Noname manuscript No.
(will be inserted by the editor)

A web-based tool for designing experimental studies to detect hormesis and estimate the threshold dose

Víctor Casero-Alonso · Andrey Pepelyshev · Weng K. Wong

the date of receipt and acceptance should be inserted later

Abstract Hormesis has been widely observed and debated in a variety of context in biomedicine and toxicological sciences. Detecting its presence can be an important problem with wide ranging implications. However, there is little work on constructing an efficient experiment to detect its existence or estimate the threshold dose. We use optimal design theory to develop a variety of locally optimal designs to detect hormesis, estimate the threshold dose and the zero-equivalent point (ZEP) for commonly used models in toxicology and risk assessment. To facilitate use of more efficient designs to detect hormesis, estimate threshold dose and estimate the ZEP in practice, we implement computer algorithms and create a user-friendly web site to help the biomedical researcher generate different types of optimal designs. The online tool facilitates the user to evaluate robustness properties of a selected design to various model assumptions and compare designs before implementation.

Keywords Approximate design · D-efficiency · Risk assessment · Toxicology · ZEP dose

1 Introduction

Hormesis is a special form of a dose-response relation which has been observed and discussed in many areas of life sciences. In the area of radiation alone, there is at least a monograph on hormesis (Luckey, 1991). Other examples can be found

Víctor Casero-Alonso
Mathematics Department, University of Castilla–La Mancha, Spain
E-mail: Victormanuel.casero@uclm.es

Andrey Pepelyshev
School of Mathematics, Cardiff University, Cardiff, UK, CF24 4AG
Faculty of Mathematics & Mechanics, St.Petersburg State University, St.Petersburg, Russia, 198504

Weng K. Wong
University of California at Los Angeles

in disciplines, such as biomedical or toxicological sciences; see for example, Rodricks (2003); Calabrese (2005); Cook and Calabrese (2006a); Thayer *et al.* (2006); Cook and Calabrese (2006b); Calabrese (2009); Foran (1988); Sielken and Stevenson (1998); Teeguarden *et al.* (2000). Hormesis is characterized by having beneficial effect when the mean response is stimulated at low doses and becomes inhibitory at high doses (Calabrese, 2005; Thayer *et al.*, 2006). For hormesis to exist, there is conceptually an assumed threshold and the question becomes whether such a threshold exists in the assumed model that matches reality (Cox, 1987; Slob, 1999). Such an issue was discussed as early as 1971 in Hatch (1971) and continues to today in disciplines such as aging, biology, crop growth, environmental science, food chemistry, material science, medicine, pharmaceutical sciences, radiation physics, technology; some recent examples include Vaiserman (2011); Radak *et al.* (2017); Zou *et al.* (2017); Sthijns *et al.* (2017); Abbas *et al.* (2017); Roullier-Gall *et al.* (2016); Ji *et al.* (2016).

Detecting the presence of hormesis can be an important problem with wide ranging implications (Cook and Calabrese, 2006a; Foran, 1988; Sielken and Stevenson, 1998; Teeguarden *et al.*, 2000). As the dose increases, the shapes of the mean toxicological response can vary from J-shaped to inverted U-shaped with different threshold models (Goetghebeur and Pocock, 1995; Hunt and Rai, 2005; Ulm, 1991). There is also recent research that suggests exercise, oxidants and antioxidants may change the shape of a belled shape hormetic response curve (Radak *et al.*, 2017).

An example that shows possible existence of hormesis in an aquatic toxicological experiment conducted by the US Environmental Protection Agency to identify effluents and receiving waters containing toxic materials. The whole effluent toxicity (WET) test is used to estimate the toxicity of waste water caused by many different species (Denton and Norberg-King, 1996; Lewis *et al.*, 1994). There are several endpoints to measure the aggregate toxic effect of an effluent. For many of these biological endpoints, toxicity is manifested as a reduction in the response relative to the control group. The WET testing involves multi-concentrations and includes several concentrations of effluent and a control group with a zero dose. More information about the WET testing can be found at `https://www3.epa.gov/`. Figure 1 shows a somewhat inverted U-shaped dose-response curve constructed using data set collected in Lewis *et al.* (1994) for a real study. The species in this experiment is Ceriodaphnia dubia, which is frequently used in toxicity testing of waste water treatment plant effluent water in the United States. The endpoint is a measure of reproduction given by the total number of young Ceriodaphnia dubia.

Experimental designs for studying the existence of hormesis are not well investigated at all. We could only locate a couple of references that discussed design issues for such studies and they include Hunt (2002a,b) and Hunt and Bowman (2004). One of their findings was that increasing the number of low-level doses improves the power for detecting hormetic effect and that current designs do not seem adequate to detect existence of hormesis. The work by Dette *et al.* (2011) appears to be the first technical piece of work to set up formal hypotheses to detect hormesis and estimate threshold level and construct different types of optimal designs for various purposes in such studies. In particular, a hypothesis to formally test whether hormesis exists using a model based approach is formulated in Dette *et al.* (2011).

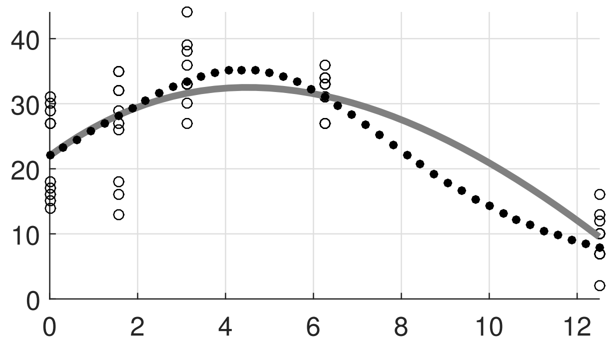

Fig. 1 Plot of the total number of young Ceriodaphnia dubia (between 0 and 45) versus a dose concentration range of $[0, 12.5]$ from a whole effluent toxicity test [circles]. The fitted curves are the extended Gompertz model with $\theta = (22, 10, 15, 0.09)$ [grey curve] and the linear logistic model with $\theta = (22, 4, 8, 4)$ [dotted curve].

A main goal of this paper is to study design issues for detecting hormesis for the two models proposed in Deng *et al.* (2001) where both models have an inverted U-shaped mean response curve. A second goal is to create a set of user-friendly codes that is freely accessible to all and allows researchers to find tailor-made optimal designs for their problems, compare competing designs and evaluate robustness properties of a selected design. In particular, the online tool allows us to readily compare efficiencies of different designs across models, including the different models proposed for studying hormesis in Dette *et al.* (2011). We expect that having an online tool is likely going to be more effective than providing computer programs to researchers in biomedicine in terms of encouraging them to explore important design issues. Our hope is that the web-based tool will facilitate researchers in using a more informed design for their studies to investigate the existence of hormesis and estimate the threshold dose.

Section 2 provides the statistical background for our model-based approach to find an optimal experiment design using theory. In Section 3, we discuss two models in Deng *et al.* (2001) and determine different types of locally optimal designs for the two models. In Section 4 we evaluate robustness properties of an optimal design to various violations in the model assumptions. This is an important task to undertake before implementing the optimal design because an optimal design can be sensitive to model assumptions and optimality criteria, see for example, Wong (1994); Moerbeek (2005). Section 5 describes our newly created website that enables users to compute and select more effective designs for detecting hormesis and the threshold value.

2 Statistical background

In this section we recall background and review theory for finding an optimal design. We present statistical models and optimality criteria based on information matrices.

2.1 Statistical models

The mean response of the extended Gompertz model is given by

$$\mu(d,\theta) = \theta_1 - \theta_2 d + \frac{\theta_3}{\theta_4}(1 - e^{-\theta_4 d}), \tag{1}$$

where θ_1 is the intercept, θ_2 is a rate constant, θ_3 is a hybridized parameter and θ_4 is a first-order rate constant. The covariate d may be the age for describing the age-specific mortality rate (Boxenbaum *et al.*, 1988; Neafsey *et al.*, 1988) or represent a dose concentration used to arrest the growth of a cancerous tumor.

An alternative model —with also 4 nonlinear parameters— is the linear-logistic model whose mean response is given by

$$\mu(d,\theta) = \frac{\theta_1 + \theta_2 d}{1 + e^{-\theta_3} d^{\theta_4}}. \tag{2}$$

This model was proposed in Vanewijk and Hoekstra (1993) where d may represent the dose concentration for testing hormesis in ecotoxology and toxicological studies. A similar model with an additional linear parameter was used in Brain and Cousens (1989) in herbicide dose-response studies.

The errors in both models are assumed to be normally and independently distributed each with mean 0 and constant variance. These assumptions are the same as those used in Dette *et al.* (2011) for designing studies to detect existence of hormesis. In both models, $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ denotes the vector of parameters in the model.

2.2 Approximate design and Information Matrix

Suppose we have resources to take a predetermined number N of observations from the study. Researchers choose several doses $d_1, d_2, \ldots, d_k$ from a given dose interval $[0, \bar{d}]$ to observe the N responses. Given a design criterion, a statistical model and a known value of $\bar{d}$, the design questions are the optimal number k and locations of these design points, along with the optimal proportion of observations w_i to take at $d_i, i = 1, \ldots, k$. We denote this generic k-point approximate design by $\xi = \{d_1, d_2, \ldots, d_k; w_1, w_2, \ldots, w_k\}$. In practice, a simple way to implement the experiment is by rounding each Nw_i to an integer N_i so that $N_1 + \ldots + N_k = N$. For more sophisticated rounding procedures, see Pukelsheim and Rieder (1992).

Approximate designs can be studied under a broad framework when the design criterion is convex or concave on the set of approximate designs. In particular, it is straightforward to use the so-called Equivalence Theorem (Kiefer and Wolfowitz, 1960), which is based on the directional derivative of the convex functional, to verify whether an approximate design is optimal among all designs on $[0, \bar{d}]$. If it is not, the theory also provides us an assessment of its proximity to the optimum, without knowing the optimum; see Section 3. Monographs on optimal design theory, such as Fedorov (1972) and Atkinson *et al.* (2007), provide details and applications of the theory to find different types of optimal designs. Wong and Lachenbruch (1996) gives a tutorial on finding optimal approximate designs for dose response studies.

Given a design ξ, the covariance matrix of the least squares estimator of θ is asymptotically proportional to the inverse of the normalized information matrix

$$M(\xi,\theta)=\sum_{i=1}^{k} w_i f(d_i,\theta)f^T(d_i,\theta),$$

where

$$f(d,\theta)=\frac{\partial\mu(d,\theta)}{\partial\theta}$$

is the vector of partial derivatives of $\mu(d,\theta)$ with respect to the model parameters. For instance, for model (1) and model (2) we have, respectively

$$f(d,\theta)=\left(1,-d,\frac{1}{\theta_4}(1-e^{-\theta_4 d}),-\frac{\theta_3}{\theta_4^2}(1-(\theta_4 d+1)e^{-\theta_4 d})\right)^T,$$

$$f(d,\theta)=\left(\frac{1}{1+e^{-\theta_3}d^{\theta_4}},\frac{d}{1+e^{-\theta_3}d^{\theta_4}},\frac{(\theta_1+\theta_2 d)e^{-\theta_3}d^{\theta_4}}{(1+e^{-\theta_3}d^{\theta_4})^2},\right.$$
$$\left.-\frac{(\theta_1+\theta_2 d)e^{-\theta_3}d^{\theta_4}\ln(d)}{(1+e^{-\theta_3}d^{\theta_4})^2}\right)^T.$$

2.3 Optimality criteria

There are different purposes in a study and the design optimality criterion should be suitably chosen. If there is interest in estimating all parameters in the mean response, an appropriate choice is the popular D-optimality criterion. In toxicology there is often interest to estimate a meaningful function of the model parameters. For instance, we may be interested to find an efficient design to estimate the zero equivalent point (ZEP) dose τ (Deng *et al.*, 2001) or a design that is most efficient for detecting existence of hormesis.

The D-optimal criterion seeks to maximize the determinant of the information matrix $M(\xi,\theta)$ over all possible designs on the dose interval of interest. Mathematically, this is the same as maximizing the log determinant of the information matrix, which is a concave function of the information matrix. For a nonlinear model, any criteria depends on some of the unknown parameters and consequently any D-optimal design that optimizes the determinant depends on the unknown parameters θ. This means that we require nominal values of the model parameters before we can compute a locally optimal design that optimizes a user-selected criterion. The nominal values typically come from expert opinions or related studies. Therefore, we denote the D-optimal design by $\xi_D^*(\theta)$. More generally, optimal designs that depend on the unknown parameter θ are called locally optimal designs. When errors are normally and independently distributed, D-optimal designs minimize the generalized variance of the estimates for all parameters and so these designs are appropriate for estimating model parameters. From Kiefer and Wolfowitz (1960), if errors have constant variances, D-optimal designs are also optimal designs for estimating the response surface across the design interval.

To define the τ-optimality design criterion for estimating the ZEP dose τ we recall the definition of the ZEP dose

$$\tau = \tau(\theta) = \max\{d \in [0, \bar{d}] : \mu(d, \theta) = \mu(0, \theta)\}.$$

Using Delta's method, the asymptotic variance of its estimate is proportional to

$$\mathrm{Var}_\tau(\xi) = b^T(\theta) M^{-1}(\xi, \theta) b(\theta),$$

where $b(\theta) = \partial\tau(\theta)/\partial\theta$ and assuming τ is differentiable with respect to θ. The locally τ-optimal design $\xi^*_\tau(\theta)$ is the approximate design that minimizes $\mathrm{Var}_\tau(\xi)$ over all other approximate designs on the given dose interval.

To detect existence of J-shaped hormesis, Dette *et al.* (2011) showed that an appropriate hypothesis to test is

$$H_0 : \mu'(0, \theta) \geq 0 \quad vs \quad H_1 : \mu'(0, \theta) < 0,$$

but for the inverted U-shaped hormesis considered here, the null hypothesis is H_0 : $\mu' \leq 0$, where $\mu'(d, \theta) = \partial\mu(d, \theta)/\partial d$ and assuming μ is differentiable with respect to d.

To maximize the power of this test, we first need an estimate of the model parameters θ; subsequently we minimize the asymptotic variance of $\mu'(d, \hat{\theta})$, which is proportional to

$$\mathrm{Var}_h(\xi) = h^T(\theta) M^{-1}(\xi, \theta) h(\theta),$$

where $h(\theta) = \partial\mu'(0, \theta)/\partial\theta$. A design $\xi^*_h(\theta)$ is a locally h-optimal design if it minimizes $\mathrm{Var}_h(\xi)$ over the set of all approximate designs on the dose interval.

The two asymptotic variances defined just above have the same functional form, $c^T(\theta) M^{-1}(\xi, \theta) c(\theta)$ for a known vector $c(\theta)$. That is, they are particular cases of the well known c-optimality criterion commonly discussed in design monographs, such as Fedorov (1972) and Silvey (1980).

3 Efficiencies and locally optimal designs

In practice, nominal values from experts or previous studies may not be accurate and it is important to evaluate how robust the optimal design is to mis-specification of the nominal values. This assessment is commonly made using the concept of the efficiency of a design. For example, suppose we have a nonlinear model with m parameters in the mean response and the vector of nominal values for the parameters is θ. If $\xi^*_D(\theta), \xi^*_\tau(\theta), \xi^*_h(\theta)$ is, respectively, the D, τ and h-optimal designs for the problem, the D-efficiency of an approximate design ξ is

$$\mathrm{Eff}_D(\xi, \theta) = \left(\frac{\det M(\xi, \theta)}{\det M(\xi^*_D(\theta), \theta)} \right)^{1/m}.$$

Similarly, the τ-efficiency of ξ is

$$\text{Eff}_\tau(\xi,\theta) = \frac{\text{Var}_\tau(\xi^*_\tau(\theta),\theta)}{\text{Var}_\tau(\xi,\theta)}$$

and the h-efficiency of ξ is

$$\text{Eff}_h(\xi,\theta) = \frac{\text{Var}_h(\xi^*_h(\theta),\theta)}{\text{Var}_h(\xi,\theta)}.$$

If the efficiency is 0.5, the design ξ has to be replicated twice to have the same criterion value as that of the optimal design. In the ideal world, we want designs with high efficiencies across different criteria, different models and reasonable changes in the nominal values.

The concept of efficiency is useful also when we do not know the optimum design. In fact when we wish to compare two designs ξ_1 and ξ_2, we compute their relative efficiency using an appropriate ratio of the values of the optimality criterion evaluated at the two designs. If the relative efficiency is close to unity, the two designs are about equally efficient. Furthermore, when the criterion is convex or concave, the equivalence theorem can also be used to confirm the optimality of any approximate design. The mathematical derivation of the equivalence theorem provides, as a side product, an efficiency lower bound for the design under investigation (Pukelsheim, 1993). This is helpful because we can check how close a design is to the optimum without knowing the optimum; if this efficiency lower bound is sufficiently high, say 99%, the practitioner may terminate the search for the optimal design, declare it as a nearly optimal design and use it in practice.

We next present locally optimal designs for the two models that we have found numerically. We also report efficiencies of the design

$$\xi_p = \{0, 1.56, 3.12, 6.25, 12.5; 1/5, \ldots, 1/5\}, \tag{3}$$

used in the WET test (Deng *et al.*, 2001) under the 3 criteria considered in the paper for different choices of nominal values of the parameter models. Practitioners can reproduce our results using the web-based tool; even more, with the available app they can evaluate the efficiency of any other design, any other values of the model parameters or a combination of both.

3.1 Locally D-optimal designs

The locally D-optimal designs are found by straightforward maximization of the determinant of the information matrix of a 4-point design using nominal values for the parameters. For approximate designs, we first search among all 4-point designs because the number of design points for a D-optimal approximate design is frequently equal to the number of the parameters in the mean response and when this is true, the weights at the design points can be shown to be equal (see Lemma on p.42 in Silvey (1980)).We may use an equivalence theorem (see White (1973) for the nonlinear models case) to verify the optimality of an approximate design under a concave functional, like D-optimality. These equivalence theorems are widely discussed in design monographs (Fedorov, 1972; Silvey, 1980). If conditions in the equivalence theorem are violated, the computed 4-point design is not optimal and we consider optimizing the information matrix over the space of all 5-point designs, and so on. The link between the number of support points and the number of parameters in the nonlinear case is not so straightforward: this observation dates back to Ford *et al.* (1992)

For model (1), we can theoretically verify that the locally D-optimal design $\xi^*(\theta)$ does not depend on θ_1, θ_2 and θ_3 and depends only on θ_4. This is because the information matrix for the model does not depend on parameters θ_1, θ_2 and θ_3 nonlinearly. In contrast, the locally D-optimal design $\xi^*(\theta)$ for model (2) depends on all parameters $\theta_1, \theta_2, \theta_3$ and θ_4.

Table 1 displays optimal design points or doses d_1, d_2, d_3 and d_4 of the equally weighted locally D-optimal design for model (1) for several nominal values of θ_4. Table 2 displays the corresponding results for model (2) for various nominal values of the parameters vector θ. The choice of the first nominal values for each model displayed in each Table comes from fitting the data in the WET test (see Figure 1). It is important to choose meaningful values of the parameters to investigate. This may not be obvious and the range of alternative values of the parameter to study can be highly model dependent. Plotting the mean response for different sets of nominal values can be helpful to arrive at a meaningful range of values in the parameters to investigate. For example, for Gompertz curve the choice of $\theta_4 = 0.15$ (maintaining the values of the other model parameters) is unacceptable because the mean response curve becomes negative when the dose levels exceed 10 units. This is problematic because the mean response is the total number of young Ceriodaphnia dubia and so it cannot take on negative values. Therefore, we chose closest values to $\theta_4 = 0.09$ for model (1). On the other hand, the range of values we chose for the model (2) seem appropriate because the mean response is not negative over the range of doses we studied.

From the two tables, we observe that the design $\xi_D^*(\theta)$ for model (1) always requires the 0 dose (placebo) and the largest admissible dose $\bar{d}(= 12.5)$. However, for model (2), the design $\xi_D^*(\theta)$ always contains the 0 dose but the largest dose may not be at $\bar{d}$. The two interior doses d_2 and d_3 in the D-optimal designs for model (1) are approximately 3 and 8.5 respectively. We note that one of the interior doses in the implemented design (3) is 3.12, which is close to one of the interior doses of the D-optimal designs. The intermediate doses of the locally D-optimal designs for model (2) are more variable; for instance, d_2 varies between 2.6 and 4.2. We also observe that the locally D-optimal design $\xi_D^*(\theta)$ for model (2) depends slightly on θ_1 or θ_2; this suggests that a slight mis-specification of the nominal values of these parameters is unlikely to cause a big drop in D-efficiencies.

The last two columns of Tables 1 and 2 show the D-efficiency of the implemented design ξ_p which is approximately 0.82 for model (1) and in the range $0.735 - 0.864$ for model (2). We report other efficiencies of this design later on and show that optimal design theory can provide us with a more efficient design for estimating parameters and also a design which is more robust to model assumptions and optimality criteria than the implemented design.

Tables 1 and 2 also show the D-efficiencies of the locally D-optimal design for the nominal value θ^0 for other nominal values of the parameters θ. We observe that,

Table 1 Locally D-optimal and h-optimal designs for model (1) for selected values of θ_4. The last two columns show D-efficiencies (first 3 rows) and h-efficiencies (last 3 rows) of ξ_p and of the locally D-optimal and h-optimal designs with a vector of mis-specified nominal values where $\theta_4^0 = 0.09$.

	design points				weights				eff	
θ_4	d_1	d_2	d_3	d_4	w_1	w_2	w_3	w_4	ξ_p	$\xi(\theta_4^0)$
D-optimal										
0.09	0	3.009	8.557	12.5	0.250	0.250	0.250	0.250	0.822	1
0.08	0	3.057	8.613	12.5	0.250	0.250	0.250	0.250	0.816	1.000
0.10	0	2.968	8.501	12.5	0.250	0.250	0.250	0.250	0.828	1.000
h-optimal										
0.09	0	2.710	8.917	12.5	0.355	0.441	0.148	0.055	0.512	1
0.08	0	2.758	8.963	12.5	0.353	0.443	0.150	0.054	0.505	0.999
0.10	0	2.675	8.859	12.5	0.354	0.444	0.147	0.056	0.520	0.999

Table 2 Locally D-optimal and h-optimal designs for model (2) for selected values of $(\theta_1, \theta_2, \theta_3, \theta_4)$ in columns 1-4. The last two columns show D-efficiencies (firsts rows) and h-efficiencies (lasts rows) of ξ_p and of the locally D-optimal and h-optimal designs with a vector of mis-specified nominal values where $\theta^0 = (22, 4, 8, 4)$.

				design points				weights				eff	
θ_1	θ_2	θ_3	θ_4	d_1	d_2	d_3	d_4	w_1	w_2	w_3	w_4	ξ_p	$\xi(\theta^0)$
D-optimal													
22	4	8	4.0	0	3.384	6.651	11.238	0.250	0.250	0.250	0.250	0.864	1
22	4	7	4.0	0	2.610	5.138	8.663	0.250	0.250	0.250	0.250	0.771	0.809
22	4	9	4.0	0	4.174	8.138	12.5	0.250	0.250	0.250	0.250	0.735	0.845
22	4	8	3.6	0	3.831	7.851	12.5	0.250	0.250	0.250	0.250	0.796	0.882
22	4	8	4.4	0	2.921	5.501	8.828	0.250	0.250	0.250	0.250	0.749	0.855
22	3	8	4.0	0	3.346	6.589	11.105	0.250	0.250	0.250	0.250	0.864	1.000
22	5	8	4.0	0	3.413	6.697	11.338	0.250	0.250	0.250	0.250	0.863	1.000
30	4	8	4.0	0	3.343	6.585	11.095	0.250	0.250	0.250	0.250	0.864	1.000
14	4	8	4.0	0	3.442	6.741	11.432	0.250	0.250	0.250	0.250	0.862	0.999
h-optimal													
22	4	8	4.0	0	3.080	6.811	12.5	0.391	0.444	0.084	0.081	0.544	1
22	4	7	4.0	0	2.384	5.295	10.769	0.388	0.439	0.078	0.095	0.609	0.336
22	4	9	4.0	0	3.809	8.286	12.5	0.395	0.453	0.094	0.058	0.378	0.789
22	4	8	3.6	0	3.475	8.033	12.5	0.385	0.449	0.103	0.063	0.439	0.886
22	4	8	4.4	0	2.685	5.646	10.722	0.397	0.444	0.072	0.087	0.614	0.603
22	3	8	4.0	0	3.044	6.760	12.5	0.390	0.442	0.084	0.084	0.555	0.999
22	5	8	4.0	0	3.109	6.847	12.5	0.393	0.445	0.083	0.079	0.535	0.999
30	4	8	4.0	0	3.042	6.754	12.5	0.389	0.442	0.084	0.084	0.556	0.998
14	4	8	4.0	0	3.137	6.881	12.5	0.394	0.447	0.082	0.077	0.527	0.997

D-optimal designs for model (1) varies when values of θ_4 are close to θ_4^0 but efficiencies remain close to 1. We note that in performing such a robustness study to ascertain sensitivities of the optimal design to nominal values, it is important to choose meaningful values of the parameters. If we evaluate the robustness of the locally D-optimal when θ_4^0 is 0.09 but the true nominal is 0.15, we obtain a D-efficiency of 0.996. While this may seem reassuring, we recall that the choice of $\theta_4 = 0.15$ is unacceptable.

Further, we observe that the D-efficiencies of $\xi(\theta^0)$ from Table 2 are close to 1 for different values of θ_1 and θ_2 when θ_3 and θ_4 are constant. This confirms that the D-optimal design depends only slightly on θ_1 and θ_2 for model (2) and are robust against small deviations of nominal values; in other words, small deviations from the real value of θ_1 or θ_2 do not seem to have an impact on the efficiency of the optimal design.

3.2 Locally τ-optimal designs

Numerical calculations show that for each model, the locally τ-optimal design $\xi^*_\tau(\theta)$ is a two-point design concentrated at the placebo dose, 0, and the ZEP dose $\tau(\theta)$ (see Table 3). Such a specific form of $\xi^*_\tau(\theta)$ was also observed for the other models in Dette *et al.* (2011).

Table 3 ZEP dose, $\tau(\theta)$, for selected values of θ and τ-efficiencies of the implemented design ξ_p.

θ_1	θ_2	θ_3	θ_4	$\tau(\theta)$	τ-eff(ξ_p)
model (1)					
–	–	–	0.09	9.714	0.374
–	–	–	0.08	10.928	0.465
–	–	–	0.10	8.742	0.361
model (2)					
22	4	8	4.0	8.153	0.453
22	4	7	4.0	5.842	0.426
22	4	9	4.0	11.379	0.490
22	4	8	3.6	11.260	0.497
22	4	8	4.4	6.370	0.460
22	3	8	4.0	7.408	0.460
22	5	8	4.0	8.783	0.444
30	4	8	4.0	7.352	0.460
14	4	8	4.0	9.479	0.432

The last column in Table 3 report the τ-efficiencies of the implemented design ξ_p. It has low τ-efficiencies, under 0.465 for model (1) and under 0.497 for model (2) for the nominal values of θ considered in this work. This may not be surprising since we are comparing the equally weighted two-point optimal designs at the placebo dose and at the ZEP dose with a design supported at 5 different doses spread over the same dose interval.

3.3 Locally h-optimal designs

Tables 1 and 2 show 4-point locally h-optimal designs $\xi^*_h(\theta)$ for the two models. We observe their optimal doses are close with those of the locally D-optimal designs. For example, for model (1) the doses for the h-optimal design when $\theta_4 = 0.09$ are 0, 2.710, 8.917 and 12.5, which are very close to the D-optimal design doses: 0, 3.009, 8.557 and 12.5 for the same nominal parameter values. From our examples, we observe that the locally h-optimal designs $\xi^*_h(\theta)$ for model (1) always include the 0 dose and the largest possible dose $\bar{d}(= 12.5)$. For model (2), the locally h-optimal design $\xi_h^*(\theta)$ always contains the 0 dose but may not include the largest admissible dose $\bar{d}$.

Unlike D-optimal designs, the h-optimal designs are not equally weighted, i.e. not every dose in the h-optimal design requires the same number or proportion of observations. For the nominal values considered, the h-optimal designs for both models require roughly 80% or more of the total observations be at the two lowest doses, d_1 and d_2. The optimal design for model (2) requires fewer than 9.5% observations at its largest dose, d_4, and the optimal design for model (1) requires fewer than 5.5% observations at its largest dose.

Tables 1 and 2 also show h-efficiencies of the implemented designs ξ_p and the h-optimal design when some of the nominal values are mis-specified and the assumed vector of the nominal values is θ^0, and more specifically, only θ_4^0 in the model (1). From both tables, the implemented design ξ_p has low h-efficiencies for all the nominal values of the tables, around 50% for model (1) and in the range 37%-61% for model (2). The design ξ_p has higher h-efficiencies than τ-efficiencies, suggesting it performs better for detecting hormesis than for estimating the ZEP dose. The tables also show ξ_p has higher D-efficiencies than h- and τ-efficiencies, implying that the implemented design ξ_p is best for estimating the model parameters among the 3 objectives.

We observe that the h-efficiencies of $\xi_h^*(\theta^0)$ seem robust with respect to mis-specification of the nominal values of the parameter θ_4 in model (1) because these efficiencies are generally very high. The same is not true for model (2). The h-efficiencies can drop to as low as 33% when nominal values of θ_3 vary. A similar but smaller effect is observed when θ_4 varies, with the h-efficiency dropping to about 60%. The h-efficiencies remain close to 1 when nominal values for the parameters θ_1 and θ_2 vary, suggesting that h-optimal designs seem to be robust to mis-specification of these parameters.

4 Robustness properties of optimal designs

In previous sections, we assumed nominal values are available and optimal designs were constructed assuming they are correct. In practice, the nominal values may be unknown or unreliable and mis-specifications of their nominal values can result in sub-optimal designs with possibly very low efficiencies. It is therefore advisable that before a design is implemented, researchers should undertake a robustness study that investigates sensitivities of the design to various model assumptions. In this section, we evaluate optimal designs for the various models when there is uncertainty in the nominal values. Ideally, we want a design that remains relatively efficient when the model is slightly mis-specified in various ways.

Tables 1 and 2 display various types of efficiencies for models (1) and (2) when some nominal values of their parameters are mis-specified. Such an investigation is always helpful because it informs us which parameters are more important to be accurately specified before constructing an optimal design. Those parameters that are less influential means that when their nominal values are slightly mis-specified, the re-

Table 4 D-efficiencies of locally D-optimal designs for model (1) with various nominal values of $\theta_4^{(1)}$ when model (2) having several nominal values of $\theta_3^{(2)}$ and $\theta_4^{(2)}$ is the true model, and vice versa.

$(\theta_3^{(2)}, \theta_4^{(2)})$ $\theta_4^{(1)}$	(8,4)	(7,4)	(9,4)	(8,3.6)	(8,4.4)
(2) as the true model					
0.09	0.8163	0.5229	0.9228	0.9348	0.5375
0.08	0.8116	0.5159	0.9245	0.9347	0.5311
0.10	0.8211	0.5298	0.9215	0.9350	0.5438
(1) as the true model					
0.09	0.7084	0.3612	0.9395	0.9513	0.3798
0.08	0.7039	0.3548	0.9414	0.9516	0.3739
0.10	0.7127	0.3675	0.9374	0.9507	0.3857

sulting optimal designs are not very different from the optimal design and, therefore, there is a slight drop in efficiency.

We next consider the design issue when there is model uncertainty. To fix ideas, suppose model (1) and model (2) are two plausible models. Figure 1 shows the mean responses from the two models for a selected choice of their model parameters. We construct and compare the optimal designs when one of the two models is postulated and is mis-specified. Rows 1-3 in Table 4 show the D-efficiencies of the optimal design for the assumed model (1) when the true model is model (2) for several sets of parameter values for both models. Rows 4-6 show the corresponding results when model (1) is the true model and the working model is (2).

Table 4 shows the D-efficiencies of the locally D-optimal designs under a mis-specified model can vary wildly from a low of about 36% in D-efficiency to a high of about 95% D-efficiency, depending on which parameters in each of the model are mis-specified. On the other hand, D-optimal designs have similar D-efficiencies for different values of the parameter $\theta_4^{(1)}$ when values of the parameters in model (2) are fixed, regardless whether model (1) is the assumed or the true model.

Analogous tables can be constructed to ascertain h-efficiencies when the model is mis-specified. In general we observe low h-efficiencies for some locally h-optimal designs when either one of these models is mis-specified when the other holds. As a matter of fact, all the h-efficiencies are lower than 80%. In particular, we observe that when the true model is model (1), the h-optimal designs for model (2) with parameters $(\theta_3^{(2)}, \theta_4^{(2)}) = (7, 4)$ have h-efficiencies around 50%. In contrast, when the true model is model (2) with parameters $(\theta_3^{(2)}, \theta_4^{(2)}) = (7, 4)$, the h-optimal designs for model (1) have very low h-efficiencies: 0.113, 0.0984 and 0.1262 when $\theta_4^{(1)}$ is 0.09, 0.08, and 0.10, respectively. For space consideration, we do not present the corresponding results in a tabular form.

We note that the above calculation is illustrative in the sense that we have somewhat arbitrarily picked nominal values for the two models in this discussion. In practice, the nominal values should be appropriately selected based on available data or from a pilot study.

5 An interactive web-based tool

Practitioners and researchers in toxicology and pharmacology may not be able to easily compute the optimal designs for their problems. To facilitate use of the proposed optimal designs, we have created a website that generates the sought optimal design for the models discussed in this paper. This website can be freely assessed through the link `http://areaestadistica.uclm.es/oed/index.php/computer-tools/`. It contains tools for finding different types of optimal designs for various scenarios. One of the tools is `OED-hormesis`, which was used to generate the optimal designs reported in this paper. Our algorithms are all based on the `R software` (R Core Team, 2015) and are available on the user-friendly interactive web app, which was created using the library `Shiny` (Chang *et al.*, 2016). The navigation bar on the app allows users to choose one or two models for comparison purposes.

To find an optimal design, the user first inputs a predefined set of design parameters for the selected model (Figure 2 top). For instance, when there is only interest in one model, the user selects the model from the given list, the dose space and the nominal values of the parameters for the model. The app also evaluates the efficiency of any generated design relative to one of the predefined designs that include the design for the WET test, the design for the toxicity study of the chemical diethylhexl phthalate (DEHP), a carcinogen, on mice given by Dette *et al.* (2011) or any desired design that the user inputs in the corresponding box.

Upon execution, the app uses the nominal values and constructs a plot of the mean response of the model in the first tab. The graph is helpful to ascertain whether the mean response has the shape that the user wants from the nominal values. In the second tab, the app displays the locally D-, h- and τ-optimal designs showed in Tables 1 and 2 (Figure 2 top). The app computes the best 4-point τ-optimal design to avoid numerical errors caused by the singularity of the information matrix when designs have fewer points than the number of parameters (see Section 3.2). The last line of the output shows various efficiencies of the user-supplied design and the estimated $\tau(\theta)$ value or ZEP dose.

The app facilitates comparison of the efficiencies of various optimal designs for the two models discussed in our paper and those studied in Dette *et al.* (2011). The tab ”Models to compare” allows the user to modify the predefined comparison (Figure 2 bottom). For particular cases, the user has to select the design interval meaningfully, choose the two models among the five available models, and carefully modify the nominal values for the model parameters. The efficiencies in Tables 1 and 2 and in Section 4 can all be obtained using this app.

6 Conclusions

Our work is an attempt to investigate a hitherto unaddressed problem in toxicology. While the phenomenon of hormesis seems to occur in various degrees across many areas in toxicology and beyond, the issue of finding an informed and efficient design to detect hormesis or accurately estimate ZEP or the threshold dose has not been

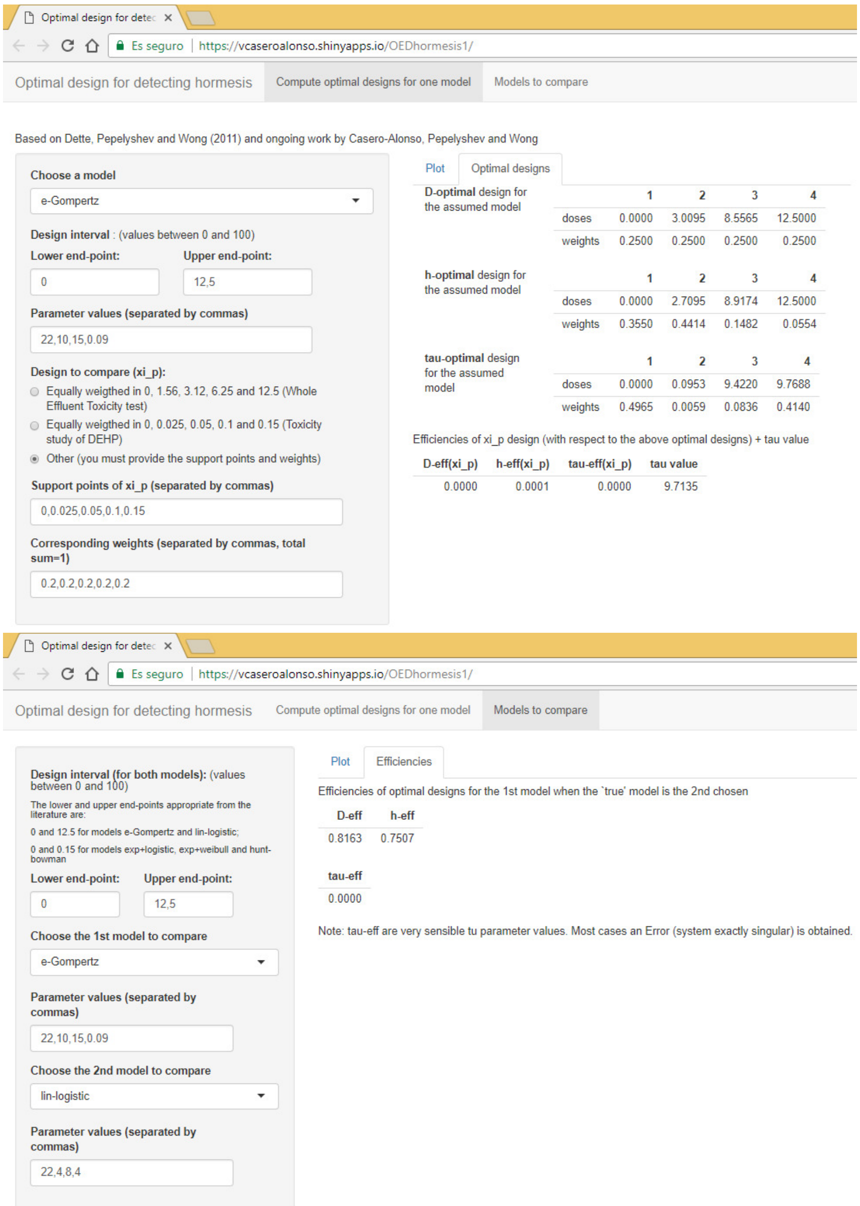

Fig. 2 Interactive web app with default values showing different types of optimal designs (top) and efficiencies of optimal designs for two models (bottom).

sufficiently addressed from the statistical viewpoint. We show that our proposed designs have advantages over currently used designs in terms of saving resources and improved precision for estimating the threshold and model parameters.

We have also shown that the design implemented in practice, ξ_p, has several doses in common (or similar) with the D-optimal or h-optimal designs for model (1). This includes the placebo dose, a dose around 3.12 (in the case of D-optimal design) and largest admissible dose. For model (2) the optimal designs have only the placebo dose in common with those of ξ_p. However ξ_p has lower efficiencies for testing the presence of hormesis than to estimate the model parameters.The performance of this design is even worst for estimating the ZEP dose when compared with the τ-optimal designs. For the nominal values considered in this work, the D-efficiency of ξ_p is higher than 73.5% for model (2) and around 82% for model (1). The h-efficiencies of ξ_p are all lower than 61% and in some cases lower than 40%. The τ-efficiencies of ξ_p are consistently unacceptably low.

Our results indicate that the proportion of observations to be taken at each dose is different from one criterion to other. For example, we have the equally weighted D-optimal designs to h-optimal designs that require less than 10% of observations at its largest dose, d_4 and more than 80% in its first and second dose, d_1 and d_2.

A limitation of our approach is that we consider models with a single independent variable. In practice, models may have two or more variables plus interaction terms. Such models will make it more difficult for us to determine the optimal designs of interest. Another limitation is that we assume that there is a single set of nominal values for the model parameters. The alert researcher should conduct a sensitivity analysis to ascertain whether the locally optimal design is sensitive to meaningful mis-specifications in the nominal values. Sometimes, in practice, there are multiple possible values for the model parameters or there are different opinions on the values of the nominal values. More sophisticated designs can be constructed using optimal design theory to incorporate the additional information. For example, if a joint distribution of the possible values of the model parameters is available, we may adopt a Bayesian approach. Usually such a joint distribution is elicited from all available information that may come from experts, pilot studies or similar studies. The construction of a Bayesian optimal design is more complicated because multiple integration is required to solve the optimization problem. Examples of Bayesian optimal designs can be found in Baek *et al.* (2006); Rodríguez *et al.* (2015); Zhu and Wong (2001).

Alternatively, we may be willing to specify an interval of possible values for each parameter and find a design that is efficient no matter which value in each of these intervals is the true value for the parameter. Such optimal designs are variously called maximin or minimax optimal designs, depending on the problem formulation. They are difficult to find because we have a nested bi-level optimization problem and techniques to search for them are beyond the scope of this paper. Some recent work in constructing minimax or maximin approach using various techniques for different types of design problems are Duarte *et al.* (2018) and Chen *et al.* (2015, 2017).

We conclude by noting that some practitioners could be interested to find an optimal design to discriminate between two or more models using the T or KL-optimality criterion. Details of this more complicated approach can be found in Atkinson and

Fedorov (1975); López-Fidalgo *et al.* (2007) and Amo-Salas *et al.* (2016) and are beyond the scope of this paper.

Acknowledgements Casero-Alonso and Wong are partially sponsored by Ministerio de Economía y Competitividad (Spain) and funds FEDER (EU), Grant MTM2016-80539-C2-1-R. Casero-Alonso is also supported by the Grant GI20174047 from UCLM. Pepelyshev is partially supported by Russian Foundation for Basic Research, project 17-01-00161-a. Wong is also partially supported by a grant from the National Institute of General Medical Sciences of the National Institutes of Health (USA) under Award Number R01GM107639. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Abbas, T., Nadeem, M. A., Tanveer, A. and Chauhan, B.S. (2017). Can hormesis of plant-released phytotoxins be used to boost and sustain crop production? *Crop Protection*, **93**, 69–76. `https://doi.org/10.1016/j.cropro.2016.11.020`

Amo-Salas, M., Delgado-Márquez, E., Filová, L. and López-Fidalgo, J. (2016). Optimal designs for model discrimination and fitting for the flow of particles *Statistical Papers*, **57**, 875–891. `https://doi.org/10.1007/s00362-016-0792-5`

Atkinson, A. C., Donev, A. N., Tobias, R. D. (2007). *Optimum Experimental Designs, with SAS*. Oxford University Press.

Atkinson, A. C. and Fedorov, V. V. (1975). The design of experiments for discriminating between two rival models *Biometrika*, **62**, 57–70 `https://doi.org/10.1093/biomet/62.1.57`

Baek, I., Zhu, W., Wu, X. and Wong, W. K. (2006). Bayesian optimal designs for a quantal dose-response study with potentially missing observations. *Journal of Biopharmaceutical Statistics*, **16**, 679–693. `https://doi.org/10.1080/10543400600860501`

Boxenbaum, H., Neaffsey, P. J., and Fournier, D. J. (1988). Hormesis, Gompertz function, and risk assessment. *Drug metabolism reviews*, **19**, 195–229. `https://doi.org/10.3109/03602538809049623`

Brain, P. and Cousens, R. (1989). An equation to describe dose responses where there is stimulation of growth at low doses. *Weed Research*, **29**, 93–96. `https://doi.org/10.1111/j.1365-3180.1989.tb00845.x`

Calabrese, E. J. (2005). Paradigm lost, paradigm found: the reemergence of hormesis as a fundamental dose response model in the toxicological sciences. *Environmental Pollution*, **138**, 378–411. `https://doi.org/10.1016/j.envpol.2004.10.001`

Calabrese, E. J. (2009). Hormesis: A Conversation with a Critic. *Environmental Health Perspectives*, **117**, 1339–1343. `https://doi.org/10.1289/ehp.0901002.`

Chang, W., Cheng, J., Allaire, J. J., Xie, Y., and McPherson, J. (2016). *Shiny: Web Application Framework for R. R package version 0.13.1*. `https://CRAN.R-project.org/package=shiny`

Chen, R. B., Chang, S. P., Wang, W., Tung, H. C., and Wong, W. K. (2015). Minimax optimal designs via particle swarm optimization methods. *Statistics and Computing*, **25**, 975–988. https://doi.org/10.1007/s11222-014-9466-0

Chen, P. Y., Chen, R. B., Tung, H. C., and Wong, W. K. (2017). Standardized maximim D-optimal designs for enzyme kinetic inhibition models. *Chemometrics and Intelligent Laboratory Systems*, **169**, 79–86. https://doi.org/10.1016/j.chemolab.2017.08.009

Cook, R., Calabrese, E. J. (2006a). The importance of hormesis to public health. *Environmental Health Perspectives*, **114**, 1631–1635. https://doi.org/10.1289/ehp.8606

Cook, R., Calabrese, E. J. (2006b). Hormesis is biology, not religion. *Environmental Health Perspectives*, **114**, A688.

Cox, C. (1987). Threshold dose-response models in toxicology. *Biometrics*, **43**, 511–523. https://doi.org/10.2307/2531991

Deng, C., Graham, R., Shukla, R. (2001). Detecting and estimating hormesis using a model-based approach. *Human and Ecological Risk Assessment*, **7**, 849–866. https://doi.org/10.1080/20018091094691

Denton, D. L. and Norberg-King, T. J. (1996). Whole effluent toxicity statistics: a regulatory perspective. In: Grothe, DR. Dickson KL, Reed-Judkins, DK (eds), *Whole effluent toxicity testing: an evaluation of methods and prediction of receiving system impacts*, 83–102. Society of Environmental Toxicology and Chemistry, Pensacola, FL, USA.

Dette, H., Pepelyshev, A., Wong, W. K. (2011). Optimal experimental design strategies for detecting hormesis. *Risk Analysis*, **31**, 1949–1960. https://doi.org/10.1111/j.1539-6924.2011.01625.x

Duarte, B. P. M., Sagnol, G. and Wong, W. K. (2018). An Algorithm Based on Semidefinite Programming for finding Minimax Optimal Designs. *Computational Statistics and Data Analysis*, **119**, 99–117. https://doi.org/10.1016/j.csda.2017.09.008

Fedorov, V. V. (1972). *Theory of optimal experiments*. Elsevier.

Foran, J. (1988). Regulatory implications of hormesis. *Human and experimental toxicology*, **17**, 441–443. https://doi.org/10.1177/096032719801700807

Ford, I., Torsney, B. and Wu, C.F.J. (1992). The use of a Canonical Form in the Construction of Locally Optimal Designs for Non-linear Problems. *Journal of the Royal Statistical Society. Series B*, **54**, 569–583.

Goetghebeur, E. J. T. and Pocock, S. (1995). Detection and estimation of J-shaped risk-response relationship. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, **158**, 107–121. https://doi.org/10.2307/2983406

Hatch, T. F. (1971). Thresholds: do they exist? *Archives of Environmental Health: An International Journal*, **22**, 687–689. https://doi.org/10.1080/00039896.1971.10665924

Hunt, D. (2002a). Simulations to improve experimental designs for u-shaped dose response modeling. *Journal of Statistical Computation and Simulation*, **72**, 737–746. https://doi.org/10.1080/00949650214266

Hunt, D. (2002b). Dose and litter allocations in the design of teratological studies to detect hormesis. *Teratology*, **66**, 309–314. https://doi.org/10.1002/tera.10106

Hunt, D. L. and Bowman, D. (2004). A parametric model for detecting hormetic effects in developmental toxicity studies. *Risk analysis*, **24**, 65–72. https://doi.org/10.1111/j.0272-4332.2004.00412.x

Hunt, D. L. and Rai, S. N. (2005). Testing threshold and hormesis in a random effects dose-response model applied to developmental toxicity data. *Biometrical Journal*, **47**, 319–328. https://doi.org/10.1002/bimj.200310129

Ji, L. L., Kang, C. and Zhang, Y. (2016). Exercise-induced hormesis and skeletal muscle health. *Free Radical Biology and Medicine*, **98**, 113–122. https://doi.org/10.1016/j.freeradbiomed.2016.02.025

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, **12**, 363–366. https://doi.org/10.4153/CJM-1960-030-4

Lewis, P. A., Klemm, D. J., Laxorchak, J. M., Norberg-King, T. J., Peltier W. H., Herber M. A. (Eds.) (1994). *Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms, 3rd edition* EPA/600/4-91/002. US Environmental Protection Agency, Cincinnati.

López-Fidalgo, J., Tommasi, Ch. and Trandafir, C. (2007). An optimal experimental design criterion for discriminating between non-Normal models. *Journal of the Royal Statistical Society, Series B*, **69**, 231–242. https://doi.org/10.1111/j.1467-9868.2007.00586.x

Luckey, T. D. (1991). *Radiation Hormesis*. CRC Press, Boca Raton, FL, USA.

Moerbeek, M. (2005). Robustness properties of A-, D-, and E-optimal designs for polynomial growth models with autocorrelated errors. *Computational Statistics and Data Analysis*, **48**, 765–778. https://doi.org/10.1016/j.csda.2004.03.014

Neafsey, P. J., Boxenbaum, H., Ciraulo, D. A. and Fournier, D. J. (1988). A Gompertz age-specific mortality rate model of aging, hormesis and toxicity: Fixed-dose studies. *Drug metabolism reviews*, **19**, 369–401. https://doi.org/10.3109/03602538808994141

Pukelsheim, F. (1993). *Optimal design of experiments*. John Wiley & Sons Inc., New York. https://doi.org/10.1137/1.9780898719109

Pukelsheim, F. and Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika*, **79**, 763–770, https://doi.org/10.1093/biomet/79.4.763

R Core Team (2015). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Radak, Z., Ishihara, K., Tekus, E., Varga, C., Posa, A.,Balogh, L., Boldoghe, I., Koltai, E. (2017). Exercise, oxidants, and antioxidants change the shape of the bell-shaped hormesis curve. *Redox Biology*, **12**, 285–290. https://doi.org/10.1016/j.redox.2017.02.015

Rodricks, J. V. (2003). Hormesis and toxicological risk assessment. *Toxicological Sciences*, **71**, 134–136. https://doi.org/10.1093/toxsci/71.2.134

Rodríguez, C., Ortiz, I. and Martínez, I. (2015). Locally and maximin optimal designs for multi-factor nonlinear models, *Statistics*, **49**, 1157–1168. https://doi.org/10.1080/02331888.2014.922562

Roullier-Gall, C., Witting, M., Moritz, F., Gil, R. B., Goffette, D., Valade, M., Schmitt-Kopplin, P., Gougeon, R. D. (2016). Natural oxygenation of Champagne wine during ageing on lees: A metabolomics picture of hormesis. *Food Chemistry*, **203**, 207–215. https://doi.org/10.1016/j.foodchem.2016.02.043

Sielken, R. L. and Stevenson, D. E. (1998). Some implications for quantitative risk assessment if hormesis exists. *Human and experimental toxicology*, **17**, 259–262. https://doi.org/10.1177/096032719801700508

Silvey, S. D. (1980). *Optimal design*. Chapman and Hall.

Slob, W. (1999). Thresholds in toxicology and risk assessment. *International Journal of Toxicology*, **18**, 259–268. https://doi.org/10.1080/109158199225413

Sthijns, M. M., Thongkam, W., Albrecht, C., Hellack, B., Bast, A.,Haenen, G. R. and Schins, R. P. (2017). Silver nanoparticles induce hormesis in A549 human epithelial cells, *Toxicology in Vitro*, **40**, 223–233. https://doi.org/10.1016/j.tiv.2017.01.010

Teeguarden, J. G., Dragan, Y., and Pitot, H. C. (2000). Hazard assessment of chemical carciogens: the impact of hormesis, *Journal of Applied Toxicology*, **20**, 113–120. https://doi.org/10.1002/(SICI)1099-1263(200003/04)20:2<113::AID-JAT641>3.0.CO;2-9

Thayer, K. A., Melnick, R., Huff, J., Burns, K., Davis D. (2006). Hormesis: a New Religion? *Environmental Health Perspectives*. **114**, A632–A633.

Ulm, K. (1991). A statistical method for assessing a threshold in epidemiological studies. *Statistics in Medicine*, **10**, 341–349. https://doi.org/10.1002/sim.4780100306

Vaiserman, A. M. (2011). Hormesis and epigenetics: Is there a link? *Ageing Research Reviews*, **10**, 413–421. https://doi.org/10.1016/j.arr.2011.01.004

Vanewijk, P. H. V. and Hoekstra, J. A. (1993). Calculation of the EC50 and its confidence interval when subtoxic stimulus is present. *Ecotoxicology and Environmental Safety*, **25**, 25–32. https://doi.org/10.1006/eesa.1993.1003

White, L.V. (1973). An Extension of the General Equivalence Theorem to Nonlinear Models. *Biometrika*, **60**, 345–348. https://doi.org/10.1093/biomet/60.2.345

Wong, W. K. (1994). Comparing robust properties of A, D, E and G-optimal designs. Computational *Statistics and Data Analysis*, **18**, 441–448. https://doi.org/10.1016/0167-9473(94)90161-9

Wong, W. K. and Lachenbruch P. A. (1996). Designing studies for dose response. *Statistics in Medicine*, **15**, 343–359. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<343::AID-SIM163>3.0.CO;2-F

Zhu, W. and Wong, W. K. (2001). Bayesian optimal designs for estimating a set of symmetrical quantiles. *Statistics in Medicine*, **20**, 123–137.

https://doi.org/10.1002/1097-0258(20010115)20:1<123::AID-SIM643>3.0.CO;2-5

Zou, X., Xiao, X., He, Y., Hu, L., Hu, C. and Huang, X. (2017). Hormetic effects of metal ions upon V. fischeri and the application of a new parameter for the quantitative assessment of hormesis, *Journal of Hazardous Materials*, **322**, 454–460. https://doi.org/10.1016/j.jhazmat.2016.09.045