Figure 2) open the prospect of conferring benefit as well as avoiding harm, whereas the threshold and LNT curves (Figure 1) only lend themselves to the latter. On the basis of this difference, it has been proposed that hormesis can be used to improve public health (Cook and Calabrese, 2006a, 2006b). If doing so were to entail either giving or allowing toxicant exposures to accrue the hormetic benefit, it would represent a shift from the principle of nonmaleficence to that of beneficence (Hoffmann and Stempsey, 2008). Public health policies based on beneficence, such as fluoridation of public drinking water and mandatory vaccinations, tend to generate controversy, and they typically have to meet a high standard with respect to efficacy and safety before they can be accepted (Hoffmann and Stempsey, 2008). A lack of precision in defining and targeting a hormetic zone would make this unlikely with respect to toxicant exposures (Thayer et al., 2005, 2006). The hurdle to acceptance would undoubtedly be smaller for hormetic stressors that pose little risk of harm, such as using mild exercise to encourage healthy aging or to reduce the likelihood or slow the onset of degenerative diseases (Arumugam et al., 2006; Radak et al., 2008; Rattan, 2008; Le Bourg, 2009). However, caution is essential, as even mild stresses are not necessarily beneficial and may sometimes have negative effects (Le Bourg, 2009).

6. CHALLENGES OF ASSIMILATING HORMESIS INTO RISK ASSESSMENT

6.1 Default Assumptions for Low-Dose Effects

Risk assessment often relies on default assumptions about what one might expect at low doses unless there is compelling evidence to the contrary. Such assumptions are based on general scientific knowledge and policy judgments when specific scientific information is lacking.
(NRC, 1994; Preston and Hoffmann, 2013). Default assumptions have been used because of the impracticality or impossibility of determining what actually occurs in the low-dose zone. Although it is hoped that mechanistic understanding of the mode of action of toxicants can replace default assumptions (EPA, 2005), current approaches to risk assessment rely on a combination of the two (Bolt and Huici-Montagud, 2008; Preston and Hoffmann, 2013).

An LNT model has been the default assumption for genotoxic carcinogens (EPA, 2005; Brenner and Sachs, 2006; Preston and Hoffmann, 2013), while thresholds are widely recognized for nongenotoxic carcinogens (Bolt and Huici-Montagud, 2008) and other toxicologic effects (Eaton and Gilbert, 2008; Hoffmann, 2009). The assertion that hormesis should be the default assumption for risk assessment (Calabrese, 2004b) is a challenge to these interpretations. The question of whether hormesis qualifies as a default assumption is complicated by the difficulty of detecting hormesis in individual cases and the diversity of mechanisms that may contribute to or diminish hormetic effects. It has been argued that hormesis would only qualify as a default assumption if it were so prevalent as to be a nearly universal phenomenon (Crump, 2001). However, there is disagreement about reported frequencies of hormetic curves in toxicological studies and claims about the generalizability of the phenomenon.

6.2 Disagreement about the Generalizability of Hormesis

While the reality of hormesis has received growing acceptance, its generalizability continues to generate debate. It has been asserted that hormesis is broadly generalizable without regard to the specific agent, organism, biological endpoint or genetic susceptibility (Calabrese 2004a, 2008, 2010; Calabrese and Baldwin, 1998, 2001b; Calabrese and Mattson, 2011; Calabrese et al., 1999). This view has been contested on the basis of ambiguity in the criteria for generalizability, mechanistic and statistical considerations, and uncertainties about the frequency and reproducibility of hormetic responses (Crump, 2001; Axelrod et al., 2004; Kitchin and Drane, 2005; Mushak, 2007; Elliott, 2011).

The evidence for hormesis is stronger for some biological endpoints than for others, and carcinogenesis has often been the focus of debate. It has been claimed that hormesis can be generalized to carcinogens (Calabrese and Baldwin, 1998; Calabrese, 2008a), but this view has been challenged (Mushak, 2007). A problem in obtaining persuasive evidence of hormesis for carcinogenicity is the difficulty of measuring decreases in the spontaneous incidence of tumors in studies of limited size. Even the massive ED_{01} study was equivocal in this respect (Bruce et al., 1983; Kodell et al., 1983; Mushak, 2007). The same is true of epidemiologic studies, where the arguments often lie in confounding variables. It might be hoped that data from short-term tests for carcinogens could prove fruitful in the detection of hormesis because they offer controlled conditions and larger sample sizes. Mutagenicity and carcinogenicity have been linked by the historic development of the fields, mechanistic considerations, many agents that exhibit both properties, and the use of mutagenicity testing as an indicator or surrogate for likely carcinogenicity. While it is clear that there are nongenotoxic carcinogens and that correlations have been overstated, there remain many overlaps between these endpoints. Convincing
examples of genotoxic effects that exhibit biphasic curves in the low-dose zone can be found in the scientific literature (Hooker et al., 2004; Gocke and Müller, 2009; Thomas et al., 2013), but there is little evidence of widespread hormesis supporting the notion of generalizability.

The analysis of databases on inhibition of growth in microorganisms has provided evidence of hormesis (Calabrese et al., 2006, 2010), so databases from mutagenicity testing may seem to offer a promising way to get comparable information for genotoxicity. The most widely used of all mutagenicity tests is the Salmonella / microsome assay, commonly known as the Ames test. It has been claimed that Ames test data from systematic chemical screening show a high incidence of hormetic curves (Calabrese et al., 2011), but the methods used in this analysis have been challenged (Zeiger and Hoffmann, 2012). The Ames test, while an excellent assay for bacterial mutagenicity, should be considered an inappropriate model for studying hormesis because there is a complex interaction between mutagenicity and toxicity, such that a reduction in colony count is apt to reflect toxicity rather than a reduction in mutation frequency below the spontaneous level (Zeiger and Hoffmann, 2012). The generalizability of the hormesis model to genetic toxicology remains in doubt.

6.3 Disagreement about the Prevalence of Hormesis

There is no agreement on the frequency of hormesis that would be needed to support the view that hormesis is a highly prevalent phenomenon. Studies using clearly defined methods and criteria to evaluate published papers for hormesis are few, but they are reasonably consistent in estimating frequencies of nonmonotonic curves. Davis and Svendsgaard (1994) reported a prevalence of 12-24%, whereas Calabrese and Baldwin (2001a, 2001b) reported 19.5-37%. Such values have been interpreted as evidence both for (Calabrese and Baldwin, 2001a) and against (Mushak, 2007) the prevalence of hormesis. These numbers may themselves be challenged but, even if accepted, they do not make a compelling case for hormesis being a default assumption for risk assessment, given that two-thirds or more of dose responses do not show hormesis. On the other hand, the fact that up to one-third of responses show hormesis argues that hormetric responses are common. This constitutes good reason to be cognizant of hormesis, while showing restraint in asserting that it is a highly prevalent or broadly generalizable phenomenon.

6.4 Uncertainties in the Quantification of Hormetic Effects

To qualify as a default assumption for risk assessment, hormesis would need to be a reliable, quantitative substitute for actual low-dose data (Mushak, 2007). It is not clear that the evidence for hormesis has gone far enough beyond an accumulation of examples (Calabrese and Blain, 2005) into the realm of systematic quantitative analysis to support such an assumption (Crump, 2001; Kitchin and Drane, 2005; Mushak, 2007). It has often been claimed on the basis of examples of hormesis in peer-reviewed literature that a deviation of 30%-60% from the control is typical of hormesis (Calabrese, 2002, 2004b, 2008a, 2010; Calabrese and Baldwin, 1998; Calabrese and Blain, 2005; Calabrese and Mattson, 2011; Calabrese et al.,
The magnitude of the response has been called "the most consistent quantitative feature of the hormetic dose response" (Calabrese, 2010). By comparison, the hormetic responses in analyses of databases from chemical screening were roughly 3-7% in one study (Calabrese et al., 2006) and 1-4% in another (Calabrese et al., 2010). The reason for the difference is unclear, but it may lie in differences among species, agents, and endpoints or proportions of responses that are actually hormetic. In any case, it suggests uncertainty with respect to the generalizability of hormesis and its quantitative consistency.

Other studies suggest that adaptive responses are also highly variable. For example, in a study of ten human lymphoblastoid cell lines in a leading cytogenetics laboratory, six showed an adaptive response for the induction of micronuclei by γ-rays, three showed no adaptive response, and one showed an amplified response (Sorensen et al., 2002). The responses varied in magnitude, and only five of the six lines that showed an adaptive response responded consistently in repeat tests (Sorensen et al., 2002). Variability among donors is also reported for an adaptive response to the alkylating agent MNNG in human lymphocytes (Morimoto et al., 1986).

6.5 Accounting for Heterogeneity in Susceptibility

Heterogeneity in susceptibility to toxicants raises difficult questions about hormesis, notably whether hormetic responses are equally likely for different genotypes, and how that might relate to ethical considerations in risk assessment. Failure to consider heterogeneity among individuals has been raised as a weakness of the hormesis model for risk assessment (Axelrod et al., 2004). Although this problem applies to all models, it presents some unique challenges under the hormesis model.

It has been suggested on the basis of published dose-response curves that sensitive genotypes do show hormesis, but it occurs at lower doses (Calabrese and Baldwin, 2002a; Cook and Calabrese, 2006a; Calabrese, 2008a). Thus, the dose-response curve is shifted to the left, as shown in Figure 3. Evidence consistent with this interpretation includes the finding that a radiation-sensitive, cancer-prone mouse strain showed a longer latency period for spontaneous lymphomas and osteosarcomas when treated with a low dose of γ rays administered at a low dose rate (Mitchel et al., 2003). The fact that 12 mutant yeast strains, altered in genes that can affect toxicant responses, all showed hormesis comparable to that in the wild-type strain supports the same view (Calabrese et al., 2006). It would be premature, however, to assume that sensitive genotypes will consistently show hormesis. Adaptive responses to alkylating agents and ionizing radiation can be blocked by mutations that alter repair functions (Kleibl, 2002; Hoffmann, 2009) and by inhibition of poly(ADP-ribose)polymerase (PARP), an enzyme involved in the repair of DNA strand breaks (Shadley and Wolff, 1987; Stecca and Gerber, 1998; Miura, 2004), respectively. Hormesis in its original sense also appears to be blocked in some genotypes. For example, radiation hormesis in a mammalian cell assay for neoplastic transformation is inhibited by inhibition of PARP (Pant et al., 2003), and mutations in genes of the insulin-like signaling pathway in the worm Caenorhabditis elegans block a hormetic response to heat stress (Cypser and Johnson, 2003). The variation in these responses leaves uncertainty about how one should expect sensitive subpopulations to respond to low doses.
FIGURE 3. Hormesis and sensitive subgroups. (A) A hormetic dose-response curve for a typical population (bold line) and a sensitive subpopulation (fine line). The zero equivalent point (ZEP) corresponds to a no observed adverse effect level (NOAEL), in that the response above this dose is an adverse effect. Even if both populations show a hormetic response, part of the hormetic zone for the general population is in the toxic zone for the sensitive population. (B) A greater difference between the populations, such that there would be little overlap between their hormetic zones. (Adapted from Hoffmann, G. R. and Stempsey, W. E., BELLE Newsl. 14 (3): 11-17, 2008.)

If the hypothesis is correct that sensitive subpopulations exhibit hormesis but that it occurs at lower doses (Calabrese and Baldwin, 2002a), it may present a unique challenge for the assimilation of hormesis into risk assessment (Thayer et al., 2005, 2006). Differences in susceptibility may be associated with genetic constitution, age, sex, and health status. Figure 3 shows a hormetic curve for a sensitive subpopulation in comparison to the general population. If an attempt were made to allow exposures in the hormetic zone, it may lead to part of the population benefiting while another part is harmed. This dilemma does not exist in the threshold and LNT models because they are monotonic. Defining an ethical policy for targeting the hormetic zone under these circumstances would be problematic. In the spectrum of likely viewpoints, one pole is to stay far below the NOAEL and forego the hormetic benefit to the majority in order to protect the minority. The opposite pole is to seek the hormetic zone and provide the greatest good for the greatest number. The principle of nonmaleficence would call for a course like the former, rather than the utilitarian ethics of the latter (Hoffmann and Stempsey, 2008; Hoffmann, 2009). Of course, intermediate courses of action might also be formulated.

If the curves are like those in Figure 3A, it may be possible to have exposures below the NOAEL that benefit both populations, as suggested by Cook and Calabrese (2006a). However, if there is a larger difference between the two groups, as in Figure 3B, that goal may be unattainable. If the difference between populations were larger than those in Figure 3B, there would be no overlap in the hormetic zones. In actual situations, unlike these idealized curves, it would be unlikely that the hormetic zone and the degree of sensitivity would be so well defined that proper levels of exposure could be identified. The prudent course of action is not to use a hormesis-based risk assessment strategy.
6.6 Interactions among Agents

Unlike controlled laboratory exposures, human chemical exposures occur in mixtures of several agents, some of which are identified and others not. Effects of toxicants may be additive or they can exhibit such interactions as potentiation, synergism, or antagonism (Eaton and Gilbert, 2008). Little is known about possible interactions when several agents are simultaneously in the hormetic zone. It has been argued that if hormesis is maximal at a dose equivalent to one-fifth of the NOAEL, then concurrent exposure to more than five such agents would move the total effect into the toxic zone (Axelrod et al., 2004; Shrader-Frechette, 2008). Without a clear justification for the assumption of additivity, this assertion remains in the realm of speculation. An alternative proposal is that hormetic maxima are in the range of 30-60% and that simultaneous hormetic exposures are apt to drive the total effect to the higher part of that range (Calabrese, 2008a). Without a clear justification for the assumption of cumulative hormetic effect, whether additive or synergistic, this assertion is also speculation.

In sum, both the possibility that simultaneous exposures to agents that have similar toxic effects might move the response from the hormetic zone to the toxic zone (Thayer et al., 2005) and the possibility of cumulative hormetic effects (Calabrese, 2008a) warrant further consideration. While interactions among agents with respect to hormesis remain largely a mystery, promising studies in plants suggest that the dimensions of hormetic stimulation of mixed treatments may be predicted, at least roughly, from that of the individual responses (Belz et al., 2008). Studies in Drosophila of combined treatments with agents that induce mild stress (cold and hypergravity) support the same interpretation (Le Bourg, 2012). However, much more needs to be learned before generalizations are reached about the likelihoods of additivity, synergy, or antagonisms with respect to hormesis.

6.7 Hormesis and Concomitant Toxicity

The problems of heterogeneity in susceptibility apply not only to individuals, but also to differences among endpoints and differences among tissues and organs. The hormetic zone is near to the classical NOAEL, so it would not be unusual for a given exposure that is hormetic for one endpoint or tissue to be in the toxic zone for others. For example, a lower incidence of testicular tumors in rats treated with cadmium chloride (Waalkes et al., 1988) has been ascribed to hormesis (Calabrese and Baldwin, 1998), but it is accompanied by an increase in prostate tumors at the same doses (Waalkes et al., 1988; Thayer et al., 2006).

An analysis of carcinogenicity data from the National Toxicology Program (NTP) bioassay program indicated that over 90% of NTP-tested chemicals in the survey showed at least one statistically significant decrease in site-specific tumor frequency. Random variability can account for some of the decreases because of the many comparisons made, but others probably reflect anticarcinogenic effects (Haseman and Johnson, 1996). If combined with carcinogenic effects at a higher dose, such decreases might be ascribed to hormesis. If associated with carcinogenic effects at another site at higher dose, the pooling of the two monotonic curves may give an artificial appearance of hormesis, as has been argued about a claim of hormesis for the carcinogenicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin (Thayer et al.,
When hormesis is associated with carcinogenicity, the carcinogenic effect would typically override a small hormetic benefit in importance. If hormesis were factored into risk assessment, one would need to be confident about whether a hormetic effect for one endpoint is accompanied by detriment elsewhere.

### 6.8 Feasibility of Hormesis-Based Risk Assessment

Over five years ago, a critic of the prospect of hormesis-based risk assessment for toxicants, including chemical carcinogens, extended a challenge to its proponents. Mushak (2007) stated that "proponents have not yet laid out a convincing methodologic schematic that actually walks the reader or risk assessor through a hormesis-based quantitative risk assessment." To my knowledge, this remains true today. The pitfalls that must be overcome suggest that hormesis-based risk assessment would certainly be premature and possibly not even feasible.

### 7. WHY AN UNDERSTANDING OF HORMESIS IS ESSENTIAL

#### 7.1 Concerns about Hormesis and Risks of Ignoring It

Concerns about possible misapplications of hormesis in risk assessment have led some critics of hormesis to emphasize gaps in the evidence for its existence, to deny its common occurrence, or to minimize its potential importance (Shrader-Frechette, 2008; Mushak, 2009). However, critics have also correctly identified substantive questions that must be resolved with respect to risk assessment (Axelrod et al., 2004; Thayer et al., 2005; Mushak, 2007, 2009; Elliott, 2011). An inadvertent consequence of this debate is that it may encourage a disregard for hormesis more globally. If so, this could be disadvantageous for public health and environmental quality, because potential benefits of mild hormetic stress may go unrealized, detrimental consequences of hormesis may be overlooked, and opportunities for better environmental policies could be lost.

#### 7.2 Optimizing the Benefits of Mild Stress Responses

There is accumulating evidence that stress responses play a role in the aging process, such that mild stress confers hormetic benefits, whereas severe or chronic stress can exacerbate the degradative aspects of aging (Rattan, 2008). An understanding of hormetic mechanisms may therefore promote a healthier aging process. The hormesis model is also relevant to understanding how responses to mild stress can benefit cardiovascular, skeletal, muscular, and neurological health (Arumugam et al., 2006; Radak et al., 2008). The biphasic nature of the responses is reflected in the common observation that regular exercise, unlike inactivity, is beneficial but that the benefit can be offset by excessive exercise and overtraining (Radak et al., 2008). Conditioning can also confer benefits for the avoidance or slowing of neurodegenerative disorders (Mattson et al., 2004; Mattson and Cheng, 2006), and biphasic responses are reported from many areas of neurobiology (Calabrese, 2008b). Diverse
mechanisms can contribute to these phenomena, notably including modulation of the formation of reactive oxygen species (Arumugam et al., 2006; Radak et al., 2008). The beneficial effects of a vegetable-rich diet are ascribable, at least in part, to antioxidant effects, but it may also include hormetic effects of small exposures to toxic phytochemicals (Mattson and Cheng, 2006).

There is growing evidence that exercise, cognitive stimulation, and calorie restriction can improve longevity and lower the risk of Alzheimer’s disease, Parkinson’s disease, stroke and other age-related disorders through hormetic mechanisms (Mattson et al., 2004; Arumugam et al., 2006). Generalizations must be reached cautiously, however, as some effects may be specific to particular organisms or genotypes. For example, calorie restriction, which is known to increase longevity in rodent studies, may not do so consistently in primates (Mattison et al., 2012). It has been hypothesized that life-history strategies are important in predicting whether dietary restriction would improve longevity, such that short-lived species that spend their lives in a small geographic area are more likely to benefit than longer-lived species that can migrate over great distances (Le Bourg, 2010).

Current understanding of the extent to which hormesis influences longevity and disease resistance is at a formative stage. A lack of understanding of hormesis or failure to recognize its occurrence can impede the optimization of practices that take advantage of natural adaptive responses. A more complete understanding of biological stress responses offers the promise of improved therapies for age-related disorders and better dietary and behavioral approaches for the improvement of public health (Mattson et al., 2004; Arumugam et al., 2006; Rattan, 2008).

### 7.3 Avoiding Unforeseen Risks to Public Health

In the context of toxicologic risk assessment, it is often assumed that hormesis refers to a beneficial effect. While this may be correct at some level, it can be misleading. Calabrese has argued persuasively that hormesis is better defined on the basis of the characteristics of its biphasic dose-response relationship and not on the basis of benefit and harm (Calabrese and Baldwin, 2002b; Calabrese, 2008a). In some instances hormetic effects can be detrimental to health, and it is important that they be recognized and avoided. For example, the hormetic effect would be deleterious if a low dose of an inhibitory drug were to stimulate a harmful hyperplasia. Thus, specific instances can differ with respect to whether the high-dose range or the low-dose range is beneficial (Calabrese, 2008a; Hoffmann, 2009).

There is widespread awareness that antibiotic use selects for antibiotic resistance, but it is less generally appreciated that antibiotics in insufficient dosages can stimulate bacterial growth through a hormetic mechanism. Davies et al. (2006) advanced the argument that low doses of antibiotics commonly serve as signaling molecules and have stimulatory effects in bacteria. If unrecognized, the public health consequences can be substantial. Linares et al. (2006) found that three different classes of antibiotics can stimulate the opportunistic pathogen *Pseudomonas aeruginosa*, which colonizes the lungs of cystic fibrosis patients causing serious deterioration. The need to understand hormetic processes is obvious if antibiotics at low doses can confer hormetic benefits on bacterial pathogens to the detriment of their human hosts.
The hormetic stimulation of surviving tumor cells by low-levels of residual chemotherapy drugs after cytotoxic therapy also deserves consideration, as hormesis is reported to be common in human tumor cells (Calabrese, 2005a). The temporary stimulation of metastatic breast cancer by tamoxifen in some patients before the drug reaches sufficient dosages to be inhibitory may also have an underlying hormetic mechanism (Brandes, 2005). A lack of understanding of hormesis or failure to recognize its occurrence or its consequences can contribute to ineffectiveness in combating detrimental hormetic effects.

While the hormesis model (Figure 2) implies that benefit and harm lie on opposite sides of the ZEP, it should not be accepted uncritically that this is necessarily true of nonmonotonic dose-response curves. Some nonmonotonic responses, such as those for endocrine disruptors, are reported to entail harmful effects both at high and low doses (Timms et al., 2005; Vandenberg et al., 2012). These responses and hormetic responses both suggest that effects at low doses are not readily predicted by effects at high doses, but they have different implications with respect to risk. While the hormetic J-shaped curve (Figure 2A) suggests that a threshold model (Figure 1A) would overestimate risk at low doses, adverse effects of endocrine disruptors at low doses have been interpreted as reason to believe that a threshold model can underestimate low-dose risk (Weltje et al., 2005).

7.4 Agricultural Productivity and Environmental Quality

An appreciation for hormesis can offer useful insight for environmental policies (Hoffmann, 2009). The hormetic stimulation of bacteria by low doses of antibiotics, which is an obvious concern for public health, can also have detrimental environmental consequences through the formation of bacterial biofilms (Linares et al., 2006). One may speculate on consequences of hormesis in disturbed ecosystems, but our understanding of hormetic effects in microorganisms in natural environments is rudimentary.

Differences among plant and animal species in susceptibility to toxicants are often not well enough known to predict low-dose responses in natural environments and agricultural settings. Many herbicides, including such widely used compounds as glyphosate, show hormetic effects in plants at low doses (Duke et al., 2006; Cedergreen et al., 2007). It is difficult to detect hormetic effects in communities of mixed species because stimulatory effects associated with exposure to a chemical may also be ascribable to altered interspecific competition (Duke et al., 2006). Nevertheless, hormesis was evident in plants even when controlling for this confounding factor (Cedergreen et al., 2007). Hormetic effects of glyphosate have been measured in several plant species, and hormesis may be the explanation underlying the use of the herbicide to stimulate the accumulation of sucrose in sugar cane (Velini et al., 2008).

Although pesticides are typically applied under field conditions at doses sufficient to be effective, pests peripheral to the treatment zone may experience hormetic effects of low doses, and this may contribute to subsequent outbreaks of the pest (Morse, 1998). Different classes of insecticides, including organochlorine, organophosphate, carbamate and pyrethroid, have all
been reported to cause hormetic stimulation in insects (Morse and Zareh, 1991). For example, insecticide doses that cause high mortality when fed to citrus thrips also suppressed their reproductive rate, but doses that caused less than 1% mortality led to increased fecundity (Morse and Zareh, 1991). It is likely that such phenomena can contribute to the resurgence of pests. Conclusions, however, are not always straightforward, as ecological manifestations of hormesis, like those in public health, may involve complex patterns where a hormetic benefit in one area is offset by detriment elsewhere. For example, a hormetic increase in numbers of offspring may be offset by their lower survival (Duke et al., 2006; Hoffmann, 2009). A recent review by Cutler (2013) identifies many possible cases of hormesis in insects and discusses other factors that can mimic hormesis in insect populations, including reduced competition from other herbivores, changes in pest behavior, altered host-plant nutrition, and increased attractiveness of the host plant. Such observations argue that an understanding of toxic effects in pests, their hosts, predators and competitors should be accompanied by a better understanding of their responses to low doses (Morse, 1998; Kefferd et al., 2008; Hoffmann, 2009; Cutler, 2013).

8. CONCLUSIONS

Hormesis describes a dose-response relationship in which effects at low doses are opposite to those at high doses. Substantial evidence supports the reality of hormesis, but much disagreement remains about its prevalence and broader implications. Much of the controversy stems from the proposal that the hormetic response should be incorporated into risk assessment and public policy with respect to toxicants. Such applications would depend on hormesis being a highly prevalent and consistent response to toxicants and on the feasibility of acquiring a hormetic benefit without unduly risking toxicity. There is insufficient evidence on these points, and heterogeneity among species, genotypes, and tissues substantially complicates extrapolations. Much evidence suggests that basing toxicologic risk assessment on the principle of hormesis would be premature and is probably not feasible. At the same time, mild stress is known to stimulate adaptive responses that can be beneficial, and there is reason to think that factors such as exercise and cognitive stimulation contribute to good health through hormetic mechanisms. On the other hand, the possibility of adverse effects occurring through hormetic stimulation of bacteria, parasites and tumors also deserves consideration. An understanding of hormesis and stress responses is therefore important for public health and medicine, and it has become increasingly clear that understanding how these phenomena function in microorganisms, plants and animals can be important for environmental policies and agriculture.

ACKNOWLEDGMENTS

The author thanks Edward Calabrese for many interesting discussions of hormesis, Justin McAlister and Errol Zeiger for helpful suggestions on the manuscript, and Darlene Colonna for excellent secretarial assistance.
REFERENCES


