TOXICOLOGICAL SCIENCES **94(2)**, 368–378 (2006)
doi:10.1093/toxsci/kfl098
Advance Access publication September 1, 2006

Hormesis Outperforms Threshold Model in National Cancer Institute Antitumor Drug Screening Database

Edward J. Calabrese,*,[1] John W. Staudenmayer,† Edward J. Stanek III,‡ and George R. Hoffmann§

*Department of Public Health, Environmental Health Sciences Program, Morrill I, N344, University of Massachusetts, Amherst, Massachusetts 01003; †Department of Mathematics and Statistics, Lederle Graduate Research Tower, University of Massachusetts, Amherst, Massachusetts 01003; ‡Department of Public Health, Biostatistics and Epidemiology Program, School of Public Health and Health Sciences, Arnold House, University of Massachusetts, Amherst, Massachusetts 01003; and §Department of Biology, College of the Holy Cross, O'Neil Hall, 107, Worcester, Massachusetts 01610-2395

Received March 31, 2006; accepted August 30, 2006

Which dose-response model best explains low-dose responses is a critical issue in toxicology, pharmacology, and risk assessment. The present paper utilized the U.S. National Cancer Institute yeast screening database that contains 56,914 dose-response studies representing the replicated effects of 2189 chemically diverse possible antitumor drugs on cell proliferation in 13 different yeast strains. Multiple evaluation methods indicated that the observed data are inconsistent with the threshold model while supporting the hormetic model. Hormetic response patterns were observed approximately four times more often than would be expected by chance alone. The data call for the rejection of the threshold model for low-dose prediction, and they support the hormetic model as the default model for scientific interpretation of low-dose toxicological responses.

Key Words: **hormesis; threshold; dose-response; yeast; NCI; U-shaped; J-shaped; bell-shaped; risk assessment; carcinogens; chemotherapeutics; cell proliferation; *Saccharomyces*.**

The threshold dose-response model has long been recognized as the dominant dose-response model in the biological sciences, including pharmacology and toxicology (Clark, 1926, 1933, 1937). The threshold model dominates discussion in the leading pharmacological (Hardman and Limbird, 2001) and toxicological textbooks (Eaton and Klaassen, 2001; Hayes, 2001), development of study designs that drive hazard assessment procedures for pharmaceutical and chemical agents, and risk assessment processes used by regulatory and public health agencies worldwide. Despite this fact and a history of broad acceptance in many biological disciplines, the assumption that the threshold model should be used has been recently challenged. An alternative model, the hormesis model, has been proposed based on evidence of its generalizability by biological system, endpoint measured, chemical class tested (Calabrese, 2004, 2005; Calabrese and Baldwin, 2001a,b,c, 2003; Calabrese and Blain, 2005; Calabrese *et al.*, 1999), and high frequency in the toxicological literature (Calabrese and Baldwin, 2001b). Using *a priori* entry and evaluative criteria, Calabrese and Baldwin (2003) reported that the hormesis model far outperformed the threshold model in a toxicological assessment using approximately 800 dose-response relationships that were broadly representative of commonly employed biological models, endpoints, and chemical agents. The present paper extends these findings by a systematic in-depth analysis of 56,914 dose-response studies in yeast. The analyses demonstrate that the hormesis dose-response model strongly outperforms the threshold model when applied to the extensive and highly standardized National Cancer Institute (NCI) tumor drug screening database using yeast as the test organism.

[1] To whom correspondence should be addressed. Fax: 413-545-4692, E-mail: edwardc@schoolph.umass.edu.

MATERIALS AND METHODS

This study utilized data from the NCI yeast anticancer drug screen, described in detail by Holbeck (2004) and at the NCI Web site (http://www.dtp.nci.nih.gov/yacds/index.html). Briefly, data from stage 2, which contains the most promising compounds based on preliminary testing, were selected for evaluation; the agents were tested at five concentrations (1.2, 3.7, 11, 33, and 100μM) in 13 yeast strains. The yeast comprises a panel of *Saccharomyces cerevisiae* strains altered in DNA damage repair or cell cycle control genes, along with the wild-type (wt) strain without such genetic alterations. The NCI Web site contains a description of the genotype of each yeast strain used. The responses reported are derived from the fraction of growth of the yeast strain exposed to the compound relative to the growth of the same yeast strain treated with solvent (i.e., DMSO) control. Yeast cells in the exponential phase of growth were inoculated into synthetic complete medium containing 2% glucose and the test chemical. The starting cell density was 10^4 cells per well containing 200 μl of medium. (Julian Simon, personal communication).

This study, like any analysis of preexisting data, has limitations based on the data that are available. Factors limiting the range of questions that we could analyze were the fact that the NCI database provides the average of two responses and the difference between them but not original optical densities or

raw data permitting us to match chemicals to plates. Possible sources of variation were differences among columns or rows in 96-well plates and the lack of randomization of treatments within plates, plates within stacks, and stacks within the incubator. In the main, however, the large size, apparent quality control, and internal consistency of the NCI database minimize risks associated with experimental protocols, and we restricted our analysis to questions where experimental variation can be properly analyzed. Moreover, in each instance where such factors may be relevant to our analysis, we specifically point it out. Whenever a choice of assumptions was possible, we made assumptions so as to ensure that our analysis was conservative, in the sense that the bias, if any, would reduce the chance of observing hormesis.

Replication Procedure

Each chemical was tested four times at the same five concentrations in each of the 13 yeast strains (Table 1). Ninety-six–well plates were used, with 80 chemicals being tested at the same concentration (1.2, 3.7, 10, 33, or 100μM) on one plate, leaving the 16 peripheral wells for controls. Each concentration for that drug was incubated over the same 12-h period on a different plate such that there were five plates run on the same chemical at the same time. Of the 16 control wells, four were allocated to unexposed controls, eight to solvent controls, and four to controls using cycloheximide. All strains were cycloheximide sensitive, and the assay was deemed invalid if there was growth in the presence of cycloheximide. The order of the plates in the growth chamber with respect to chemical tested was systematic and not randomly allocated. The relative position of the controls on each plate was constant. The variability in response was designed to be maximized by using a different source of chemical on a different day and different yeast cultures in each different test (Julian Simon, personal communication). Slightly greater evaporation from the peripheral wells containing the controls may have caused control values to be several percentage points above a normal background. This potential bias was systematic for all plates and, if present, would introduce a negative bias with respect to discerning possible stimulatory responses (Faessel *et al.*, 1999). To ensure that our analysis was conservative with respect to the detection of hormesis, no correction was made for this factor. Factors other than evaporation from peripheral wells may contribute to variation in data from 96-well plates. An analysis of such factors by Faessel *et al.* (1999) found that differences associated with positions of plates in stacks and of stacks in an incubator tended to be smaller than differences between the middle and edges of plates. Moreover, the importance of these sources of variation is minimal for our purposes as each plate contained its own set of controls.

The response data consisted of a ratio of the optical density (OD) of the response well for the treatment divided by the mean of the OD readings of the eight solvent control wells for each concentration. OD readings were at 600 nm. This process was repeated on the second day, and the ratios from the 2 days were averaged. We refer to the average response as the replication response. Two replication responses were produced for each concentration and for each strain/experiment. Data on the NCI Web site provide the average of the two response values and the difference between the two values, but the original OD values are not available. Data were also not available to match chemicals to plates to take advantage of plate effects. It is known that even if the control and treatment responses have the same mean, then the mean of their ratio will be greater than 100%, and the mean of the ratio approaches 100% as the variance of controls gets smaller (Casella and Berger, 2002). This creates a slight bias in favor of a hormetic model, but since the design used eight controls for each response, the effect is slight. Assuming a lognormal model, if the coefficient of variation is 10%, then the mean of the ratio of a control and treatment response with the same mean is 100.1%.

Evaluation Strategy

Since the goal of this research is to evaluate whether there is nonrandom biological activity as measured by cell proliferation below the toxicological threshold, it is necessary to evaluate individual dose-response relationships. In the five-concentration protocol of NCI, the dose-response study should ideally have at least one concentration in the toxic (i.e., above threshold) domain, a concentration with a response that approximates the control response (i.e., the so-called no observed effect level [NOEL] or the highest dose that does not differ in a significant manner from the control) and several lower concentrations that would be evaluated for biological activity below this NOEL or threshold-like value. The NCI yeast database is unique and useful for this purpose but not ideal, in that some experiments show toxicity but insufficient doses below the toxicologic threshold. Our *a priori* entry criterion required at least one measurement at a concentration below that that was used to estimate the

TABLE 1
NCI Yeast Screening Replication Methodology

	Replication 1	Replication 2
Day 1	• Five concentrations, each single concentration on a different 96-well plate • Eight solvent control wells per 96-well plate • The single treatment value for each concentration is divided by the average of the eight solvent control values for each well plate	
Day 2	• Five concentrations, each single concentration on a different 96-well plate • Eight solvent control wells per 96-well plate • The single treatment value for each concentration is divided by the average of the eight solvent control values for each well plate The average of day 1 and 2 values creates what is designated as replication 1	
Days 3 and 4		• The procedures on days 3 and 4 were identical to those on days 1 and 2 The average of day 3 and 4 values creates what is designated as replication 2

toxicological threshold. A response in the toxic zone at the highest concentration was considered desirable but not essential since lack of toxicity at the highest concentration in a screening bioassay may simply reflect an inadequate concentration range.

A two-stage approach was used to assess possible below–toxic threshold biological activity. The first involved estimation of the toxic threshold, while the second assessed the distribution of responses at concentrations lower than the estimated threshold concentration.

Threshold Estimation Strategies

Two strategies were used to estimate the toxicological threshold: the Benchmark Dose (BMD) and the NOEL. Since the results were similar, we only present the BMD results here. The Benchmark Dose 10 (BMD(10)) is the dose at which the response is estimated to have decreased 10% below control value (Crump, 1984). The BMD(10) was selected since 10% bounded the variability of the most variable yeast strain (i.e., response SD ranged from 3.0 to 7.5% for the 13 strains). Use of a BMD based on less than 10% (e.g., 2.5–7.5%) would yield progressively higher estimates of the frequency of hormesis for all parameters estimated (see ''Results'' section, Fig. 3 and ''Supplementary Data'' section), suggesting that the current approach (i.e., BMD(10) method) would lead to an overall underestimation of hormesis frequency. An identical analysis using the BMD(2.5) for all parameters reported for the BMD(10) analysis is given within the ''Supplementary Data'' section. Our method of estimating a BMD(10) is explained below (Fig. 1). Since our goal is to classify toxicity, we did not calculate the lower bound of a confidence interval for that dose. The BMD(10) approximates the control but probably entails a low degree of toxicity. It corresponds to a dose that is slightly higher than the toxicological threshold. This suggests that a dose immediately below and very close to the BMD(10) may itself be within the toxic zone (i.e., a slightly higher concentration than the actual toxicological threshold). This would become less likely with increasing distance between the BMD(10) and the concentration below the BMD(10). For example, for agents with a BMD(10) near 3.7μM, the 1.2μM dose would be close to the toxicity threshold. In contrast, for agents with a BMD(10) approaching 100μM, the 1.2μM dose would be nearly two orders of magnitude below the toxicity threshold.

A BMD(10) was calculated for each of the 28,457 (2189 chemicals and 13 strains) dose-response experiments using the average of two replications as response. The BMD(10) was estimated through the following procedure.

1. The largest concentration with an average response below 90% is identified. Let this concentration be C_{below}, and let the associated response be R_{below}.

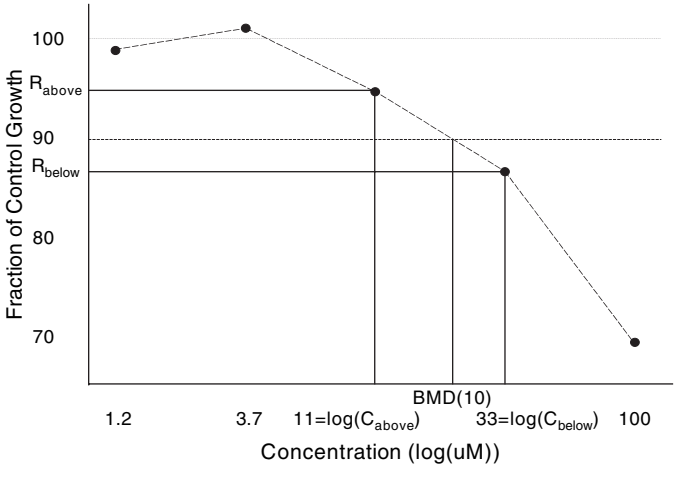

FIG. 1. General scheme used for the derivation of the BMD(10) used in the present paper.

2. If the average response at the next smallest concentration is at least 90%, then let this concentration be C_{above}, and let the associated response be R_{above}. The BMD(10) is estimated by linear interpolation on the log concentration scale:

$$\text{BMD}(10) = \exp[\log(C_{\text{above}}) + (0.90 - R_{\text{above}})(\log(C_{\text{below}}) - \log(C_{\text{above}}))/(R_{\text{below}} - R_{\text{above}})].$$

3. If the average response at the next lowest concentration below C_{below} is less than 90%, then let this concentration be C_{below} with response R_{below} and return to step 2.

If all responses for a particular chemical-strain experiment were above or below 90% then ''greater than 100μM'' or ''less than 1.2μM,'' respectively, was reported.

In contrast to a linear or nonlinear regression approach to calculating BMD, the procedure described above is ''local'' in the sense that the BMD(10) is only calculated using the responses at concentrations that are adjacent to the BMD(10). Further, when this approach is used, the two concentration-response pairs that surround the BMD(10) are chosen using only concentrations above the BMD(10) and one concentration below the BMD(10). Responses at concentrations used to estimate below-threshold responses were not used in the estimation of the BMD.

Assessing the Distribution of Responses below the Toxic Threshold

We used two approaches to assess evidence of stimulated biological activity at concentrations below the estimated threshold of toxic response. The first is a pattern analysis that counts how often both replicates in each experiment were above and below (or equal to) 100% and compares those counts to expected values, assuming a threshold model. The second approach compares the frequency of responses at various levels above and below 100%.

Pattern analysis. This approach categorized each concentration-response replication for all chemical-strain combinations and analyzed the patterns of responses at each concentration that were above and below (or equal to) 100%. In this approach, each chemical-strain repetition can express one of two different responses: H (response above 100%) and L (response less than or equal to 100%). A simple ''fair coin'' model was posited for the responses below the BMD(10) where each single replication would have a 50% chance of being above or below (or equal to) 100% and there is statistical independence across responses. This model assumes that the responses at concentrations below the BMD(10) have a median of 100%. It therefore describes a threshold model with minimal distributional assumptions. The fidelity between the observed data and this hypothesized model was tested.

Comparison of above/below-control values in the subtoxic zone of the dose-response. The threshold dose-response model predicts that responses below the toxicological threshold should randomly vary on either side (i.e., above or below) of control group values (100% response). The hormetic model predicts that there should be a nonrandom stimulatory response (i.e., responses greater than 100%) below the toxic threshold. In order to test which model best accounts for the observed data, above-control (> 100% response) to below-control (≤ 100% response) ratios were detailed for all yeast strains in the various BMD(10) classifications. The nonrandom distribution predicted by the hormesis model would be reflected in a greater frequency of responses above than below the control and in the magnitude of the deviation from the control.

Comparisons were made to responses above 100, 105, 110, 115, and 120% and then to below-the-appropriate-control group response using the formula:

$$\frac{\text{Control}}{\text{Above Response Level}} = \text{Below Response Comparison}$$

$$\left(\text{e.g.,}\frac{100\%}{120\%} = 83.33\%\right).$$

This methodology is based on the observation that the 100% control value is 83.33 of 120%. This model indicates that a 20% increase in response over the

100% value is equivalent to a 16.7% decrease from the control. Using this approach, ratios of counts comparing the following levels were made > 100%/≤ 100%, > 105%/≤ 95.24%, > 110%/≤ 90.91%, > 115%/≤ 86.96%, and > 120%/≤ 83.33%. This methodology was used to take into account the possibility of an unrestricted stimulatory response while the maximum inhibitory response was fixed at zero.

RESULTS

Figure 2 describes the concentration-response relationships of the 13 yeast strains to the 2189 chemical agents tested. While there was little change on average from the control response at the lowest concentration (1.2μM), indications of average toxicity start to become evident at 3.7μM, progressing in dose-dependent fashion over the next three concentrations (11, 33, and 100μM). Table 2 shows the number of chemicals with BMD(10) values within each of six BMD(10) classification ranges for each of the 13 strains. The more toxic chemicals are included in the low BMD(10) range (e.g., < 1.2μM), while the chemicals with the lowest toxic potential comprise the highest BMD(10) categories. The data indicate that the wild type and SPY50780 yeast strains had the lowest number (139/2189 and 143/2189) of concentration-responses with BMD(10)s < 1.2μM, indicating that they were the least susceptible strains, a perspective that is supported by plotting of the overall data in Figure 2. In contrast, strains carrying *rad50*, *rad50EPP+*, *rad18*, *rad52*, and *sgs1* were the most susceptible (Fig. 2, Table 2).

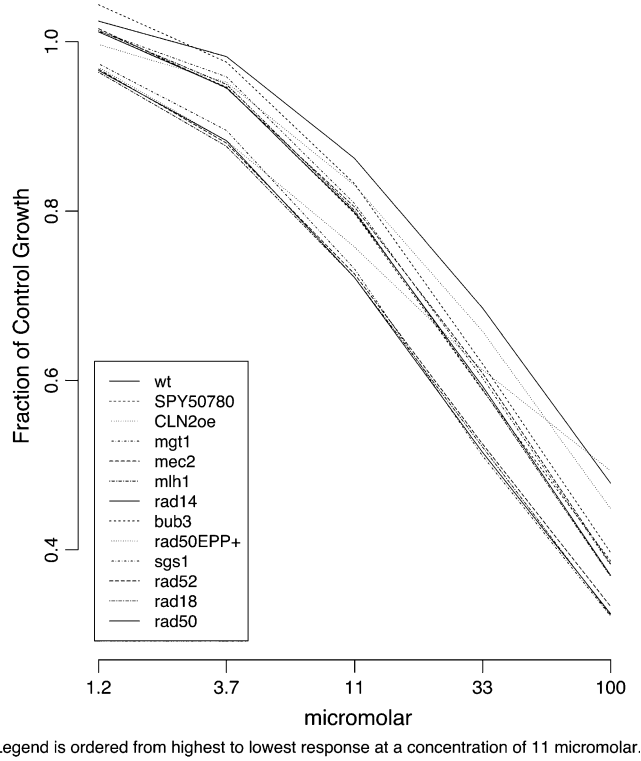

FIG. 2. Average concentration-response of 2189 chemicals on the 13 yeast strains.

Table 3 is a summary of the below-BMD(10) mean responses and SD for each of the 13 yeast strains. The numbers of chemical-concentration relationships satisfying *a priori* entry criteria using the BMD(10) methodology are shown in Table 2. Of 28,457 concentration-responses to the 2189 chemicals, 16.7% (4763) gave no evidence of toxicity, having BMD(10) values ≥ 100μM. Assessments were performed on this subgroup of responses under the assumption that at higher concentrations, a toxic response would have occurred. When the BMD(10) is less than 3.7μM, there is no concentration that can be assessed for biological activity below the BMD(10), thereby not satisfying our entry criteria. There were 7558 such responses (3798 with BMD(10) < 1.2μM and 3760 with 1.2 ≤ BMD(10) < 3.7), accounting for 26.6% of the total responses. Therefore, 73.4% of the total dose-responses were evaluated. Similar findings were observed with the NOEL methodology (data not shown).

BMD(10) Response Evaluation

We averaged responses at concentrations when the concentration was below the BMD(10) for each strain. This resulted in averages of between 196 and 572 responses, with the mean and SD (Table 3). The below-BMD(10) mean values (Table 3) are generally consistent across each of the 13 yeast strains within a specific BMD(10) classification as well as across BMD(10) classifications. However, the mean values are modestly lower in the 3.7 ≤ BMD(10) ≤ 11μM group than in the other groups ($p < 0.001$), which do not differ significantly from each other. These trends are consistent with median values as well. Consequently, all 13 strains with each of the four BMD(10) classifications had average responses significantly greater than the control ($p < 0.001$ for each of the four columns in Table 3). These findings are consistent with a nonrandom distribution of responses in the direction of the hormetic dose-response. Findings with the NOEL methodology were similar (data not shown) except that the responses were usually several percentage points higher per strain than for the BMD(10) methodology. Similarly, if a smaller BMD(2.5–7.5) cutoff point were used instead of the BMD(10), the mean responses become progressively higher as the BMD value decreases (Fig. 3 and supplementary data, Table S1).

According to the threshold dose-response model, the distribution of responses below the estimated threshold (e.g., BMD(10)) should approach a 1:1 ratio for above- and below-control values. This was assessed for each BMD(10) classification group for responses > 100/≤ 100%, > 105/≤ 95.24%, > 110/≤ 90.91%, > 115/≤ 86.96%, and > 120/≤ 83.33%. Alternatively, one could use a different model and assume the equivalency of a symmetrical response (e.g., > 120%/≤ 80% rather than > 120%/≤ 83.3%), but this paper used the prior and more conservative approach. This approach was

TABLE 2
Number of Chemicals Tested Per Yeast Strain Classified on the Basis of BMD(10)

Yeast strains	BMD(10) < 1.2	1.2 ≤ BMD(10) < 3.7	3.7 ≤ BMD(10) < 11	11 ≤ BMD(10) < 33	33 ≤ BMD(10) < 100	BMD(10) ≥ 100	Totals
Wild type	139	249	365	443	462	531	2189
SPY50780	143	253	411	536	430	416	2189
CLN2oe	236	246	379	456	428	444	2189
mgt1	218	256	408	551	380	376	2189
mec2	259	269	399	462	446	354	2189
mlh1	227	265	417	572	363	345	2189
rad14	227	285	423	550	361	343	2189
bub3	244	274	453	488	363	367	2189
rad50EPP+	414	367	405	330	196	477	2189
sgs1	435	291	454	498	241	270	2189
rad52	424	334	398	452	289	292	2189
rad18	419	321	403	464	302	280	2189
rad50	413	350	411	464	283	268	2189
Totals	3798	3760	5326	6266	4544	4763	28457

selected in order to make hormesis more difficult to detect. Table 4, A–D, indicates the distribution of responses below the BMD(10) for the chemicals in the various BMD(10) ranges. For example, the distribution of responses in the 33 ≤ BMD(10) < 100μM classification range (Table 4, C) for the 13 yeast strains is nonrandomly distributed in the direction of a hormetic response regardless of the degree of variability in the data. The findings are inconsistent with the threshold model, which predicts a ratio closely approximating 1:1. A comparison of the respective BMD(10) classification groups reveals that in the large variation comparisons (i.e., > 110%/≤ 90.91%), the proportion of stimulatory responses exceeds those on the "below" side by threefold to over 10-fold. The comparisons are generally similar among the 11 ≤ BMD(10) < 33μM, 33 ≤ BMD(10) < 100μM, and BMD(10) ≥ 100 classifications. While the 3.7 ≤ BMD(10) < 11μM classification (Table 4, A) also shows an excess of above-control values, the magnitude of the above/below differential is notably less. The most likely explanation for the reduced response in the 3.7 ≤ BMD(10) < 11μM classification is that responses at the 1.2μM concentration in these experiments may have displayed toxicity for some of the chemicals since 1.2μM is very close to

TABLE 3
Below-BMD(10) Mean Responses (%) by Yeast Strain and BMD(10) Grouping[a]

	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Yeast strains	3.7 ≤ BMD(10) < 11		11 ≤ BMD(10) < 33		33 ≤ BMD(10) < 100		BMD(10) ≥ 100	
Wild type	102.6	(12.7)	107.2	(15.1)	105.8	(12.9)	105.1	(11.1)
SPY50780	106.1	(13.6)	108.3	(15.0)	108.8	(14.7)	105.5	(11.8)
CLN2oe	101.7	(10.3)	103.7	(11.4)	104.6	(10.8)	104.8	(10.2)
mgt1	102.7	(13.1)	106.6	(14.6)	106.5	(13.9)	105.0	(11.2)
mec2	105.4	(16.8)	107.3	(16.4)	105.8	(14.4)	106.0	(14.7)
mlh1	103.8	(15.2)	107.2	(15.5)	105.9	(14.7)	104.5	(11.0)
rad14	103.9	(12.9)	107.4	(13.7)	106.5	(13.3)	106.4	(12.5)
bub3	104.8	(13.0)	106.0	(12.5)	106.8	(12.2)	106.0	(10.6)
rad50EPP+	102.2	(10.5)	105.3	(16.1)	106.3	(14.4)	107.7	(15.4)
sgs1	103.3	(11.0)	106.7	(14.4)	106.8	(14.9)	104.8	(11.4)
rad52	103.6	(12.6)	106.8	(15.3)	105.9	(13.6)	104.0	(10.8)
rad18	103.9	(12.5)	106.2	(14.2)	106.5	(14.2)	106.6	(12.2)
rad50	102.9	(12.7)	105.5	(13.8)	104.4	(14.9)	104.7	(11.4)
Overall	103.6	(13.0)	106.6	(14.5)	106.2	(13.7)	105.5	(12.1)

[a]The number of concentration-responses for each mean value is given in Table 2. The average number of concentration-responses on which a single mean value is based is 402 (196–572 range). The 3.7 ≤ BMD(10) < 11 column is based on responses at 1.2μM, the 11 ≤ BMD(10) < 33 column is based on responses at 1.2 and 3.7μM, the 33 ≤ BMD(10) < 100 column is based on responses at 1.2, 3.7, and 11μM, and the BMD(10)≥100 column is based on responses at 1.2, 3.7, 11, and 33μM.

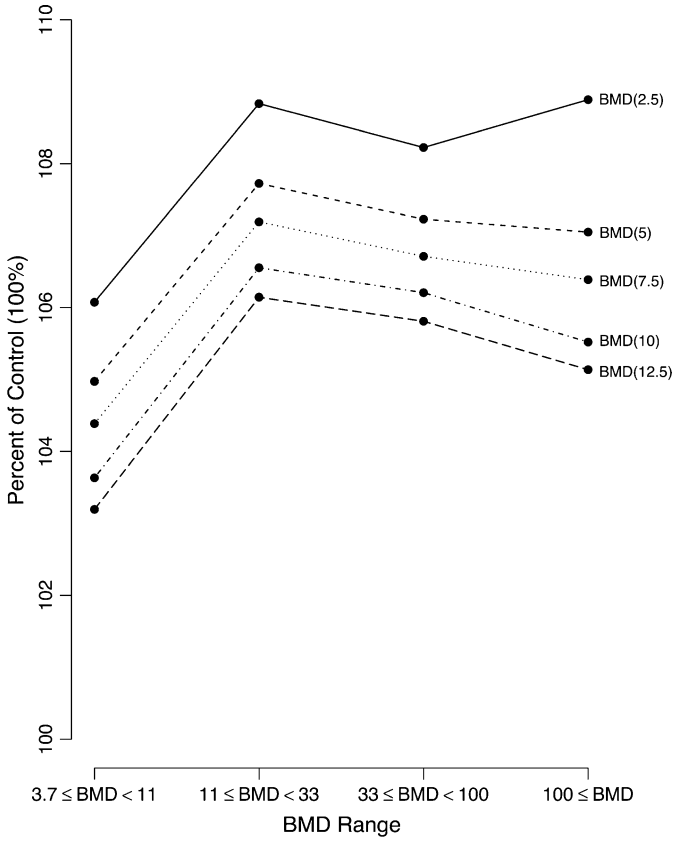

FIG. 3. BMD cutoff point influence on mean yeast growth for 13 yeast strains. (Data table for all yeast strains is found in supplementary data, Table S1).

the BMD(10) value. Figure 4 provides a simplifying summary of the information in Table 4. The comparisons in Table 4 were also used to compute weighted average estimates of overall ratios of above-control responses to below-control responses. For experiments in the $3.7 \leq \text{BMD}(10) < 11\mu\text{M}$ classification range, above-control responses were seen 1.96 times as often as below-control responses. For the $11 \leq \text{BMD}(10) < 33\mu\text{M}$, $33 \leq \text{BMD}(10) < 100\mu\text{M}$, and $\text{BMD}(10) \geq 100\mu\text{M}$ classifications, above-control responses were seen 4.64, 4.22, and 6.94 times as often as below-control responses, respectively. A weighted average calculated across all four BMD(10) classifications revealed that the above-control responses predicted by the hormesis model are 4.39 times as frequent as below-control responses. A similar assessment was performed using the BMD(2.5) (see supplementary data—Table S2A–D and Figure S1) with findings consistent with the BMD(10) analysis but even more supportive of the hormetic model.

Pattern Analysis

Patterns of response below the BMD(10) were compared for six log-spaced ranges of BMD(10)'s. Figure 5 presents these comparisons, based on observed responses that were not used in the calculation of the BMD(10), along with the expected counts under the fair coin threshold model. For instance, for BMD(10)s in the range $3.7 \leq \text{BMD} < 6.3$, the responses at 3.7μM and 11μM were used to calculate the BMD(10)s, and the figure summarizes the pattern of three responses for the replicates at 1.2μM. For the $11 \leq \text{BMD} < 33\mu\text{M}$ range, there were five possible patterns for the four replicates at 1.2μM: 4H 0L, 3H 1L, 2H 2L, 1H 3L, and 0H 4L. Under the fair coin threshold model, we would expect the fractions of responses that fit those patterns to be 1/16, 4/16, 6/16, 4/16, and 1/16, respectively. Similar procedures were used with BMD(10) values ranging from 33 to 100μM.

Strikingly, for each BMD(10) category, the observed responses markedly skew toward patterns that have more "H" (> 100%) responses, and the skewness increases as the BMD(10) increases and toxicity decreases ($p < 0.0001$). Considering the all-H patterns (left most pattern in each panel), the observed patterns are 1.3, 1.7, 5.3, 6.8, 20.7, and 25.1 times more frequent than the expected counts as the BMD(10) increases from the lowest range (3.7–6.3μM) up to the highest (57–100μM), where 57μM is halfway between 33 and 100μM on the log scale. Further, there is a strong general pattern with the BMD(10) ranges with counts tending to decrease monotonically as the number of Ls in the pattern increases. An assessment using the BMD(2.5) reveals similar findings to the BMD(10) (Fig. 5) but even more supportive of the hormetic model (supplementary data, Figure S2).

Figure 6 plots the fraction of H responses for each strain at the lowest concentration (1.2μM) as a function of the log distance below the BMD(10). As the distance of the 1.2μM concentration below the BMD(10) increases, the frequency of having both replicated responses at this first concentration (1.2μM) being greater than 100% increases to almost 60%, while only 25% would have been expected by chance assuming a threshold model. The figure shows that as the 1.2μM concentration reaches a value of ≤ 1/4th of the BMD(10), the probability of responses consistent with the hormesis model become far more common than chance for all strains; for the more highly toxic agents in the lowest BMD(10) category, the response at 1.2μM is often below control values. A similar assessment was performed using the BMD(2.5). It revealed similar findings to the BMD(10) which were even more supportive of the hormetic model (supplementary data, Figure S3).

DISCUSSION

The data indicate that responses to concentrations below the toxicological threshold for each of the 13 yeast strains tested with many hundreds of chemically diverse agents are non-randomly distributed with respect to the control. A variety of complementary methodological evaluations (Tables 3 and 4, Figs. 4–6) support the same interpretation. These findings indicate that the threshold dose-response model inadequately accounts for biological activity below the threshold. However, the results are consistent with predictions of the hormesis

TABLE 4

Evaluation of Below–Concentration Threshold Responses Based on Threshold Dose-Response Model Predictions. The Table Summarizes the Distribution of Responses Below the BMD(10) Level, Comparing the Frequencies of Levels Above and Below the 100% Control Value. The Above and Below Cutoffs that We Consider Are Comparison 1, > 100%/≤ 100%; Comparison 2, > 105%/≤ 95.2%; Comparison 3, > 110%/≤ 90.9%; Comparison 4, > 115%/≤ 87.0%; and Comparison 5, > 120%/≤ 83.3%. The Threshold Model Predicts a Ratio Closely Approximating 1:1 in Each of the Five Comparisons

	Comparison 1 (> 100%/≤ 100%)	Comparison 2 (> 105%/≤ 95.2%)	Comparison 3 (> 110%/≤ 90.9%)	Comparison 4 (> 115%/≤ 87.0%)	Comparison 5 (> 120%/≤ 83.3%)
(A) $3.7 \leq$ BMD(10) $< 11\mu M$[a]					
Observed above/below–control group ratios	0.995/1	1.70/1	3.08/1	4.78/1	6.13/1
Number of responses above the designated percentile	2656 (49.9%)	1282 (24.1%)	850 (16.0%)	627 (11.8%)	472 (8.9%)
Number of responses below the designated percentile	2670 (50.1%)	754 (14.2%)	276 (5.2%)	131 (2.5%)	77 (1.4%)
(B) $11 \leq$ BMD(10) $\leq 33\mu M$[b]					
Observed above/below–control group ratios	1.75/1	3.29/1	6.81/1	11.67/1	15.77/1
Number of responses above the designated percentile	7930 (63.3%)	4072 (32.5%)	2866 (22.9%)	2182 (17.4%)	1687 (13.5%)
Number of responses below the designated percentile	4602 (36.7%)	1239 (9.9%)	421 (3.3%)	187 (1.5%)	107 (0.8%)
(C) $33 \leq$ BMD(10) $< 100\mu M$[c]					
Observed above/below–control group ratio	2.03/1	3.78/1	6.58/1	9.85/1	11.01/1
Number of responses above the designated percentile	9133 (67.0%)	4217 (30.9%)	2755 (20.2%)	2099 (15.4%)	1596 (11.7%)
Number of responses below the designated percentile	4499 (33.0%)	1116 (8.2%)	419 (3.1%)	213 (1.6%)	145 (1.1%)
(D) BMD(10) $\geq 100\mu M$[d]					
Observed above/below–control group ratio	1.96/1	4.60/1	40.79/1	Ratio can not be calculated	Ratio can not be calculated
Number of responses above the designated percentile	12618 (66.2%)	4961 (26.0%)	3141 (16.5%)	2399 (12.6%)	1849 (9.7%)
Number of responses below the designated percentile	6434 (33.8%)	1078 (5.7%)	77 (0.4%)	0 (0.0%)	0 (0.0%)

[a]Total of 5326 responses at 1.2μM.
[b]Total of 12,532 responses; 6266 responses each at 1.2 and 3.7μM.
[c]Total of 13,632 responses; 4544 responses each at 1.2, 3.7, and 11μM.
[d]Total of 19,052 responses; 4763 responses each at 1.2, 3.7, 11, and 33μM.

dose-response model. The conclusions based on this large database in yeast are similar to those made in earlier reports of Calabrese and Baldwin (2001b, 2003) using toxicological data representative of a broad range of biological models, endpoints, and chemical agents. The evidence of hormesis in yeast, along with the previous studies, is biologically significant and has potentially important implications for essentially all drug and chemical hazard assessment studies and risk assessments worldwide.

The similar overall response patterns below the toxic zone in all 13 yeast strains for the large number of chemicals in this extensively evaluated public database is a novel finding. Previous publications (see Holbeck, 2004, for a review) with this database have focused on the nature of the above-threshold (rather than below-threshold) responses and their underlying toxicological mechanisms since the goal of the NCI has been principally oriented toward identifying possible antitumor drugs rather than assessing the nature of the dose-response in the low-dose zone.

The pattern analysis assessment indicated that the total number of concentration-responses below the BMD(10) are skewed strongly in the direction predicted by the hormesis model (Fig. 5). These findings are consistent with the earlier reports of Calabrese and Baldwin (2001b, 2003) indicating a similar relationship for data derived from the toxicological literature.

The quantitative nature of the dose-response in the hormetic zone in the present study is also consistent with findings reported for hormesis with other biological models, endpoints, and chemical agents. That is, the hormesis response is usually modest, with the maximum response typically being only 30–60% greater than the controls (Calabrese and Blain, 2005). For example, the data for the wild-type yeast strain in the $11 \leq$ BMD(10) $< 33\mu M$ and $33 \leq$ BMD(10) $< 100\mu M$ classifications indicate that the proportion of responses exceeding 120% was 26.3 and 16.5%, respectively. In the case of strain SP47080, the respective proportion of responses >120% were 26.2 and 23.7%, respectively.

The present findings illustrate the importance of study design in the assessment of hormesis. A comparison of the lowest concentration (1.2μM) to the BMD(10) revealed that the chance of a stimulatory response becomes greater as the difference between 1.2μM and the BMD(10) increases. Approximately 60% of the time, both replicate responses at 1.2μM exceeded control values for BMD(10) values between 50 and 100μM, compared to only 16% for BMD(10) values between 1.2 and 3.7μM. These findings are consistent with the

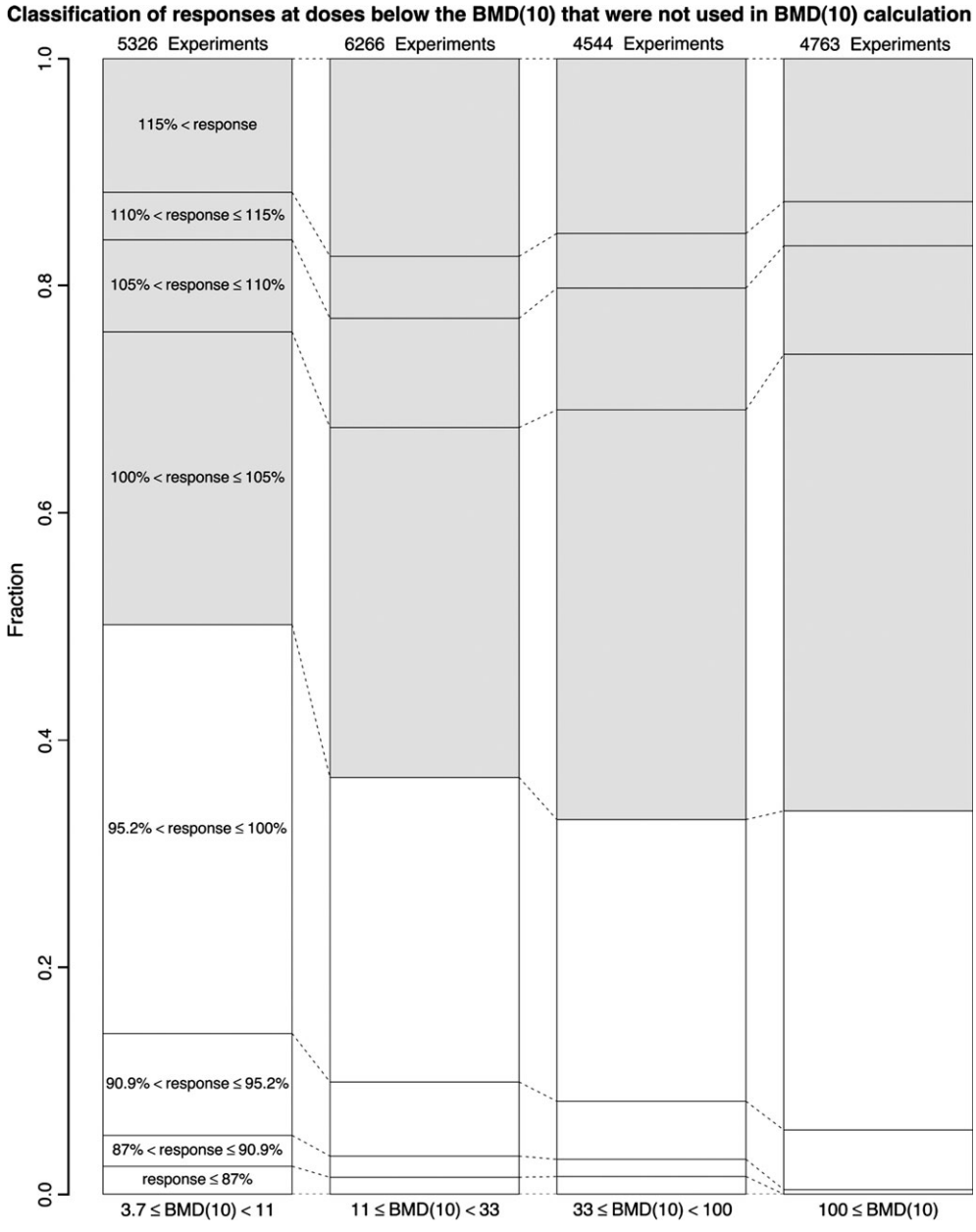

FIG. 4. Distribution of responses below the BMD(10). Shaded panels represent above-control responses, and clear panels are below-control responses.

observations of Calabrese and Baldwin (2001b) that there is an optimal range of hormetic responses starting at about 1/3–1/4 of the estimated toxic threshold. It is likely that the low response at the concentration below the BMD(10) in the $3.7 \leq$ BMD(10) $< 11\mu$M classification (Table 3) was due to there being a substantial proportion of such responses within the toxicity zone. The likelihood of a hormetic response increases as the distance from the BMD(10) increases, at least up to the limits presented in the present database. Since the semilog concentration spacing covered only a 100-fold concentration range, usually including toxicity at the high concentrations, it was not possible to explore the concentration-response range at which a return to control values would be expected. This would have required several concentrations lower than 1.2μM. Based on the hormesis database (Calabrese and Blain, 2005), about 80% of the hormetic responses are within 100-fold of the dose of the toxic threshold. Regardless of the genetic differences among the 13 yeast strains, the overall response to the 2189 chemicals was similar. These findings suggest that the hormetic response is a general one, unrelated to a specific cell cycle regulatory mechanism or DNA repair pathway. Similar quantitative features of the hormetic dose-response occur in models representing broad phylogenetic diversity and various cancer and noncancer-related endpoints (Calabrese and Baldwin,

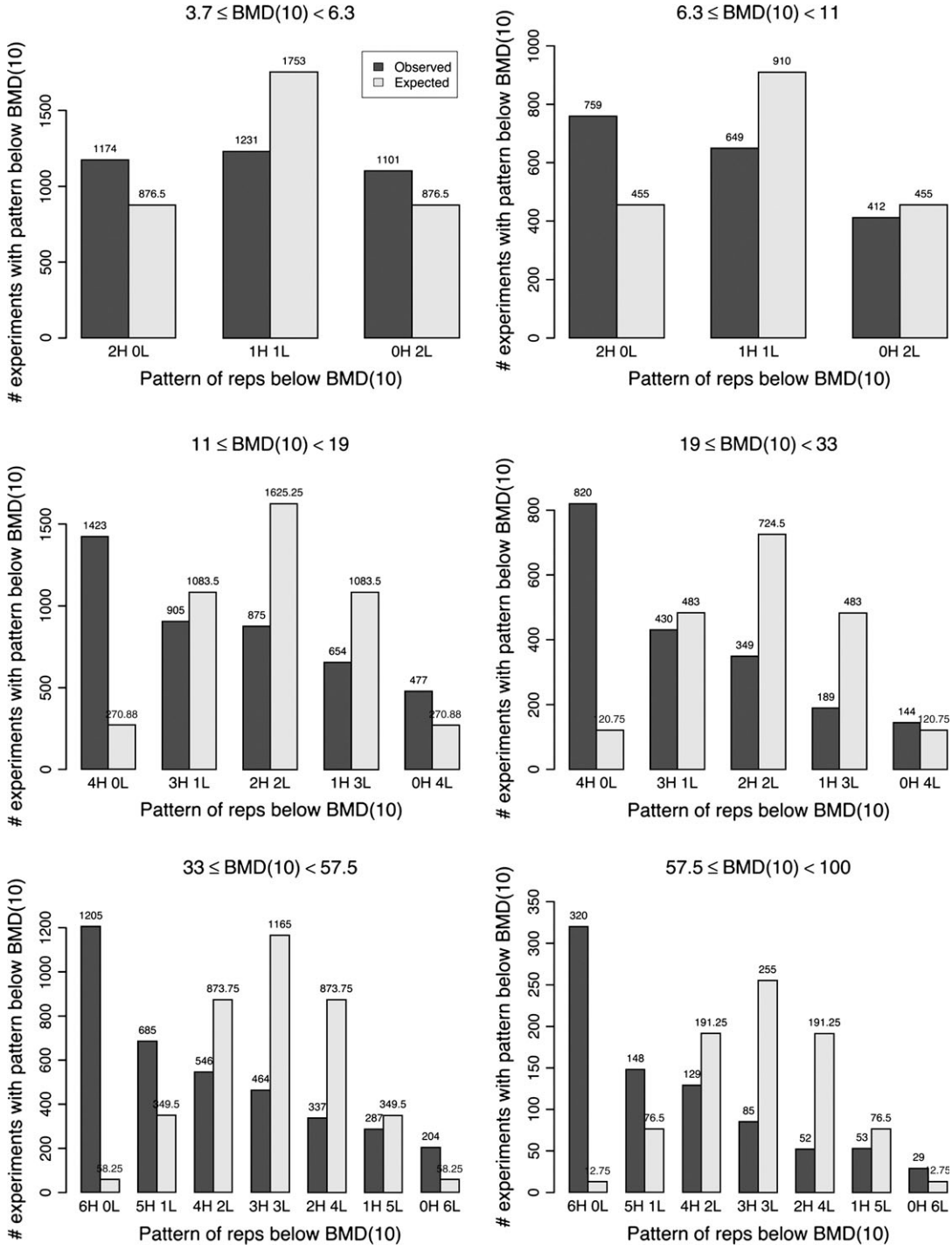

FIG. 5. Pattern analysis of below-BMD(10) responses: a test of threshold and hormetic dose-response model predictions. The levels 3.7, 6.3, 11, 19, 33, 57.5, and 100μM are approximately evenly spaced on the log scale. For example, take the panel labeled $11 \leq \text{BMD}(10) < 19$. There are responses at two concentrations below the BMD(10) (1.2 and 3.7uM) and two replications at each concentration. An experiment falls into the 3H 1L pattern if three replication responses were greater than 100%, and one was less than or equal to 100%. The dark bar is the observed count in that pattern, and the lighter bar is the expected count assuming a threshold model holds where H and L each occur with a probability of 1/2 at each replication below the BMD(10).

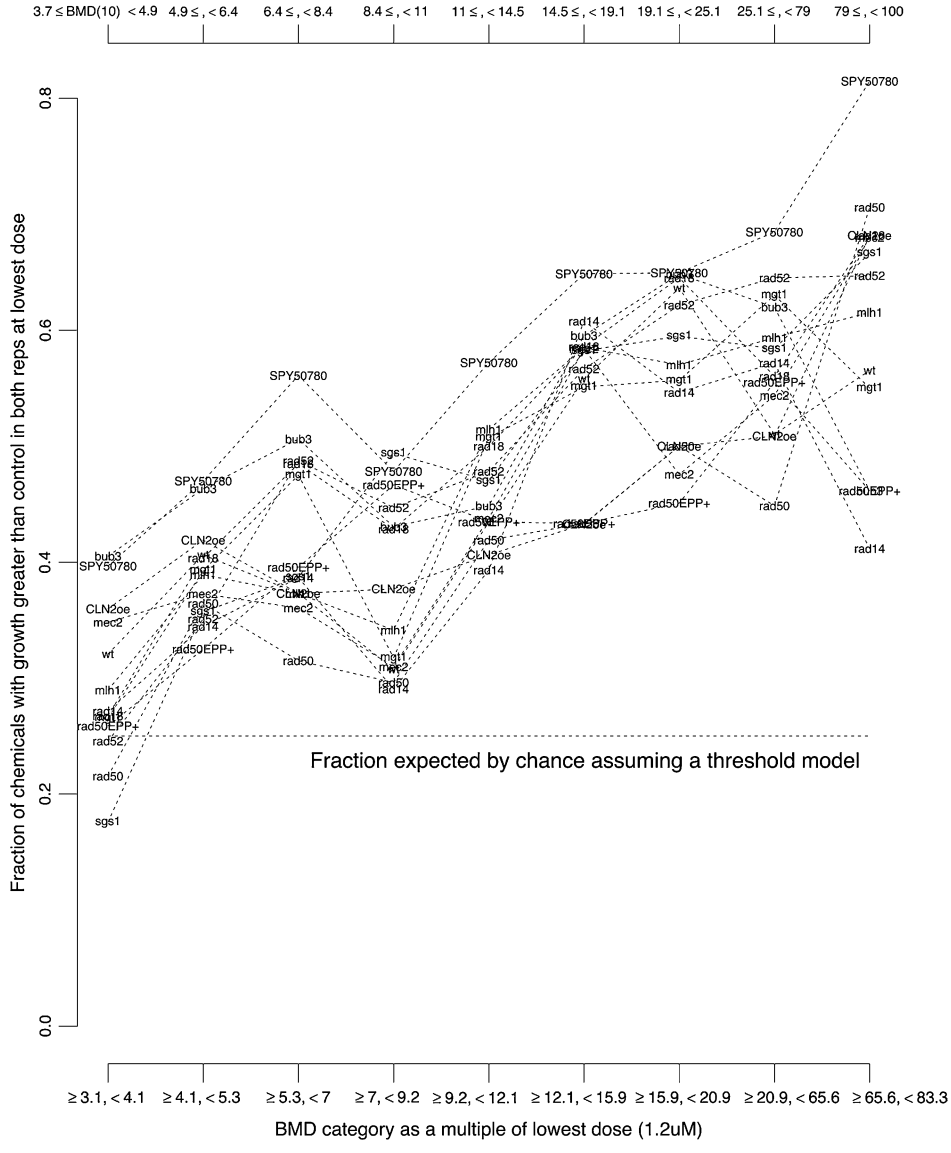

FIG. 6. Frequency of a hormetic response at 1.2μM in relation to the distance from the BMD(10).

2001b, 2003). The findings with yeast are consistent with those seen with NCI cancer drug screening data for 70 human tumor cell lines and up to 55,000 chemicals, involving over 3.3 million dose-responses (Calabrese, Staudenmayer, and Stanek, in preparation) (http://dtp.nci.nih.gov). Our present findings are drawn from the U.S. NCI database for screening of potential antitumor agents, and the consistently observed stimulation of proliferation in the below-threshold zone may have significant implications for the design of new antitumor drugs, drug testing, and the management of patients in clinical settings (Calabrese *et al.*, 2006).

The current findings are particularly important because they demonstrate the inadequacy of the traditional threshold dose-response model in predicting below-threshold responses. They also indicate that the hormetic model is consistent with these subtoxic responses. The findings suggest that the hormetic responses are more fundamental than threshold responses and support recent arguments that the hormesis model should be considered as the default dose-response model for scientific interpretation of toxicological responses (Calabrese, 2004). Several features of the study design and/or methodological evaluation (e.g., peripheral placement of controls on the 96-well plate, use of a nonsymmetrical model to assess above/below 100% responses, use of the BMD(10) instead of the NOEL or BMD's with lower cutoff points [2.5, 5.0, and 7.5]) favored conservative estimates and may have caused an underestimation of the frequency of hormesis. Thus, the inadequacies of the threshold model are probably greater than presented, while the predictive capacity of the hormesis model in the below-threshold zone exceeds that reported. The findings

argue for a paradigm shift in our understanding of the dose-response relationship, the central pillar of pharmacology and toxicology.

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

ACKNOWLEDGMENTS

Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, United States Air Force (USAF). The assistance of Dr Julian Simon, Fred C. Hutchinson Cancer Research Center, is gratefully acknowledged.

REFERENCES

Calabrese, E. J. (2004). Hormesis: From marginalization to mainstream. A case for hormesis as the default dose-response model in risk assessment. *Toxicol. Appl. Pharmacol.* **197,** 125–136.

Calabrese, E. J. (2005). Paradigm lost, paradigm found: The re-emergence of hormesis as a fundamental dose response model in the toxicological sciences. *Environ. Pollut.* **138,** 379–412.

Calabrese, E. J., and Baldwin, L. A. (2001a). Hormesis: U-shaped dose-responses and their centrality in toxicology. *Trends Pharmacol. Sci.* **22,** 285–291.

Calabrese, E. J., and Baldwin, L. A., Eds. (2001b). Introduction: Scientific foundations of hormesis. *Crit. Rev. Toxicol.* **31,** 351.

Calabrese, E. J., and Baldwin, L. A. (2001c). The frequency of U-shaped responses in the toxicological literature. *Toxicol. Sci.* **62,** 330–338.

Calabrese, E. J., and Baldwin, L. A. (2003). The hormetic dose response model is more common than the threshold model in toxicology. *Toxicol. Sci.* **71,** 246–250.

Calabrese, E. J., Baldwin, L. A., and Holland, C. (1999). Hormesis: A highly generalizable and reproducible phenomenon with important implications for risk assessment. *Risk Anal.* **19,** 261–281.

Calabrese, E. J., and Blain, R. (2005). The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: An overview. *Toxicol. Appl. Pharmacol.* **202,** 289–301.

Calabrese, E. J., Staudenmayer, J., and Stanek E. J. (2006). Drug development and hormesis. Changing conceptual understanding of the dose response creates new challenges and opportunities for more effective drugs. *Curr. Opin. Drug Discov. Devel.* **9,** 117–123.

Casella, G., and Berger, R. L. (2002). *Statistical Inference*, 2nd ed. p. 190. Duxbury, Pacific Grove, CA.

Clark, A. J. (1926). *Applied Pharmacology.* 1st ed. J & A Churchill, London.

Clark, A. J. (1933). *Mode of Action of Drugs*. Arnold, London.

Clark, A. J. (1937). *Handbook of Experimental Pharmacology.* Verlig Von Julius Springer, Berlin.

Crump, K. S. (1984). A new method for determining allowable daily intakes. *Fundam. Appl. Toxicol.* **4,** 854–871.

Eaton, D. L, and Klaassen, C. D. (2001). Principles of Toxicology. *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 6th ed. (C. D. Klaassen, Ed.). pp. 11–34. McGraw-Hill, New York.

Faessel, H. M, Levasseur, L. M., Slocum, H. K., and Greco, W. R. (1999). Parabolic growth patterns in 96-well plate cell growth experiments. *In Vitro Cell. Dev. Biol.* **35,** 270–278.

Hardman, J. G., Limbird, L. E., and Gilman, A. G. Eds. (2001). *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, 10th ed. McGraw-Hill, New York.

Hayes, A. W., Ed. (2001). *Principles and Methods of Toxicology*, 4th ed. Taylor and Francis, Philadelphia, PA.

Holbeck, S. L. (2004). Update on NCI *in vitro* drug screen utilities. *Eur. J. Cancer* **40,** 785–793.