Experimental Design for In Vitro Drug Combination Studies

Gregory Hather[1], Huaihou Chen[2], Ray Liu[1]
[1]Millennium: The Takeda Oncology Company, 35 Landsdowne Street, Cambridge, MA 02139
[2]New York University, 1 Park Avenue, New York, NY, 10016

Abstract
In vitro drug combination studies typically involve a large number of wells with various concentrations of two drugs added together. To gain the most information from an experiment, what should the drug concentrations be? Here, we consider the case where the single drug response curves are known beforehand, but no previous data is available from the combination. We consider several designs, including C and D-optimal designs, and a factorial design. We evaluate these designs based on the expected variance of the synergy score for a large set of in vitro experiments performed at Millennium. Based on the results, we were able to identify which design was the most efficient and robust.

Key Words: experimental design, drug combination studies, synergy

1. Background

Drug combinations have become an important part of cancer care and antiviral therapy. To identify synergistic drug combinations, and to understand the combined behaviour, scientists often perform in vitro drug combination studies. In the oncology setting, in vitro studies usually involve a cell viability assay applied to a cancer cell line. The assay usually involves a microtiter plate, where cells and various amounts of drugs are added to each well. The plate is then incubated, after which the cell viability is measured.

In the case where two drugs are considered, various methods have been used to analyze the data [1-4]. Some methods involve fitting a response surface model to describe the viability as a nonlinear function of the two drug concentrations [5, 6]. The fitted response surface can then be summarized by a single number to describe the synergy.

The choice of drug concentrations used in the experiment can affect the quality of the results. If the drug concentrations don't cover a reasonable range, then the response surface will be poorly estimated, so the synergy measure will be highly uncertain. In this paper, we will propose and evaluate several different designs for drug combination studies.

Experimental design for response surface estimation has a rich history in the literature. In particular, design methods for nonlinear response surface surfaces have been explored, both in a general context [7, 8], and in the context of drug combination studies [9, 10]. In the nonlinear setting, the optimal design for finding the response surface parameters actually depends on the parameters. If a small scale experiment has been done previously, then the parameter estimates from this experiment can be used to design a larger

experiment. However, combination studies are often performed without previous combination data. In this paper, we consider the case where previous combination data is unavailable, but where previous single agent data has been collected.

2. Analyzing Cell Viability Data

Millennium scientists performed a number of combination studies using 384 well microtiter plates (Figure 1). To analyze this data, we normalized the viability by scaling so that the median of the negative controls was 0 and the median of the positive controls was 100. More formally,

$$V_i = 100 \frac{U_i - \text{median}(U_-)}{\text{median}(U_+) - \text{median}(U_-)}$$

where V_i is the normalized viability of the i^{th} well, and U_i is the corresponding raw viability measurement. After normalization, the controls were discarded.

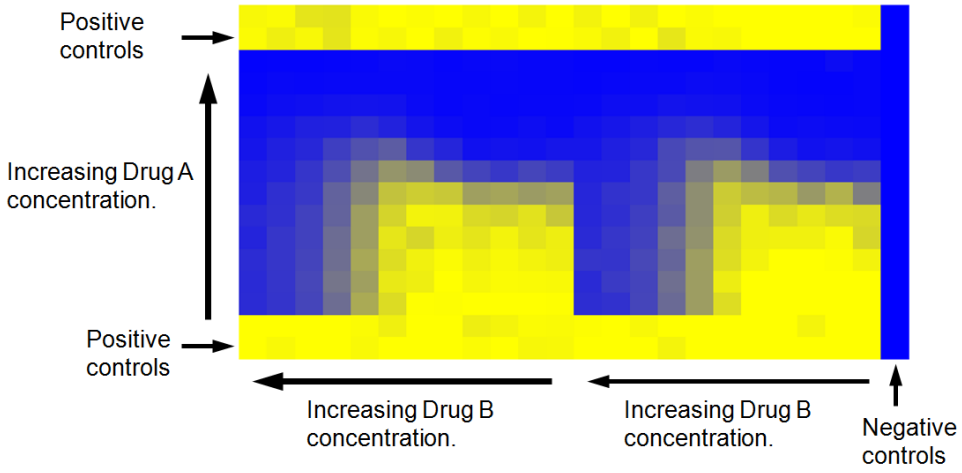

Figure 1: A heatmap showing the viability measurements for a drug combination study performed at Millennium. Here, yellow corresponds to high viability and blue corresponds to low viability.

2.1 Single Drug Experiments

Some of the experiments only involved single agents. For these cases, we assumed that the dose response curve had the form of the Hill equation [11]. Since the data was normalized, we assumed that the highest viability (the upper plateau) was 100. Thus, the viability was modeled as

$$V = 100 - \frac{E_{\max}}{1 + (I/C)^S} + error$$

where V is the normalized viability measurement. Here, E_{max} is the maximum drug effect, I is the inflection point, S is the slope, and C is the drug concentration. We assumed that the error values were independent and identically distributed normal random variables. We used the nlm() function in the R software package [12] to minimize the sum of the squared residuals and estimate the lower plateau, the slope, and the inflection point.

2.2 Combination Experiments

To describe the relationship between the normalized viability and the drug concentrations, we used a response surface model similar to that of [5], which is an extension of the Hill equation. For a given plate, let

$$C = (C_A / I_1) + (C_B / I_2)$$
$$x = (C_A / I_1) / C$$
$$E_{\max} = E_1 + E_2 x + E_3 x^2 + E_4 x^3$$
$$I = 1 + I_3 x(1 - x)$$
$$S = S_1 + S_2 x + S_3 x^2 + S_4 x^3$$
$$V = 100 - \frac{E_{\max}}{1 + (I / C)^S} + error$$

where E_1, E_2, E_3, E_4, I_1, I_2, I_3, S_1, S_2, S_3, and S_4 are parameters, C_A and C_B are the respective concentrations of drugs A and B, and V is the normalized viability measurement. This model has the property that along any line of constant dose ratio, the model has the form of a Hill equation. We assumed that the error values were independent and identically distributed normal random variables. We also included two constraints, thus yielding a model with 9 degrees of freedom.

We fit the data to our model by minimizing the residual sum of squares using the used the nlm() function in R. Next, we developed a measure of synergy that was a function of the 9 response surface parameters. We refer to this measure as the synergy score.

3. Designing Combination Experiments

For a given number of wells, we wish to choose doses that will minimize the variance of the synergy measure. Unfortunately, the best choice depends on the shape of the response surface, which we don't have until we do the experiment. However, if we have past single drug data, and we assume that there is no interaction, then we have a guess for the parameters. For a model with no interaction, we assume that $I_3 = 0$. We also assume that for $x = 0.5$, the resulting slope and E_{max} are found by averaging the slope and E_{max}, respectively, for the individual drugs. With these constraints, one can uniquely identify parameters for the model.

3.1 A C-Optimal Design

Figure 2 shows a strategy for finding an optimal study design. Given single agent data, one can produce a guess for the response surface parameters. In addition, given the dose choices for a proposed design, one can compute the likelihood function and the Fisher Information matrix, evaluated at the initial parameter guess. Using the Cramer-Rao lower bound [13, 14], the Fisher Information matrix can be used to estimate the variance matrix of the parameter estimates with the proposed design. Since the synergy score is simply a function of the response surface parameters, the one can use the Delta method [13] to estimate the expected variance of the synergy score under the proposed design. Then, one can adjust the proposed design to minimize the estimated variance.

The Delta method uses a first order (i.e. linear) approximation to relate the synergy score to the response surface parameters. Therefore, this approach minimizes the variance of a linear combination of the parameters. In the literature, this is called a C-optimal design [7].

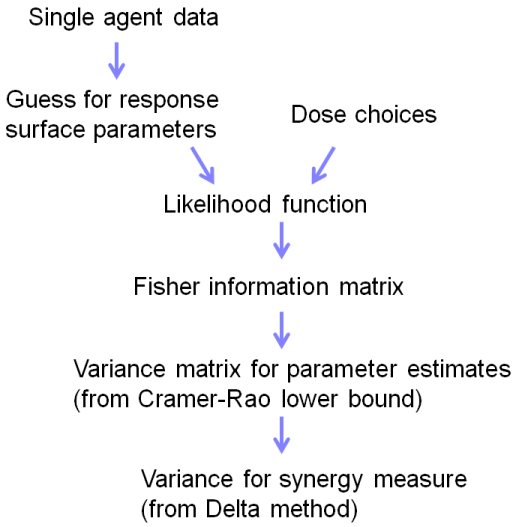

Figure 2: The process for choosing a design.

3.2 Computational Approach

The experimental design specifies the concentration for each of the two drugs in each microtiter well. Thus, minimizing the estimated variance of the synergy score would require a search over a space with dimension equal to twice the number of wells. Finding the global optimum in this space would require far too much computational power. Therefore, we developed a heuristic approach to this optimization problem.

Start with 6x6 log-spread grid points, search over the 50x50 candidate grid points and choose the point which minimizes the variance of the synergy measure estimate. Repeat the above search step until we obtain the number of design points needed. Remove the 36 starting grid points and search another 36 design points over the 50x50 candidate grid points.

3.3 Evaluating Designs

To evaluate our C-optimal design procedure, we obtained data from a set of 100 drug combination studies performed at Millennium using a variety of drugs and cancer cell lines. The studies were done with 384 well microtiter plates. The doses were arranged in a factorial design, with dose ranges manually chosen by the scientists. In addition, we obtained single agent data for each drug from previous experiments.

Figure 3 shows our method for evaluating the design. For each of the 100 experiments, we found the C-optimal design based on the past single agent data. Next, we took the corresponding combination data, estimated the response surface parameters, and used these estimates to predict the variance of the synergy measure. We believe this is a realistic assessment of the design procedure because it uses real single agent data to create the designs and real combination data to evaluate the designs.

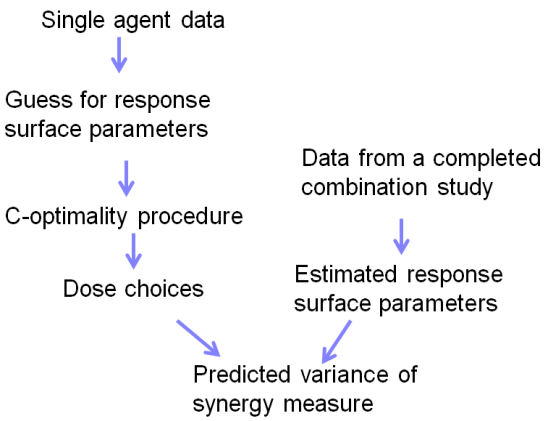

Figure 3: The process for evaluating the C-optimal design.

3.4 Other Designs

In addition to the C-optimal design, we considered two other designs, which we call the D-optimal design and the Automated Factorial design. The D-optimal design [7] is similar to the C-optimal design, except it minimizes the determinate of the estimated parameter covariance matrix.

The Automated Factorial design is a factorial design with dose ranges chosen based on the past single agent data. From the past single agent data, we estimate the slope and inflection point, and we used the following formula to select a dose range.

$$\text{Log dose range} = \left(\log(I) - a - \frac{b}{\text{slope}}, \log(I) + a + \frac{c}{\text{slope}} \right)$$

where I is the inflection point, and a, b, and c are constants. The dose levels are evenly distributed on the log scale along the selected dose range. Note that the dose range is centered around the inflection point, and the width of the range decreases with increasing slope. This ensures that the dose range covers the region when the response is changing.

4. Results and Discussion

The results are shown in Table 1. We found that the C-optimal design actually underperformed the designs manually created by the scientists. The D-optimal design is only slightly better than the manual designs. We believe these optimal designs underperform because there is a difference between the parameters used to generate the designs and the parameters used to evaluate the designs. The parameters used to generate the designs are found using the prior single agent data. This data is from a different batch than the combination data, so the parameter values may have shifted. Also, the initial parameter guesses assume that there is no drug synergy. Therefore, the initial parameter guesses may differ substantially from the parameters estimated from the combination data. We believe the C and D-optimal designs are not robust to these differences. Furthermore, previous studies have found that C and D-optimal designs can perform poorly if there is a high level of uncertainty in the initial parameter guesses [8].

Table 1: Predicted variance of the estimate for the synergy score for the various designs. The variance is expressed as a fraction of the variance expected under the manual factorial designs.

Experiment	*C-optimal*	*D-optimal*	*Automated factorial*
1	0.63	1.00	0.43
2	1.02	0.38	0.66
3	1.26	0.25	0.56
4	1.31	0.52	0.72
5	2.04	0.06	0.26
6	2.18	0.60	0.45
7	0.29	0.55	0.62
8	0.19	0.94	0.17
9	2.64	0.73	0.33
⋮	⋮	⋮	⋮
100	0.98	0.45	0.45
Mean	*4.60*	*0.83*	*0.49*

To confirm our view, we evaluated the C and D-optimal designs using the same parameters that were used to generate these designs. The results are shown in Table 2. As expected, the C and D-optimal designs perform well in this scenario.

Table 2: Predicted variance of the estimate for the synergy score for the C and D-optimal designs. In this case, the designs were evaluated using the same parameters that were used to generate the designs. The variance is expressed as a fraction of the variance under the manual factorial designs.

Experiment	*C-optimal*	*D-optimal*
1	0.10	0.22
2	0.12	0.24
3	0.08	0.20
4	0.11	0.34
5	0.04	0.10
6	0.05	0.15
7	0.15	0.39
8	0.02	0.004
9	0.02	0.07
⋮	⋮	⋮
100	0.21	0.45
Mean	*0.112*	*0.216*

In Table 1, the Automated Factorial design outperforms the manual factorial designs. The variance of the synergy measure is reduced by a factor of 2, which means the number of wells could be cut in half if the Automated Factorial design is used.

5. Conclusion

We presented a method to evaluate different designs for drug combination studies. We found that the C-optimal and D-optimal designs were not robust to misspecification of the response surface parameters. The automated factorial design showed a 2 fold reduction in the variance of the synergy measure.

Acknowledgements

We would like to thank the Double Agent team at Millennium for their support.

References

[1] Chou TC, Talalay P (1984). Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. Adv Enzyme Reg, 22, 27-55.

[2] Berenbaum MC (1985). The expected effect of a combination of agents: The general solution. J Theor Biol, 114, 413-431.

[3] Peterson JJ, Novik SJ (2007). Nonlinear Blending: A Useful General Concept for the Assessment of Combination Drug Synergy. Journal of Receptors and Signal Transduction, 27, 125-146.

[4] Bliss CI (1939). The toxicity of poisons combined jointly. Ann Appl Biol, 26, 585-615.

[5] Minto CF, Schnider TW, Short TG, Gregg KM, Gentilini A, Shafer SL (2000). Response Surface Model for Anesthetic Drug Interactions. Anesthesiology, 92, 1603-1616.

[6] Greco WR, Park HS, Rustum YM (1990). Application of a New Approach for the Quantitation of Drug Synergism to the Combination of cis-Diamminedichloroplatinum and 1-b-d-Arabinofuranosylcytosine. Cancer Res, 50, 5318-5327.

[7] Pukelsheim F (1993). Optimal design of experiments. New York: J. Wiley.

[8] Chaloner K and Verdinelli I (1995). Bayesian Experimental Design: A Review. Statistical Science, 10, 273-304.

[9] Tallarida RJ, Stone DJ Jr., Raffa RB (1997). Efficient designs for studying synergistic drug combinations. Life Sciences, 61, PL417-425.

[10] Chou TC (2006). Theoretical Basis, Experimental Design, and Computerized Simulation of Synergism and Antagonism in Drug Combination Studies. Pharmacol Rev 58, 621-681.

[11] Hill AV (1910). The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. J Physiol, 40:iv-vii.

[12] R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

[13] Cramer H (1946). Mathematical Methods of Statistics. Princeton, NJ: Princeton Univ. Press.

[14] Rao CR (1945). Information and the accuracy attainable in the estimation of statistical parameters. Bulletin of the Calcutta Mathematical Society 37, 81-89.