Efficient experiment design and nonparametric modeling of drug interaction

Hong-Bin Fang[1], Tinghui Yu[1,2], Ming Tan[1]

[1]Division of Biostatistics, University of Maryland Greenebaum Cancer Center and Department of Epidemiology and Preventive Medicine, 10 South Pine Street, MSTF Suite 261, Baltimore, MD 21201, [2]Department of Mathematics, University of Maryland, College Park, MD 20742

TABLE OF CONTENTS

1. ABSTRACT

The design and analysis of drug combination studies continue to be an area requiring further methodological developments. Faessel *et al.* (1998) studied the joint effects of the combinations of trimetrexate (TMQ) and the GARFT inhibitor AG2034 to inhibit the growth of HCT-8 human ileocecal adenocarcinoma cells. Their experiments provide a rich data resource to validate the performance of new experimental design and analysis methods for future experiments. In this paper, we first re-analyze the same data with a nonparametric model and briefly review the experimental design used in the original paper. By comparing the analysis results, we found that the fixed ratio design and the usage of the parametric model for estimating the interaction index are based on an assumption not supported by the data. We then show how the efficiency of the experiments would be improved had the maximal power experimental design based on uniform measures been used. The usage of the proposed maximal power experimental design is further supported by simulation studies.

2. INTRODUCTION

Faessel *et al.* (1) studied the synergistic effect of the combination of trimetrexate (TMQ) and the GARFT inhibitor AG2034 to inhibit the growth of HCT-8 human ileocecal adenocarcinoma cells. Two experiments were carried out to test the interaction between the two drugs. The first experiment included $2.3\mu M$ folic acid in the medium, and the second experiment included $78\mu M$. An absorbance measurement (ranging about 0.0-2.0) was recorded for each well as the treatment effect when the growth of HCT-8 human ileocecal adenocarcinoma cells was observed. Since the methods in the two experiments are the same, this paper focuses on the experiment with lower folic acid concentration ($2.3\mu M$)

Eleven wells with five repetitions on each concentration were used for each drug alone. Faessel *et al.* (1) assumed that the single drug responses follow the Hill equatio

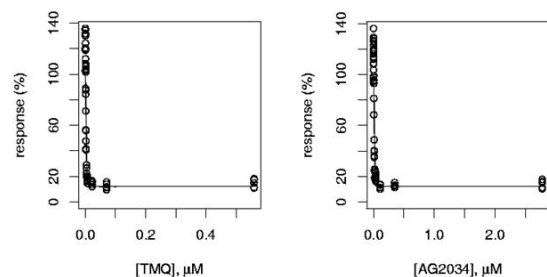$$E = \frac{(E_{con} - B)(D/ID_{50})^m}{1 + (D/ID_{50})^m} + B \quad , \tag{1}$$
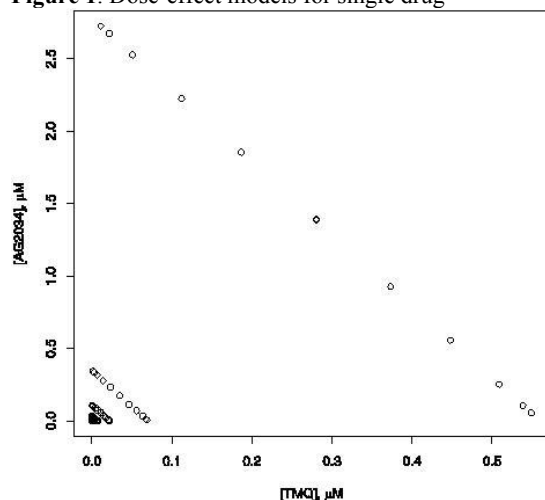
Figure 1. Dose-effect models for single drug

Figure 2. Plot of the combinations of TMQ and AG2034 based on the ray design.

where *Econ* is the control absorbance at a drug concentration of 0; *D* is the concentration of the drug used in the treatment, *B* is the extrapolated background at infinite drug concentration, ID_{50} is the median effective concentration of the drug and *m* is the slope parameter. Based only on the observations using no drugs or single drugs, the estimates are *Econ* = 1.199, *B* = 0.144 for both drugs. For TMQ alone $ID_{50,TMQ} = 1.253e{-}3$, $m_1 = -2.124$ and for AG2034 alone $ID_{50,AG2034} = 5.927e{-}3$, $m_2 = -2.953$. The fitted models and the observed data are plotted in Figure 1, with the vertical axes standing for percentages (in log scale) of the estimated *Econ* parameter.

The combination experiments of TMQ and AG2034 were then implemented according to a ray design (multiple fixed ratios) Eleven ratios between the two drugs were chosen from 1:0.004 to 1:10, with 11 dilutions in each ratio group except that 22 dilutions were administered for the ratio $d_1{:}d_2 = 1{:}5$, where d_1 and d_2 are the doses of TMQ and AG2034 in a combination, respectively. Then, a total 132 combinations were involved in the experiments. The deployment of the design points are shown in Figure 2.

Figure 2 shows that most design points are located around the origin. Specifically, more than 72% of the 132 points are in the region: $d_1 < 0.007\mu M$ and $d_2 < 0.034\mu M$. Although we may have some pharmacological basis of the two inhibitors to speculate which dose ranges

should be studied, it is not certain which dose region should be considered for the combination. This extremely unbalanced design might have put too much weight in the lower left corner of the dose region, leaving the points in the middle behaving like outliers in parameter estimation and resulting in unstable statistical analysis. Indeed, the fixed ratio design points were selected along the straight lines

$$d_2 = 5d_1 + c , \quad (2)$$

with eleven choices of intercepts *c*, where 5 is the estimated slope. We find that points corresponding to large *c* values dominate the estimation of the interaction indices. Including or excluding these points, which comprise a very small proportion of the data set, leads to very different estimated dose responses.

The fixed-ratio design has been widely employed in combination studies for its simplicity (2-4) In the fixed-ratio design of two drugs, the individual drugs are administered in amounts that keep the proportions of each drug constant. To investigate the joint action of two drugs at different ratios, Gennings *et al.* (5) proposed the ray design where a ray corresponds to one fixed ratio. The basic assumption for ray design is that synergism is a function of the proportions in the combination, i.e., one proportion may be markedly synergistic while another is simply additive (6) However, this assumption is violated in many drug combination studies, including (1), which reports the inhibition of growth of the HCT-8 cell line by TMQ and AG2034, as showed in Section 3 below. Furthermore, the potential non-constant relative potency of drugs is ignored in the fixed ratio design, which results in non-uniformly scattered combinations, and the power to detect synergism is thus undermined. Thus, the fixed ratio design may miss an apparent interaction at a particular combination design point and thus may be inadequate for combination studies of multiple drugs.

In this article, Section 3 analyzes the interaction index surface using the data from the combination study of TMQ and AG2034 to inhibit the growth of the HCT-8 cell line with a nonparametric model and demonstrates the consequences of the fixed-ratio experimental design used in the original paper. We then suggest the maximal power experimental design based on uniform measures to improve the efficiency of the experiments in Section 4. A simulation is followed to demonstrate the properties of the maximal power design. We conclude with a discussion in Section 5.

3. NONPARAMETRIC ESTIMATION OF THE INTERACTION INDEX SURFACE

In the combination study of TMQ and AG2034 against the HCT-8 cell line, 5 replicates at each of combination were administered and the whole data set consists of 660 observations with nonzero concentrations for both drugs. Faessel *et al.* (1) assumed a parametric nonlinear model for the interaction response surfac
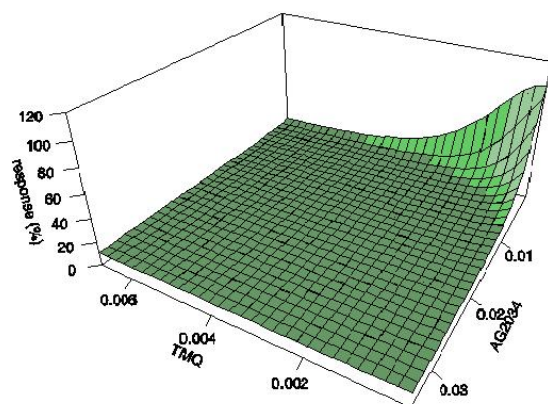
Figure 3. The estimated dose-response surface based on model (3)
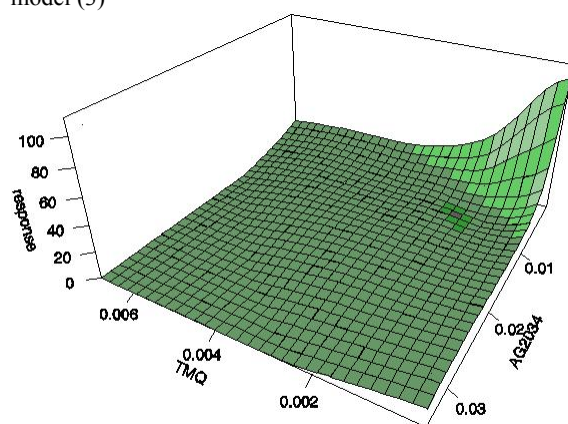
Figure 4. The estimated dose-response surface using the nonparametric thin plate spline

$$1 = \frac{d_1}{ID_{50,TMQ}\left(\frac{E-B}{E_{con}-E}\right)^{1/m_1}} + \frac{d_2}{ID_{50,AG2034}\left(\frac{E-B}{E_{con}-E}\right)^{1/m_2}} + \frac{\alpha d_1 d_2}{ID_{50,TMQ}ID_{50,AG2034}\left(\frac{E-B}{E_{con}-E}\right)^{(1/2m_1+1/2m_2)}}$$

. (3)

The synergism-antagonism parameter α is to be estimated from the observed data. The two drugs are considered additive if $\alpha = 0$, synergistic (antagonistic) if $\alpha > 0$ ($\alpha < 0$) In this case, the synergism effect is not identifiable for each combination design point (d_1, d_2) Instead, it is a label on the whole data set depending only on the estimated value (and its standard deviation) of α. In Faessel *et al.* (1), it was estimated that $\alpha = 1.5 \pm 0.25$ at 95% significance level, therefore the two drugs are significantly synergistic.

For the given estimated paramet $\hat{\alpha} = 1.5$, The surface of the predicted responses generated from equation (3) is shown in Figure 3 with 480 observations after dropping those combinations with $c > 0.1$ as defined in equation (2) To compare with the parametric modeling, we use the nonparametric (thin plate spline) estimation where no specific form of dose response

relationship is assumed (7) The fitted response E predicted with an arbitrary dose combination (d_1, d_2) is given b

$$E(d_1, d_2) = q_1 d_1 + q_2 d_2 + q_3 + \sum_{i=1}^{K} \phi((d_1 - d_{1,i})^2 + (d_2 - d_{2,i})^2) c_i$$

,   (4)

on the choice of the kernel function ϕ and the coefficients q_1, q_2, q_3, c_1, …, c_K admitting a best fit on the K observations $(d_{1,i}, d_{2,i}, E_i)$, $i = 1, …, K$. Ruppert, Wand and Carroll (8) showed that this estimate can also be derived from the mixed-effect model, hence leading to a convenient formulation of the surface estimation. Given the same 480 observations in the form of $(d_{1,i}, d_{2,i}, E_i)$, the fitted response surface is shown in Figure 4. It is worth noting that the surface in Figure 4 asymptotes (or actually reaches zero); whereas, the parametric surface in Figure 3 does not. Then, there is a difference between the estimated dose-response surface based on the parametric approach and that based on non-parametric approach. Since there is a lack of rigorous goodness of fit test for the nonparametric and parametric models, we calculated that the ratio of the mean squares of the model to the residual mean squares for the parametric model is 2.35 folds greater than that for the nonparametric model while the mean square error for the parametric model is 10% greater than that for the nonparametric model.

Although the nonparametric response surface provides more a complete description on the joint action (9-10), the approach does not give any summary measure on drug interaction. The determination of drug interaction was based on the visualization of whether the contours of the response surface were concave up or down (11) Prichard and Shipman (12) proposed using the differences of the theoretical additive surface and the response surface to reveal regions of synergy and antagonism. However, the theoretical additive surface (12) ignores experimental variations of the single drug dose-effect curves and the assessment of drug interaction does not take into account the experimental variation in the combination study (e.g., no confidence surface is given)   Thus, we shall use the three-dimensional interaction index surface and its confidence region to assess drug interactions. Berenbaum (13) defined the interaction index

$$\tau = \frac{d_1}{D_{E,1}} + \frac{d_2}{D_{E,2}},   (5)$$

where $D_{E,1}$ and $D_{E,2}$ denote the doses of TMQ and AG2034 alone yielding the same effect E, respectively, and (d_1, d_2) is the combination of concentrations of TMQ and AG2034 that yields the same effect E. If $\tau = 1$, we say that the two drugs are additive, in Loewe's sense, at the combination (d_1, d_2) If $\tau < 1$ (>1), the two drugs are synergistic (antagonistic) at the combination (d_1, d_2) Note that τ is a function of (d_1, d_2), say $\tau = h(d_1, d_2)$ Usually τ varies at different combinations.

Since the parameters in equation (1) are estimated for both TMQ and AG2034, for a given effect
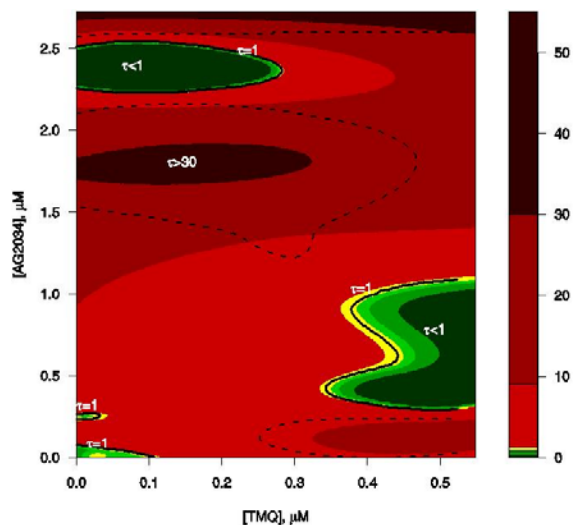
Figure 5. The nonparametric interaction index surface estimated based on the complete dataset
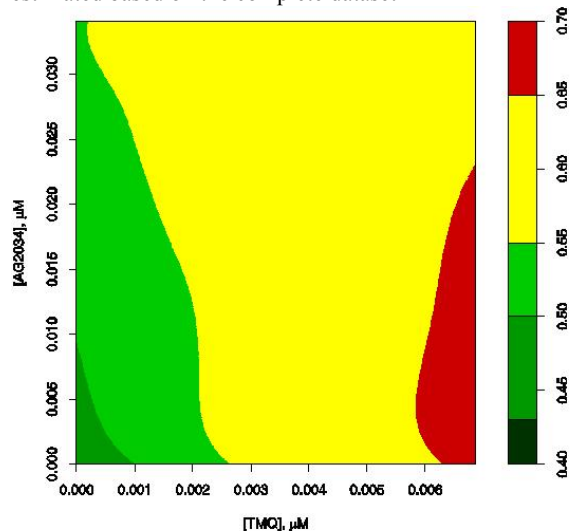
Figure 6. Close-up look of the interaction index surface in the small region of TMQ $< 0.007 \mu M$ and AG2034 $< 0.034 \mu M$ in Figure 5.

level E, the dose D_E yielding the effect level E is obtained by

$$D_E = \left(\frac{E - B}{E_{con} - E} \right)^{1/m} ID_{50} . \quad (6)$$

Thus, using equation (5), we can calculate the interaction index τ_i for each combination $(d_{1,i}, d_{2,i})$ according to the observation in the form $(d_{1,i}, d_{2,i}, E_i)$, $i = 1, \ldots, 660$. Fang *et al.* (14) developed a method to directly estimate the interaction index surface nonparametrically using thin plate splines. Sühnel (15) also discussed a similar concept of response surface for interaction.

We first use the method proposed by Fang *et al.* (14) to estimate the interaction index surface $h (d_1, d_2)$

based on the complete dataset of 660 observations. The contour plot of the fitted surface is shown in Figure 5. Green regions in this contour plot indicate synergism, and red regions indicate antagonism between the two drugs. The darkness of the colors corresponds to the intensity of the interactions. The yellow regions consist of those dose combinations with index values around 1. Particularly the solid curves in the yellow region indicate the combinations with exact additivity (say, $\tau = 1$), and the dotted curves outline the 95% confidence intervals (regions) of additivity. Remarkably the estimated indices on the 660 observed combinations have a maximum of 12.71, while the maximum value of the fitted surface is about 55.47. This discrepancy combined with the observation of extremely large confidence regions suggests that the unbalanced experimental design results in large variations in the interaction index surface estimations.

This observation is consistent with the model defined in equation (4) The fitted surface is smoothed out by a kernel function applied to the Euclidean distances between the design points $(d_{1,i}, d_{2,i})$ and the dose combination (d_1, d_2) If there are some design points lying far away from the others, the spline estimates around these outliers will be completely dominated by the closest design point. The errors (variations) in this observation will be inherited by all the points in the neighborhood of the outlier and never be diminished with the increasing sample size. In this case the interaction between very large doses is not very interesting since they are so sparsely deployed. Figure 6 thus shows the amplified version of Figure 5 in the smaller dose region of interest.

The data generated according to the fixed-ratio experimental design manifested a significant negative effect on the estimation of interaction index surface. If we drop those doses with $c = 2.8$ as defined in equation (2), the fitted interaction index surface from the 600 remaining observations is very different, as shown in Figure 7. The maximal absolute difference between the two fitted surfaces shown in Figure 6 and Figure 7 is 0.2456, while the biggest index value in Figure 6 is 0.6165. Even further, dropping two lines of large doses with $c = 2.8$ and $c = 0.34$, the index surface generated from the 540 remaining observations is shown in Figure 8. The maximal absolute difference between the surface from the complete data set (660 observations) and that from 540 observations is 0.3461. Dropping those highly scattered design points of high doses, we have a more accurate nonparametric estimation of the interaction index surface. Hence we can proceed to have a posterior investigation on the validity of the original experimental design.

In Faessel et al. (1), it was assumed that the interaction index values are constants along the straight lines with fixed ratios. Without this assumption, the design along very few straight lines is groundless. If the fitted response surface is a good approximate to the unknown true model, a test can be readily derived to assess if this assumption is valid. Figure 9 shows the nonparametric
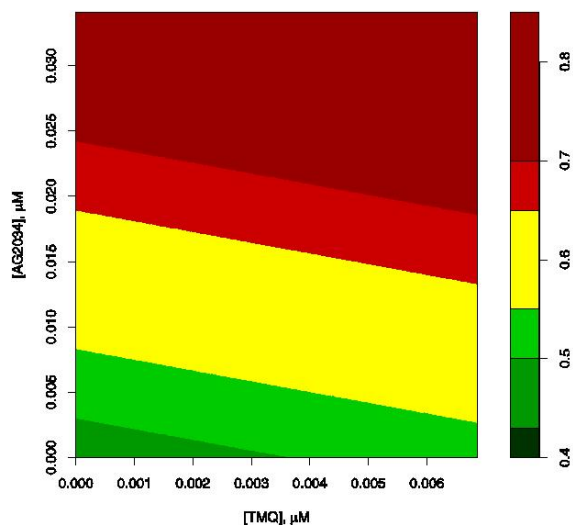
Figure 7. The nonparametric interaction index surface estimated based on 600 observations.
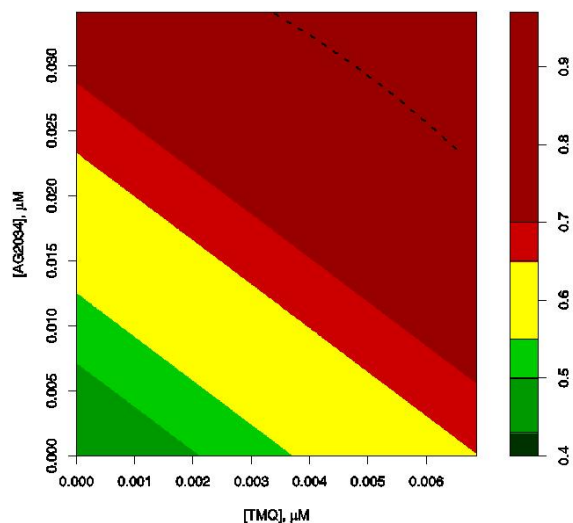
Figure 8. The nonparametric interaction index surface estimated based on 540 observations.

interaction index values along certain fixed dose ratios d_1:d_2 versus the corresponding TMQ doses.

It is easy to see from these plots that the interaction indices are far away from constant functions. An anonymous referee also brought to the authors' attention that the index is supposedly close to one when virtually only one of the two drugs is administered (i.e. when the ratio d_1:d_2 is extremely large or small and the dose effect is virtually contributed by one drug only) However, the thin plate splines model is highly affected by the unbalanced deployment of the design points. When we are only focusing on reanalyzing the original data and no new experiments can be carried out, we can simply point out that the experiment results does not support the assumption needed for the fixed ratio design. New data is needed to control the effect of the variations in each observation if

one wishes to have the fitted response surface as a better approximate to the true shape of the underlying model. This observation makes the design of fixed ratios open to questions. We may ask obviously, if there is a more efficient design that does not depend on the assumption of fixed ratio effects.

4. PERFORMANCE OF THE MAXIMAL POWER DESIGN

We pointed out in section 3 that the experimental design used in the original paper is not only very inefficient but also resulted in unstable model estimation. In this section, we propose a new experimental design based on uniform measures. Based on the dose-responses of the single drugs, Tan *et al.* (16) and Fang *et al.* (14) proposed the maximal power design for combination studies. Recently, Shiozawa *et al.* (17) utilized the methodology and computer program to identify synergistic combinations of suberoylanilide hydroxamic acid combined with cytosine arabinoside and etoposide for treating acute leukemias. This design has been shown to maximize the minimum power of the statistical test (14, 16) to detect departures from additive action, and at the same time minimize the maximum bias due to lack of fit among all potential departures of a given meaningful magnitude.

Given that the ultimate goal of a preclinical combination study is to translate its results to the clinic, we focus on studying those doses yielding certain levels of effect (*e.g.*, 20% -80% treatment effects) Based on the estimated dose-response model for single drugs, equation (1), the dose ranges are: from $7.043e{-}4$ to $3.709e{-}3$ (μM) for TMQ and from $3.916e{-}3$ to $1.294e{-}2$ (μM) for AG2034. In these dose ranges, the single dose-responses can be fitted by the log-linear models

$$E = -207.704 - 39.806\log(D_1) \quad \text{and}$$
$$E = -237.47 - 57.07\log(D_2) \quad (7)$$

for TMQ and AG2034, respectively, where E is $100 \times$ viability (% of control), and D_1 and D_2 are the doses of TMQ and AG2034 respectively. The potency of AG2034 relative to TMQ is $\rho(D_2) = 2.11255 D_2^{0.4337}$, which is non-constant and depends on doses. The predicted additive model at the combination (d_1, d_2) is

$$E(d_1, d_2) = -207.704 - 39.806\log(d_1 + \psi(d_1, d_2)d_2) \quad (8)$$

where ψ (d_1, d_2) is determined by $\psi(d_1, d_2) = 2.11255(\psi^{-1}(d_1, d_2)d_1 + d_2)^{0.4337}$.

To obtain the maximal power design for testing the joint action of TMQ and AG2034, the dose range is chosen such that the dose effect (the viability) is from 20% to 80% for TMQ. Then, the total dose ranges from $7.043e{-}4$ to $3.709e{-}3$ (μM) in TMQ. The pooled variance from the two single drug experiments is 1577.806. For a meaningful difference η of 20% (the viability) and 5
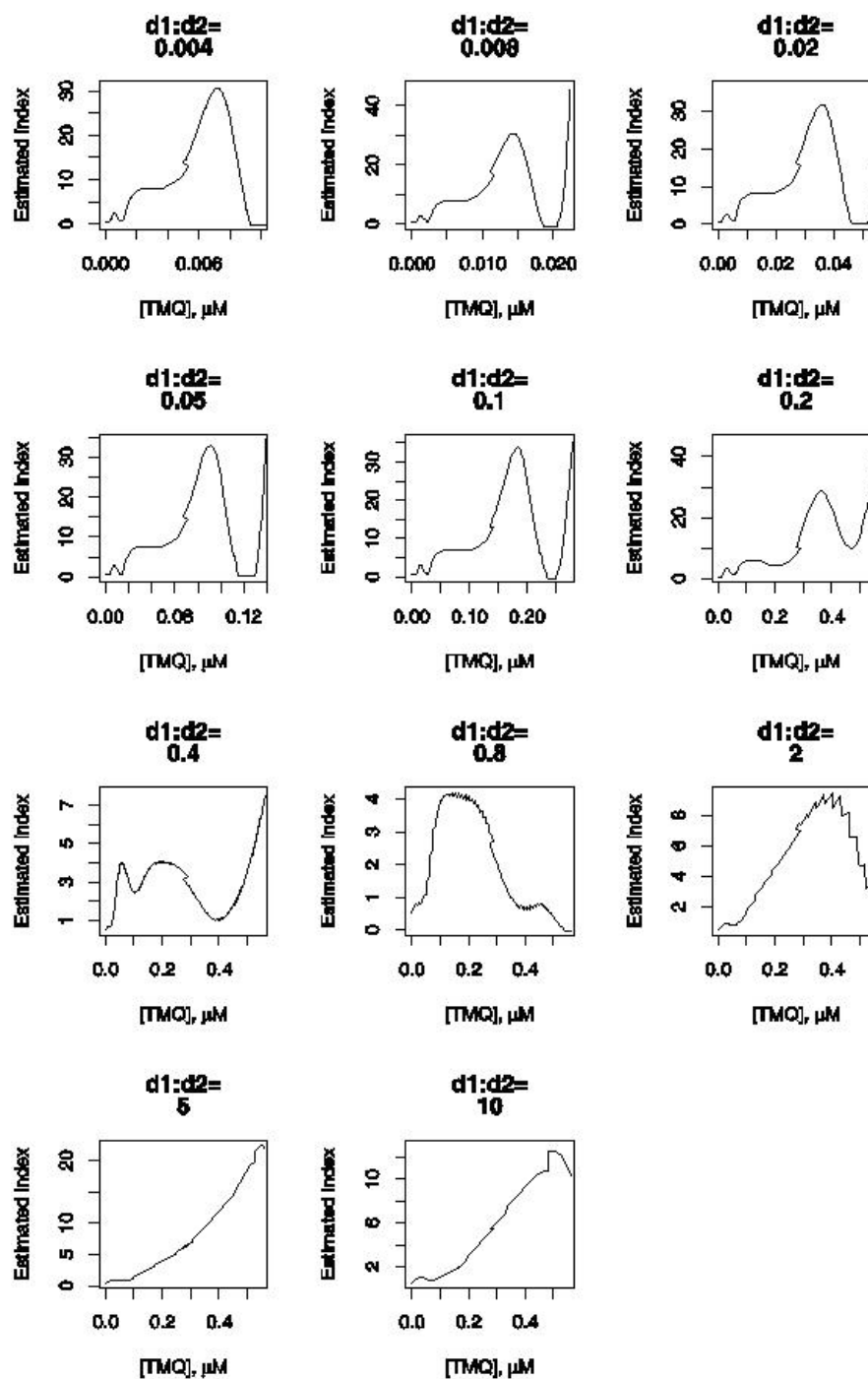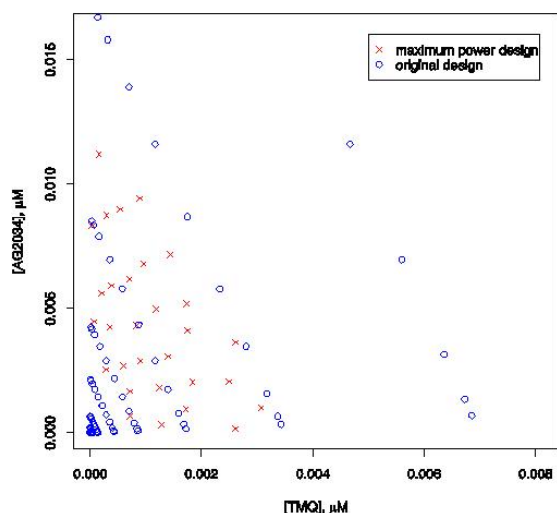
Figure 9. Estimated nonparametric interaction index curves along the lines of fixed ratios.

replications for each mixture, with type I error rate 0.05 and power 0.80, we need to study 19 combinations in the experiment in order to detect synergy/antagonism in the combination of TMQ and AG2034. Figure 10 shows the 19 combinations following the algorithms provided in (14) Here the red crosses are the proposed new design points (concentrations) and the blue circles are those used in the original experiment.

Although the power to detect synergism is maximized and proved statistically, we further demonstrate the efficiency of the proposed design by a simulation experiment. Since the dose effects from 20% to 80% are of interests, we will focus on the total dose ranges from $7.043e{-}4$ to $3.709e{-}3$ (μM) in TMQ. We assume that the ``true'' model in this dose range is given by the nonparametric thin plate spline estimation (Figure 4). With
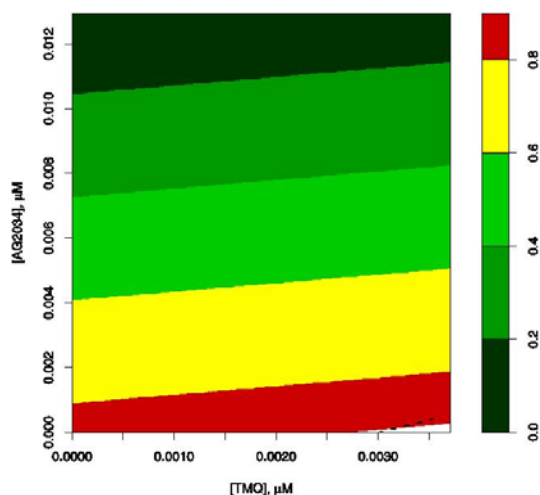
Figure 10. The plot of 19 combination design points based on the maximal power design.
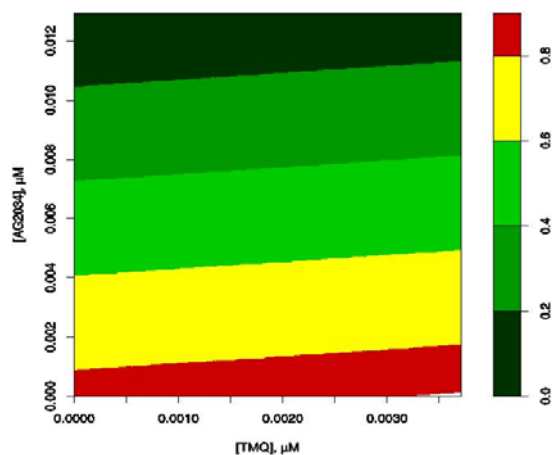
Figure 11. The nonparametric interaction index surface estimated based on the maximal power design.

Figure 12. The nonparametric interaction index surface estimated based on 160 observations by the ray design.

the estimated response surface and standard deviations, we generate 5 response values at each combination and a set of 95 response values. Using the method proposed by Fang *et al.* (14), we obtain a new estimation of the interaction index surface and the corresponding contour plot of this index surface is shown in Figure 11. Within the experimental domain that consists of the polygon with the vertices (0, $3.916e{-}3$), (0, $1.294e{-}2$), ($7.043e{-}4$, 0) and ($3.709e{-}3$, 0), there are only 160 observations, about 24% of the original design, falling into this region of practical importance. Figure 12 shows the contour plot of the estimated interaction index surface based on those 160 observations. The maximal absolute difference between the two surfaces in Figures 11 and 12 is 0.008675. The maximal power design with only 95 observations captures almost all the information contained in the experiment of 160 observations in the original design. Thus, the maximal power design is much (40%) more efficient than the fixed-ratio design, which is based on the assumption of constant interaction index values along straight lines of fixed ratios.

5. DISCUSSION

The study of the joint action of drugs has become important in drug development due to its potential to increase therapeutic index. The goal is to find which combinations are additive, synergistic or antagonistic. To assess drug interaction, an efficient and effective design for combination experiments is very important to obtain more information on the dose-response surface. Through the re-analysis of the trimetrexate (TMQ) and the GARFT inhibitor AG2034 combination, we found that data obtained based on the original experimental design resulted in unstable analysis of the interaction and reduced reproducibility.

We took advantages of recent developments in the design of combination studies, in particular, the maximal power design which is derived by means of uniform measures based on a general nonparametric statistical framework where the joint effect model does not depend on knowing the true parametric form. This design maximizes the minimum power of the statistical test to detect any departures from additive action, and at the same time minimizes the maximum bias due to lack of fit among all potential departures of a given meaningful magnitude. We generated the design and obtained the data from the best fitted dose-response surface. Since the join effect model used in the design is nonparametric, the design does not depend on knowing the true parameters describing the interaction in the model. Subsequent data analysis of the data generated by the maximal power design demonstrates the necessity of the design in estimating the dose-response adequately and the power to detect synergism. The simulation study based on Faessel's dataset further verifies the optimality of the proposed design.

6. ACKNOWLEDGEMENT

7. REFERENCES

1. H. M. Faessel, H. K. Slocum, R. C. Jackson, T. J. Boritzki, Y. M. Rustum, M. G. Nair, and W. R. Greco: Super *in vitro* synergy between inhibitors of dihydrofolate reductase and inhibitors of other folate-requiring enzymes: the critical role of polyglutamylation. *Cancer Research* 58, 3036-3050 (1998)

2. P. K. Gessner: The isobolographic method applied to drug interactions. In *Drug Interactions*, ed. By P. L. Morselli, S. Garattini, and S. N. Cohen, pp. 349-362, Raven Press, New York (1974)

3. T. C. Chou and P. Talalay: Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv. Enzyme Regul.* 22, 27-55 (1984)

4. R. J. Tallarida: Statistical analysis of drug combinations for synergisim. *Pain* 49, 93-97 (1992)

5. C. Gennings, W. H. Jr. Carter, E. W. Carney, G. D. Charles, B. B. Gollapudi, and R. A. Carchman: A novel flexible approach for evaluating fixed ratio mixtures of full and partial agonists. *Toxicol. Sci.* 80, 134-150 (2004)

6. R. J. Tallarida: *Drug Synergism and Dose-Effect Data Analysis*. Chapman and Hall/CRC, New York (2000)

7. P. J. Green and B. W. Silverman*: Nonparametric Regression and Generalized Linear Models*. Chapman & Hall, London (1994)

8. D. Ruppert, M. P. Wand, and R. J. Carroll: *Semiparametric Regression*. Cambridge University Press, Cambridge, United Kingdom (2003)

9. W. R. Greco, G. Bravo, and J. C. Parsons: The search for synergy: a critical review from a response surface perspective. *Pharmcological Reviews* 47, 331-385 (1995)

10. M. Kong and J. J. Lee: A generalized response surface model with varying relative potency for assessing drug interactions. *Biometrics* 62, 986-995 (2006)

11. J. Sühnel: Evaluation of synergism or antagonism for the combined action of antiviral agents. *Antiviral Research* 13, 23-40 (1990)

12. M. N. Prichard and C. Jr. Shipman: A three-dimensional model to analyze drug-drug interactions. *Antiviral Research* 14, 181-206 (1990)

13. M. C. Berenbaum: Synergy, additivism and antagonism in immunosuppression. *Clin. Exp. Immunol.* 38, 171-179 (1977)

14. H. B. Fang, D. D. Ross, E. Sausville, and M. Tan: Experimental design and interaction analysis of combination studies of drugs with log-linear dose responses. *Statistics in Medicine* 27, 3071-3083 (2008)

15. J. Sühnel: Zero interaction response surfaces, interaction functions and difference response surfaces for combinations of biologically active agents. *Arzneimittel Forschung* 42 (10), 1251-1258 (1992)

16. M. Tan, H.B. Fang, G. L. Tian, and P. J. Houghton: Experimental design and sample size determination for testing synergy in drug combination studies based on uniform measures. *Statistics in Medicine* 22, 2091-2100 (2003)

17. K. Shiozawa, T. Nakanishi, M Tan, H.B. Fang, W.C. Wang, M.J. Edelman, D. Carlton, I. Gojo, E.A. Sausville, D.D. Ross. Preclinical Studies of Vorinostat (Suberoylanilide Hydroxamic Acid) Combined with Cytosine Arabinoside and Etoposide for Treatment of Acute Leukemias. *Clin Cancer Res*, 15:1698-1707 (2009)

Send correspondence to: Ming Tan, Division of Biostatistics, University of Maryland Greenebaum Cancer Center, 10 South Pine Street, MSTF Suite 261, Baltimore, MD 21201, Tel: 410-706-8506, Fax: 410-706-8548, E-mail: mttan@som.umaryland.edu