Received: 8 July 2021 | Revised: 8 March 2022 | Accepted: 12 April 2022

DOI: 10.1002/sim.9423

RESEARCH ARTICLE

Statistics in Medicine WILEY

Orthogonal array composite designs for drug combination experiments with applications for tuberculosis

Jose Luna[1] | **Jessica Jaynes**[2] | **Hongquan Xu**[1] | **Weng Kee Wong**[3]

[1]Department of Statistics, University of California, Los Angeles, Los Angeles, California, USA

[2]Department of Mathematics, California State University, Fullerton, Fullerton, California, USA

[3]Department of Biostatistics, University of California, Los Angeles, Los Angeles, California, USA

Correspondence
Jessica Jaynes, Department of Mathematics, California State University, Fullerton, CA 92831, USA.
Email: jjaynes@fullerton.edu

Funding information
NSF, Grant/Award Number: 2123380

Abstract

The aim of this article is to provide an overview of the orthogonal array composite design (OACD) methodology, illustrate the various advantages, and provide a real-world application. An OACD combines a two-level factorial design with a three-level orthogonal array and it can be used as an alternative to existing composite designs for building response surface models. We compare the *D*-efficiencies of OACDs relative to the commonly used central composite design (CCD) when there are a few missing observations and demonstrate that OACDs are more robust to missing observations for two scenarios. The first scenario assumes one missing observation either from one factorial point or one additional point. The second scenario assumes two missing observations either from two factorial points or from two additional points, or from one factorial point and one additional point. Furthermore, we compare OACDs and CCDs in terms of *I*-optimality for precise predictions. Lastly, a real-world application of an OACD for a tuberculosis drug combination study is provided.

KEYWORDS

drug combination, experimental design, factorial design, response surface model

1 | INTRODUCTION

Tuberculosis (TB) is in the top 10 leading causes of death in the world with about 1.4 million deaths in 2019.[1] TB is a treatable and curable disease but in some cases, severe drug resistance can develop. For these patients, there are limited treatment options and some can be expensive and toxic, and often require up to 2 years of chemotherapy treatment.[2] To reduce drug resistance and toxicity, various drug combinations are being used. One challenge in identifying an optimal combination of drugs is that the number of possible combinations increases exponentially as the number of drugs increases.[3] In the search to find the best drug-dose combinations, Silva et al[4] developed a second generation feedback system control (FSC.II) technology that dramatically reduces the number of iterations required for drug-dose optimization. In turn, this implies reduced time, effort, and costs for the treatment.

The FSC.II methodology has four phases. In Phases 1 and 2, dose response curves are generated for each drug and drug screening experiments are performed, respectively. Recently, Vazquez et al[5] proposed two-level designs with intermediate run sizes that can be considered in Phase 2 for drug screening. Phase 3 of the FSC.II methodology considers iterative drug combination testing and requires exploration of all possible combinations of numerous drugs at various levels. In Phase 3, typically an orthogonal array composite design (OACD) is used to develop a second-order model to describe the relationship between the drugs of interest and the response. Lastly, Phase 4 involves an optimization of the drug combinations and drug ratios.

Our work is focused on the OACDs in Phase 3. The novel development of an OACD was originally introduced by Xu et al[6] as an alternative to existing composite designs. An OACD combines a two-level factorial or fractional factorial design and a three-level orthogonal array (OA), and provides a good balance between run size economy and parameter estimation efficiency. Since the introduction of an OACD, research related to OACDs has been increasing including many theoretical developments and applications to enhance the use of other designs. Examples include: definitive screening designs,[7] uniform composite designs,[8] designs for third-order models,[9] as well as exploring the robustness to missing data.[10] Moreover, given the growing literature on the use of design of experiments in medical applications,[11,12] there are also numerous and diverse applications specifically of OACDs in various disciplines. Some examples include determining micropollutants for water contamination,[13] a 12-drug application to determine the optimal combination therapy against SARS-CoV-2,[14] an application with T-cell lymphoma,[15] optimization of antioxidant drug combinations in skin cancer,[16] studying prostate cancer drugs,[17] lipid accumulation for biodiesel fuel,[18] carcinoma cancer applications,[19-21] among others.[22-24]

The aim of this article is to demonstrate the OACD methodology by building upon the foundation set by Xu et al[6] and illustrate the advantages of using an OACD as an alternative to an existing popular class of composite designs. Section 2 introduces the OACD methodology and reviews the construction method. Section 3 compares OACDs with popular central composite designs (CCDs) in terms of D-efficiency for estimating model parameters. Section 4 compares OACDs with CCDs in terms of I-optimality for making precise predictions of the response. Section 5 demonstrates that OACDs can have better robustness properties to missing observations than CCDs. Section 6 illustrates the construction and analysis of an OACD for a TB drug combination study. This study was selected from the various applications of OACDs in the TB literature.[2-4,25] Additionally, in the context of the TB application, the built-in ability to perform cross-validation and the robustness to missing observations from either a factorial or additional point are shown. The article concludes with some remarks in Section 7.

2 | ORTHOGONAL ARRAY COMPOSITE DESIGNS

Before we discuss the OACD methodology in more detail, we provide a brief background on factorial designs and composite designs. Factorial designs are efficient for studying the effects of two or more factors simultaneously. For experiments with k factors, each at two levels, full factorial designs have 2^k observations or runs that enable the estimation of the linear effect of each factor on the response over the range of the factor levels chosen, these are known as main effects. Additionally, a two-factor interaction effect occurs when the effect of one factor on the response changes when the setting of a second factor changes. Furthermore, all higher-order interactions are similarly defined. Collectively, main effects and all interaction effects are called factorial effects.[26]

A full factorial design enables the estimation of all main effects and interaction effects. However, they are often impractical or too expensive because the sample size grows exponentially with the number of factors. For economical reasons, a fraction of the runs in the full factorial design can be used, these are called fractional factorial designs.[12] Fractional factorial designs are broadly classified into two categories: regular designs and nonregular designs. A regular fractional factorial design is defined for k factors, each at two levels, as a 2^{-p} fraction of the full 2^k factorial design and denoted by 2^{k-p}. It has 2^{k-p} runs and because of the smaller sample size it cannot estimate all possible higher-order interaction effects. In particular, the factorial effects become aliased with one another. To overcome this issue, it is typically assumed that higher-order interactions are negligible because they are less likely to be important than lower-order interactions.[26] The amount of aliasing in a fractional factorial design is measured by its resolution. A larger resolution means less restrictive assumptions regarding which interactions are negligible to obtain a unique interpretation of the data. For selection of a regular fractional factorial design, the minimum aberration criterion is recommended, which maximizes the resolution of the design.[26]

A regular fractional factorial design has what is known as a simple aliasing structure, in that any two effects are either orthogonal or fully aliased. In contrast, a nonregular design has a complex aliasing structure, such that there exist effects that are neither orthogonal nor fully aliased. Nonregular designs are often considered for their run size economy and can be considered if a smaller design is needed. The generalized minimum aberration criterion[27,28] can be used for the selection of nonregular designs. Generalized minimum aberration designs minimize the overall contamination of nonnegligible interactions on the estimation of main effects.[27,28]

Factorial designs provide an efficient way to identify significant factors that affect the response of interest. After the significant factors have been identified, the next step is to model the response keeping in mind that there may be a curvilinear

relationship with the factors, in which case we may be interested to estimate combination levels of the factors that maximize or minimize the expected response variable. To this end, we entertain a second-order model with k quantitative factors of the form

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \epsilon, \tag{1}$$

where β_0, β_i, β_{ii}, and β_{ij} are the intercept, linear, quadratic, and bilinear (or two-factor interaction) terms, respectively. The error term ϵ is assumed to have mean zero and constant variance and all errors are assumed to be independent. A composite design may be used to estimate all parameters in the second-order model (1). Such a design has three parts: (i) n_f factorial points $(x_1, \ldots, x_k)$ with all $x_i = -1$ or 1 from a two-level full factorial or fractional factorial design; (ii) n_a additional points with all $x_i = -\alpha, 0, \alpha$, that are commonly known as axial points; (iii) n_0 center points with all $x_i = 0$. Without loss of generality, we assume $\alpha = 1$ and focus on one particularly useful class of composite designs in the rest of the article.

An OACD is a class of composite designs based on a two-level factorial design and a three-level OA.[6] An OA of N runs, k columns, s levels, and strength t, denoted by $\mathrm{OA}(N, s^k, t)$, is an $N \times k$ matrix in which all s^t level-combinations appear equally often in every $N \times t$ submatrix.[26] We omit the strength t when $t = 2$. For example, a 2^k factorial design can be viewed as $\mathrm{OA}(N, 2^k, t)$ with $N = 2^k$ and $t = k$. Similarly, a three-level OA can be written as $\mathrm{OA}(N, 3^k, t)$. Thus, an OACD is a composite design which consists of a two-level factorial design as its factorial points, a three-level OA as its additional points, plus any number of center points.

The choice of the two-level factorial design and three-level OA should be based on good or optimal two-level and three-level designs using criterion such as the minimum aberration or generalized minimum aberration. Once a two-level and a three-level design are chosen, they can be put together to form an OACD. To construct an OACD, the simplest approach is to simply stack the two-level and three-level designs together, and the resulting OACD often has good or optimal properties with respect to some criterion. For an OACD, it has been shown that column permutations or level permutations of an OA may lead to different geometrical structures and efficiencies.[6,29] Therefore, one can permute columns and levels of the three-level portion in an OACD to obtain better efficiency with respect to some criterion. For the k-factor three level portion, the total number of permutations is $k!(3!)^k$, which is computationally expensive or infeasible for moderate or large k. Therefore, we can find a (nearly) best permutation through a stochastic search algorithm such as the threshold accepting algorithm.[30] Otherwise, if the total number of permutations is feasible, we can use an enumeration method to search for the best permutation such that the permuted OACD is more efficient with respect to some criterion. In particular, by combining a two-level and a three-level design in an OACD, we are able to fit a second-order model to screen important factors by studying the linear, quadratic, and bilinear effects.[6] For this reason, an OACD is sometimes referred to as a second-order design.

An appealing feature of the OACD is the built-in ability to perform cross-validation on the analysis results. Using an OACD, we can fit three types of models: a model with linear and bilinear terms based on the two-level design, a model with linear and quadratic terms based on the three-level OA, and a full second-order model based on the entire OACD. Therefore, each linear effect is estimated three times and each bilinear or quadratic effect is estimated twice, providing the ability to check the consistency of their estimates. With this built-in ability to perform cross-validation it is important to note that the estimates of the parameters of the full second-order model will be correlated with the estimates of the parameters from the other two models. Additionally, the OACD can be implemented sequentially, similar to many existing composite designs.

3 | *D*-EFFICIENCY OF COMPOSITE DESIGNS

When constructing a design for an experiment one must consider the possible trade-offs between the choice of the design and the design efficiency based on some statistical criterion. A widely accepted criterion is D-optimality for estimating model parameters. In this section, we briefly discuss the concept of D-efficiency of a design and refer the reader to Kiefer[31] for additional details. Let d be a second-order design with N runs and let $\boldsymbol{X} = (\boldsymbol{1}, \boldsymbol{L}, \boldsymbol{Q}, \boldsymbol{B})$ be the model matrix of d, where $\mathbf{1}$ is a column of ones, and $\boldsymbol{L}, \boldsymbol{Q}, \boldsymbol{B}$ denote the columns corresponding to the linear, quadratic, and bilinear terms, respectively, in Equation (1). Let $\boldsymbol{M}(d) = \boldsymbol{X}'\boldsymbol{X}/N$ be the (normalized) information matrix for d. The D-optimality criterion

seeks to maximize the determinant of the information matrix $|\boldsymbol{M}(d)|$ by choice of a design. This in turn minimizes the generalized variance of the parameter estimates in Equation (1) and so the obtained estimates are as precise as possible given the sample size N.

For any k-factor design d, its D-efficiency is defined by

$$D_{\text{eff}}(d) = \left(\frac{|\boldsymbol{M}(d)|}{|\boldsymbol{M}(\xi^*)|} \right)^{1/p},$$

where ξ^* is an approximate D-optimal design and $p = (k+1)(k+2)/2$ is the number of parameters in the second-order model (1). It is instructive to consider approximate designs, which can be viewed as probability measures on the design space and are easier to study analytically.[32] For example, the determinant of the information matrix of the approximate D-optimal design ξ^* on the design space $[-1, 1]^k$ has a closed form and is given by[31,33]

$$|M(\xi^*)| = u^k v^{k(k-1)/2}(u-v)^{k-1}(u+(k-1)v-ku^2), \tag{2}$$

where

$$u = \frac{k+3}{4(k+1)(k+2)^2} \times \left((2k^2+3k+7) + (k-1)(4k^2+12k+17)^{1/2} \right),$$
$$v = \frac{k+3}{8(k+2)^3(k+1)} \times \left((4k^2+8k^2+11k-5) + (2k^2+k+3)(4k^2+12k+17)^{1/2} \right).$$

As a result, we can use (2) to compute the D-efficiency for a given design.

When only a subset of the model parameters is of interest, we find a D_s-optimal design that provides the most accurate estimates for parameters in the subset only. If s is the subset of interest, the D_s criterion value is defined by

$$D_s(d) = \frac{1}{N} \left| \boldsymbol{X}_s'\boldsymbol{X}_s - \boldsymbol{X}_s'\boldsymbol{X}_{(s)} \left(\boldsymbol{X}_{(s)}'\boldsymbol{X}_{(s)} \right)^{-1} \boldsymbol{X}_{(s)}'\boldsymbol{X}_s \right|^{1/|s|}, \tag{3}$$

where $\boldsymbol{X}_s$ and $\boldsymbol{X}_{(s)}$ are the submatrices of $\boldsymbol{X}$ corresponding to the parameters in s and not in s, respectively, and $|s|$ is the number of parameters in s.

We consider the subsets consisting of the k linear parameters, the k quadratic parameters, and the $k(k-1)/2$ bilinear parameters, respectively. The D_s-efficiency of a design d can be calculated as[7]

$$D_{L,\,\text{eff}}(d) = D_L(d),\ D_{Q,\,\text{eff}}(d) = 4D_Q(d),\ D_{B,\,\text{eff}}(d) = D_B(d),$$

where D_L, D_Q, and D_B are computed as in (3) for the linear, quadratic, and bilinear terms, respectively.

Table 1 provides the design generating information for both the two-level factorial portion and three-level OA portion for $k = 3, \ldots, 11$ factors, where the two-level design is either a full factorial design or has at least resolution V. We compare the D-efficiency and D_s-efficiency of the OACDs in Table 1 with the popular CCDs. As described in Section 2, a composite design has a total of $N = n_f + n_a + n_0$ points, and has either three or five levels depending on whether $\alpha = 1$ or not. For $k = 3, \ldots, 11$, the factorial points used in the OACDs and the CCDs are the same, but the additional points in the CCD are obtained with $n_a = 2k$. Figure 1 compares the performances of the D-efficiencies and the D_s-efficiencies for these composite designs with $n_0 = 0$. The overall D-efficiency of the OACD outperforms the popular CCD for $k > 4$. Furthermore, we observe that the D_s-efficiencies of the OACDs outperform those of the CCDs for all k in both the linear and bilinear terms. For the quadratic terms, OACD achieves higher D_Q-efficiency than CCD for $k \geq 5$.

So far we assume that the two-level portion is a regular resolution V design, which may have too many runs for some applications when $k > 5$. To reduce the number of runs, one can choose a resolution III or IV design, regular or nonregular, as the two-level portion. It is well known that a regular resolution IV design cannot be used as the two-level portion for a CCD as this does not lead to a second-order design.[26] In contrast, we can use a regular resolution IV design as the two-level portion to construct a smaller OACD with high D-efficiency.[6,7]

TABLE 1 Select OACDs for $k = 3, \ldots, 11$ factors

	Two-level factorial portion			Three-level OA	
k	Design	n_f	Columns and generators	n_a	Columns
3	2^3	8	-	OA(9)	(1-3)
4	2^4	16	-	OA(9)	(1-4)
5	2_V^{5-1}	16	$x_5 = x_1x_2x_3x_4$	OA(18)	(2-6)
6	2_{VI}^{6-1}	32	$x_6 = x_1x_2x_3x_4x_5$	OA(18)	(1-6)
7	2_{VII}^{7-1}	64	$x_7 = x_1x_2x_3x_4x_5x_6$	OA(18)	(1-7)
8	2_V^{8-2}	64	$x_7 = x_1x_2x_3x_4x_5$, $x_8 = x_1x_2x_3x_6$	OA(27)	(1-8)
9	2_V^{9-2}	128	$x_8 = x_1x_2x_3x_4x_5$, $x_9 = x_1x_2x_3x_6x_7$	OA(27)	(1-9)
10	2_V^{10-3}	128	$x_8 = x_1x_2x_3x_4x_5$, $x_9 = x_1x_2x_3x_6x_7$, $x_{10} = x_1x_2x_4x_6$	OA(27)	(1-10)
11	2_V^{11-4}	128	$x_8 = x_1x_2x_3x_4x_5$, $x_9 = x_1x_2x_3x_6x_7$, $x_{10} = x_1x_2x_4x_6$, $x_{11} = x_1x_3x_5x_7$	OA(27)	(1-11)

Note: OA(9), OA(18), and OA(27) are available in Appendix Table A1.

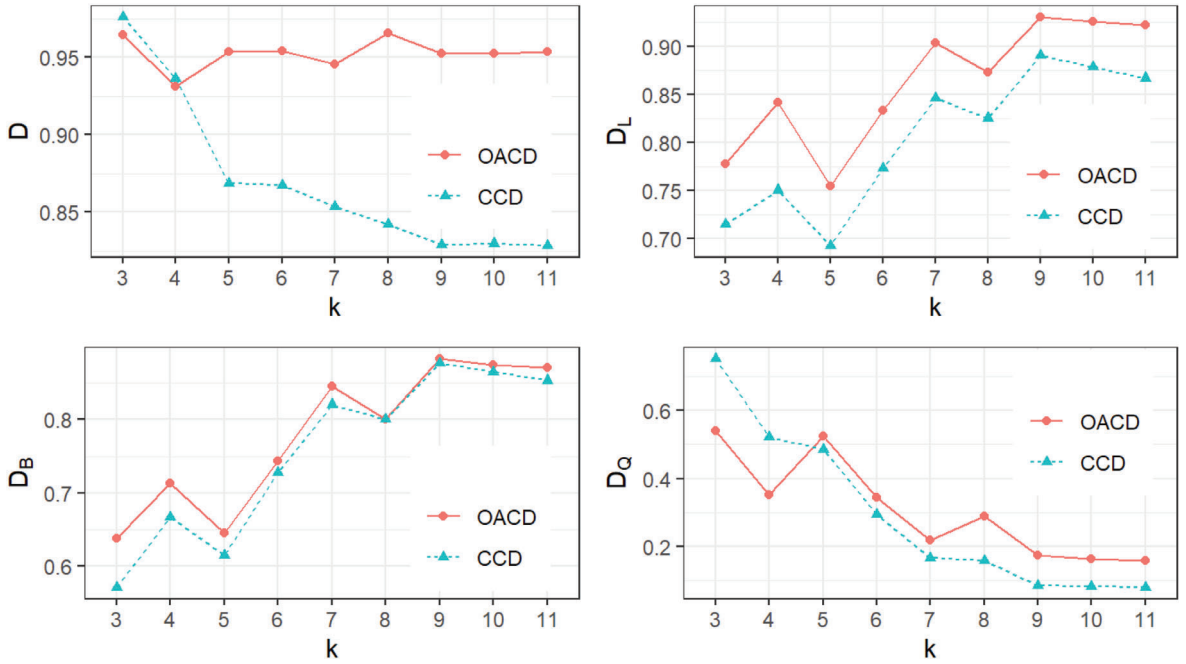

FIGURE 1 D-efficiencies and D_s-efficiencies for OACDs and CCDs of various sizes with $3 \le k \le 11$ and $n_0 = 0$

4 | *I*-OPTIMALITY OF COMPOSITE DESIGNS

Given that the purpose of an OACD is to fit a second-order model, I-optimality is another appropriate design criterion to use to compare OACDs and CCDs. Whereas D-optimality minimizes the generalized variance to obtain precise parameter estimates for all the model parameters, I-optimality minimizes the average prediction variance and is used to make precise predictions of the response.

Given an N-run design d and the second-order model (1), the prediction variance at an arbitrary point $\boldsymbol{x}$ is $\text{var}(\hat{y}(\boldsymbol{x})) = \sigma^2\boldsymbol{f}'(\boldsymbol{x})(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{f}(\boldsymbol{x}) = \sigma^2\boldsymbol{f}'(\boldsymbol{x})\boldsymbol{M}(d)^{-1}\boldsymbol{f}(\boldsymbol{x})/N$, where $\boldsymbol{f}'(\boldsymbol{x}) = (1, x_1, \ldots, x_k, x_1^2, \ldots, x_k^2, x_1x_2, \ldots, x_{k-1}x_k)$. An I-optimal design minimizes the normalized average or integrated prediction variance

$$\text{I(d)} = \frac{N}{\sigma^2}\int \text{var}(\hat{y}(\boldsymbol{x}))\mu(\boldsymbol{x}) = \int \boldsymbol{f}'(\boldsymbol{x})\boldsymbol{M}(d)^{-1}\boldsymbol{f}(\boldsymbol{x})\mu(\boldsymbol{x}),$$

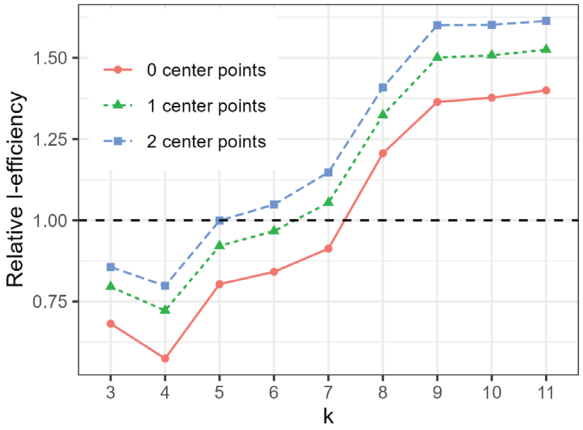

FIGURE 2 I-efficiency of OACDs relative to CCDs with $3 \leq k \leq 11$ and $n_0 = 0, 1, 2$

where μ is the uniform measure on the experimental region with total measure 1. When the experimental region is $[-1, 1]^k$, the I-value can be computed as[34]

$$\mathrm{I(d)} = \mathrm{trace}[\boldsymbol{M}(d)^{-1}\boldsymbol{A}],$$

where

$$\boldsymbol{A} = \begin{bmatrix} 1 & \mathbf{0} & \frac{1}{3}\mathbf{1}_k' & \mathbf{0} \\ \mathbf{0} & \frac{1}{3}\boldsymbol{I}_k & \mathbf{0} & \mathbf{0} \\ \frac{1}{3}\mathbf{1}_k & \mathbf{0} & \frac{4}{45}\boldsymbol{I}_k + \frac{1}{9}\boldsymbol{J}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{9}\boldsymbol{I}_q \end{bmatrix} \tag{4}$$

is the moment matrix, $\mathbf{0}$ is a vector or matrix of appropriate dimension containing all zeros, $\mathbf{1}_k$ is column vector containing k ones, $\boldsymbol{I}_k$ is a $k \times k$ identity matrix, $\boldsymbol{J}_k$ is a $k \times k$ matrix of ones, and $q = k(k-1)/2$. Unlike D-optimality, an approximate I-optimal design over a cube $[-1, 1]^k$ is unknown in the literature. To compare an OACD with a CCD, we use the relative I-efficiency defined as

$$\text{Relative } I\text{-efficiency} = \frac{I(\text{ccd})}{I\ (\text{oacd})} = \frac{\mathrm{trace}[\boldsymbol{M}(\text{ccd})^{-1}\boldsymbol{A}]}{\mathrm{trace}[\boldsymbol{M}(\text{oacd})^{-1}\boldsymbol{A}]}.$$

Figure 2 shows the relative I-efficiency comparing OACDs to CCDs with $k = 3, \ldots, 11$ and $n_0 = 0, 1, 2$. With $n_0 = 0, 1, 2$ center runs, OACDs achieve higher I-efficiency than CCDs for $k \geq 8$, $k \geq 7$, and $k \geq 6$, respectively. If the goal of the experiment is prediction, then it is recommended to add at least one or two center runs to the design since it provides information on the error variance and helps stabilize the variance of the predicted response. On the other hand, adding center runs will reduce the D-efficiency slightly for both OACDs and CCDs.[7]

5 | ROBUSTNESS OF OACDS

For any design, missing observations are sometimes unavoidable even from a well-planned experiment. Missing observations can happen for a variety of reasons including errors in the data collection or outliers. Smucker et al[35] explore the robustness of classical and optimal designs to missing observations and provide some empirical results in both screening and response surface settings. Akhtar and Prescott[36] introduce a loss criterion based on D-efficiency for the CCD to assess the robustness when observations are missing. The aim of this criterion is to minimize the maximum loss for missing observations. Our attention focuses on minimizing the effect of missing observations and comparing the robustness of OACDs and CCDs with missing observations.

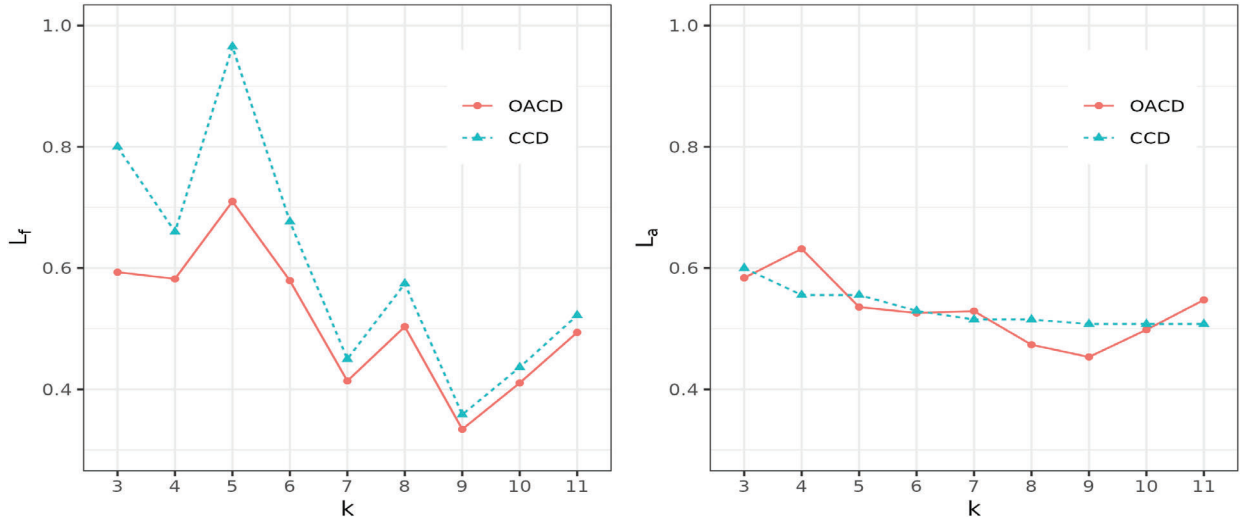

FIGURE 3 Relative loss for one missing factorial point (L_f) and one missing axial point (L_a) for OACDs and CCDs with $3 \leq k \leq 11$ and $n_0 = 0$

In general, if $m = 1, 2, \ldots$ observations are missing from an experiment, then m rows in the model matrix $\boldsymbol{X}$ will be missing. If $\boldsymbol{X}_m$ is the $m \times p$ matrix of the m missing rows corresponding to the m missing observations, and $\boldsymbol{X}_r$ is the $(N - m) \times p$ reduced model matrix, then the information matrix can be expressed as $\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}_m'\boldsymbol{X}_m + \boldsymbol{X}_r'\boldsymbol{X}_r$. It is well known that the determinant of the reduced information matrix, $|\boldsymbol{X}_r'\boldsymbol{X}_r|$, does not exceed the determinant of full information matrix, $|\boldsymbol{X}'\boldsymbol{X}|$. The relative reduction in $|\boldsymbol{X}'\boldsymbol{X}|$ due to a missing observation is called *l*oss. The relative loss can be defined as L_i, which represents the loss for the ith combination of $\binom{N}{m}$ when m runs are missing: $L_i = \frac{|\boldsymbol{X}'\boldsymbol{X}| - |\boldsymbol{X}_r'\boldsymbol{X}_r|}{|\boldsymbol{X}'\boldsymbol{X}|} = 1 - \frac{|\boldsymbol{X}_r'\boldsymbol{X}_r|}{|\boldsymbol{X}'\boldsymbol{X}|}$; $i = 1, \ldots, \binom{N}{m}$. As $0 \leq L_i \leq 1$, a small value of L_i indicates a low reduction in the determinant of the information matrix and, in this sense, less loss of information. Thus, a design with a smaller loss due to missing observations is judged more robust.[37]

Based on the composite structure of an OACD we classify the loss of a single observation as follows: L_f as the loss for one missing factorial point, and L_a as the loss for one missing axial point. Whether the missing observation is a factorial point or an axial point, the losses are averaged to one value since empirical studies have shown that there is little difference among the losses in each type.[10] Lastly, the loss of a center point is not considered in this study as their losses are small relative to the losses of a factorial or an axial point.

Figure 3 displays the loss of a single factorial point L_f and the loss of a single axial point L_a for OACDs and CCDs with $k = 3, \ldots, 11$ and $n_0 = 0$. When an OACD and a CCD are constructed with the same factorial portions (see Table 1) there will be lower loss of L_f for an OACD than a CCD for all k. However when considering the loss from an axial point, there is a decreasing pattern in L_a for CCDs compared to OACDs. In particular, this constant decrease in L_a results in CCDs achieving a lower loss of L_a for $k = 4, 7$, and 11. Note that when an OACD uses the same OA for its axial points consecutively for each additional factor, the run size will only increase by the number of factorial points being added. On the other hand, a CCD increases its axial points by two runs for each additional factor, which can result in lower loss in L_a when the OACD and CCD share the same two-level factorial portion.

Next, the loss from two missing observations is considered. In general, there are $\binom{N}{2}$ combinations of two missing runs from an N-run design. We classify the pairs of missing observations as follows: L_{ff} as the loss for missing two factorial points; L_{fa} as the loss for missing a factorial and an axial point; L_{aa} as the loss for missing two axial points. Similar to the case with one missing observation, the losses within each respective group are averaged to one value.

Figure 4 displays the loss of information when two observations are missing, classified as L_{ff}, L_{aa}, L_{fa} for both OACD and CCD. If we consider the loss of two factorial points, OACD achieves lower loss for all $k = 3, \ldots, 11$. When two axial points are missing the CCD has a decreasing pattern and results in a lower L_{aa} for $k = 4, 7$, and 11, which is similar to the case when one observation is missing. However, when a factorial and an axial point are missing, the decreasing pattern diminishes and the OACD achieves lower L_{fa} except for $k = 4$ and 11. It should be noted that the addition of center runs (n_0) could potentially reduce the loss of information due to a missing axial point, depending on the number of center

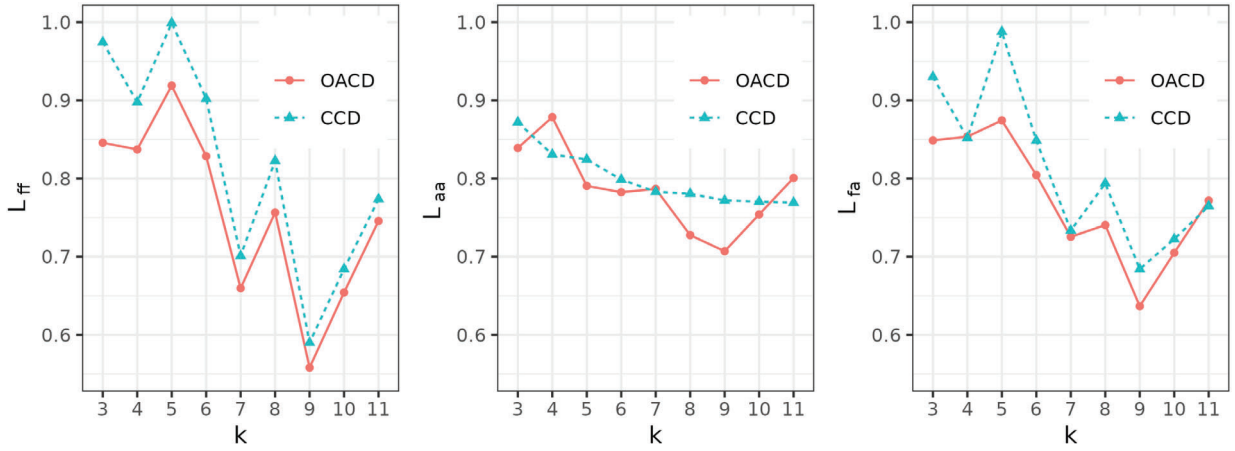

FIGURE 4 Relative loss for two missing factorial points (L_{ff}), missing one factorial point and one axial point (L_{aa}), and two missing axial points (L_{fa}) for OACDs and CCDs with $3 \leq k \leq 11$ and $n_0 = 0$

runs added to the design. Furthermore, at the expense of nine additional runs, for $k = 11$ factors, a 36-run OA could be considered instead of a 27-run OA for the three-level OA portion of an OACD, which would achieve lower loss in L_a, L_{aa}, L_{fa} compared to a CCD. In general, if one can afford adding a few additional runs the resulting design would be more robust against the loss of information from possible missing observations.

Lastly, we examine the impact of missing observations on the predictive capability of OACDs compared to CCDs. To compare the full design and reduced design in terms of the I-optimality criterion, we define the relative loss of I-efficiency as follows:

$$\text{RL} = 1 - \frac{\text{trace}\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{A}\right]}{\text{trace}\left[(\boldsymbol{X}_r'\boldsymbol{X}_r)^{-1}\boldsymbol{A}\right]},$$

where $\boldsymbol{X}$ and $\boldsymbol{X}_r$ are the model matrices of the full design and reduced design, respectively, and $\boldsymbol{A}$ is the moment matrix defined in (4). As $0 \leq \text{RL} \leq 1$, a small value of RL indicates a lower loss in average prediction variance in our experimental region. Figure 5 displays the loss of I-efficiency for a single missing factorial point (RL_f), and the loss of I-efficiency from a single missing axial point (RL_a) for OACDs and CCDs with $k = 3, \ldots, 11$ and $n_0 = 1$. As before, we average the loss over all factorial points or all axial points when calculating RL_f or RL_a. With one missing factorial point, the OACD achieves lower loss for $k \geq 3$, however the performance between OACD and CCD are nearly identical for $k \geq 7$. For $k = 5$, the CCD has notably high loss with one missing factorial point. If a single axial point is missing, the OACD achieves lower loss for $3 \leq k \leq 11$ except $k = 4$. Furthermore, the loss of I-efficiency can be explored for two missing observations. While not pictured here, we did explore two missing observations and the trends in the loss of I-efficiency for $3 \leq k \leq 11$ and the results are very similar to the situations when only one observation was missing, as described above.

6 | TB APPLICATION

To illustrate the construction, analysis, built-in ability to perform cross-validation, and robustness of an OACD, we consider the TB study in Silva et al.[4] This study applies the FSC.II methodology to a high throughput macrophage cell culture model of TB using mycobacterium tuberculosis that expresses isopropyl β-D-1-thiogalactopyranoside-inducible green fluorescent protein (Mtb-iGFP). The efficacy of the TB drug combinations in inhibiting intramacrophage Mtb-iGFP using fluorescence-based assay with high-throughput imaging analysis is measured using inhibition defined as follows:

$$\text{Inhibition} = 1 - \left(\frac{\text{Integrated GFP fluorescence intensity per nucleus of treated sample}}{\text{Integrated GFP fluorescence intensity per nucleus of untreated sample}}\right).$$

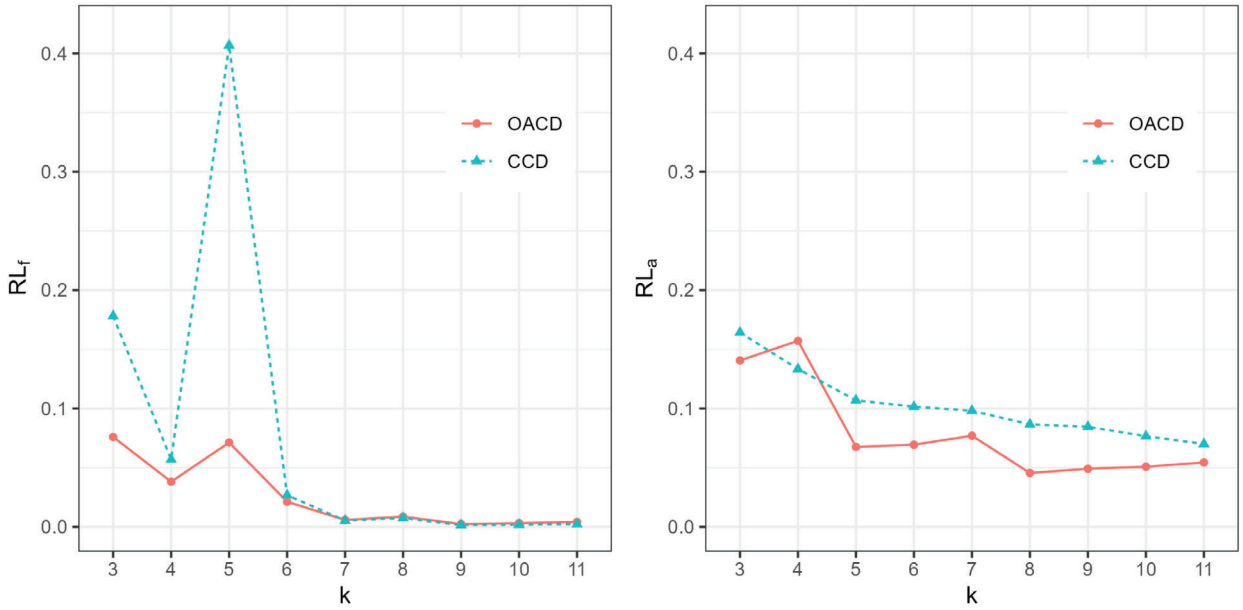

FIGURE 5 Relative loss of I-efficiency for one missing factorial point (RL_f) and one missing axial point (RL_a) for OACDs and CCDs with $3 \leq k \leq 11$ and $n_0 = 1$

TABLE 2 TB drug abbreviations

Abbreviation	Drug name
AMC	Amoxicillin/clavulanate
CLZ	Clofazimine
CYS	Cycloserine
EMB	Ethambutol
INH	Isoniazid
LZD	Linezolid
MXF	Moxifloxacin
PA824	PA-824
PAS	Para-aminosalicylic acid
PRO	Prothionamide
PZA	Pyrazinamide
RIF	Rifampicin
SQ109	SQ-109
TMC	TMC-207

A series of experiments are performed and the percent of inhibition is recorded for each. The data from each experiment is used to fit a second-order model to identify the individual and combinations of drugs that produce the highest percentage of inhibition. Positive model coefficients indicate a higher percentage of inhibition, and negative model coefficients indicate a lower percentage of inhibition.

In the TB study by Silva et al,[4] the four phases of the FSC.II methodology are implemented, starting with 14 TB drugs. This study provides a unique application of OACDs with a large number of factors, as many of the applications in the literature are for a much smaller number of factors. Additionally, the particular approach of the FSC.II methodology has the capability to support larger designs. The initial 14 drugs and their abbreviations are presented in Table 2. In Phase 2 of Silva et al,[4] a two-level screening test is used to reduce the number of drugs to the following

TABLE 3 Parameter estimates for the nine-drug experiment

	155-Run	128-Run	27-Run
CLZ	0.0601***	0.0605***	0.0661**
EMB	0.0550***	0.0553***	0.0589**
PA824	0.0229***	0.0234***	0.0506*
PAS	0.0037	0.0031	0.0178
PRO	0.0405***	0.0409***	0.0494*
PZA	0.0269***	0.0277***	0.0394*
RIF	0.0154***	0.0155***	0.0178
SQ109	0.0608***	0.0617***	0.0589**
TMC	0.0412***	0.0423***	0.0461*
CLZ^2	−0.0559**		−0.0372
EMB^2	−0.0243		−0.0189
$PA824^2$	−0.0168		−0.0339
PAS^2	0.0068		0.0044
PRO^2	−0.0073		−0.0006
PZA^2	0.0106		0.0328
RIF^2	0.0257		0.0311
$SQ109^2$	0.002		0.011
TMC^2	0.0249		0.0361
CLZ:EMB	−0.0112**	−0.0114**	
CLZ:PA824	0.0026	0.002	
CLZ:PAS	−0.0017	−0.002	
CLZ:PRO	−0.0142***	−0.0152***	
CLZ:PZA	−0.0053	−0.0063	
CLZ:RIF	−0.0123***	−0.0138***	
CLZ:SQ109	−0.0281***	−0.0291***	
CLZ:TMC	−0.0024	−0.0028	
EMB:PA824	0.003	0.0028	
EMB:PAS	−0.0014	−0.0025	
EMB:PRO	−0.0113**	−0.0113**	
EMB:PZA	−0.0116**	−0.0127**	
EMB:RIF	0.0039	0.0036	
EMB:SQ109	−0.0465***	−0.0473***	
EMB:TMC	−0.0054	−0.0061	
PA824:PAS	−0.002	−0.0022	
PA824:PRO	−0.0053	−0.0069	
PA824:PZA	−0.002	−0.0017	
PA824:RIF	−0.0025	−0.003	
PA824:SQ109	0.0028	0.0033	
PA824:TMC	−0.0044	−0.0052	

TABLE 3 Continued

	155-Run	128-Run	27-Run
PAS:PRO	0.0011	−0.0006	
PAS:PZA	0.0047	0.0036	
PAS:RIF	0.0033	0.0027	
PAS:SQ109	−0.0085*	−0.0092*	
PAS:TMC	0.0018	0.0014	
PRO:PZA	−0.0101**	−0.0105**	
PRO:RIF	−0.0042	−0.0045	
PRO:SQ109	−0.022***	−0.022***	
PRO:TMC	−0.0081*	−0.0089*	
PZA:RIF	−0.0081*	−0.0084*	
PZA:SQ109	−0.0029	−0.0041	
PZA:TMC	−0.0036	−0.005	
RIF:SQ109	−0.0085*	−0.0091*	
RIF:TMC	−0.0087*	−0.01**	
SQ109:TMC	−0.007	−0.0069	
Residual error $\hat{\sigma}$	0.042	0.04264	0.07013
R^2	0.9462	0.9471	0.9123
Significance levels:	0'***'	0.001'**'	0.01'*' 0.05 '.'

nine drugs: CLZ, EMB, PA824, PAS, PRO, PZA, RIF, SQ109, TMC. The focus of this article is on Phase 3 of the FSC.II methodology, which consists of a series of iterative experiments using OACDs. Specifically, we focus on Iterations I and II from Silva et al,[4] which use a 155-run OACD for these nine drugs and a 50-run OACD for six out of these nine drugs.

6.1 | Nine-drug 155-run OACD

For simplicity, we will call these nine drugs: $x_1, x_2, \ldots, x_9$ (CLZ, EMB, PA824, PAS, PRO, PZA, RIF, SQ109, TMC), which is common in factorial design literature. Each drug is tested at three levels: zero dose, half dose, and full dose, of the concentrations of the drug that achieved 15% inhibition in the screening test, see Silva et al[4] for additional details. For model fitting purposes, each drug is coded as −1 for the zero dose, 0 for the half dose, and +1 for the full dose.

The design for the 155-run OACD with $k = 9$ is presented in Table 1. This OACD combines a 128-run two-level fractional factorial design, such that the levels of each drug are coded as −1 and +1, and a 27-run three-level OA, with levels coded as $-1, 0$, and +1. The 128-run two-level factorial design is a one-quarter fraction of a full 2^9 factorial design with design generators $x_8 = x_1x_2x_3x_4x_5$ and $x_9 = x_1x_2x_3x_6x_7$. This fractional factorial design has resolution V, which enables the estimation of all nine linear effects plus all 36 bilinear effects. The 27-run OA is a subset of the commonly used OA$(27, 3^{13})$, which is a 27×9 matrix such that all 3^2 level combinations appear equally often in all 27×2 submatrices. This OA has the ability to estimate all nine linear effects plus nine quadratic effects for each drug. Table S1 in the supplementary material includes the 155-run OACD and data.

To illustrate the built-in ability to perform cross-validation, we fit three models using the 155-run OACD. First, using all 155 runs we fit a second-order model to the data to estimate the linear, quadratic, and bilinear effects. Second, the first 128 runs corresponding to the two-level fractional factorial design are used to fit a model containing all nine linear effects plus all 36 bilinear effects. Third, using the last 27 runs from the three-level OA we fit a model with all nine linear effects and all nine quadratic effects. To distinguish the three models we use the design run sizes and refer to them as 155-run model, 128-run model, and 27-run model, respectively.

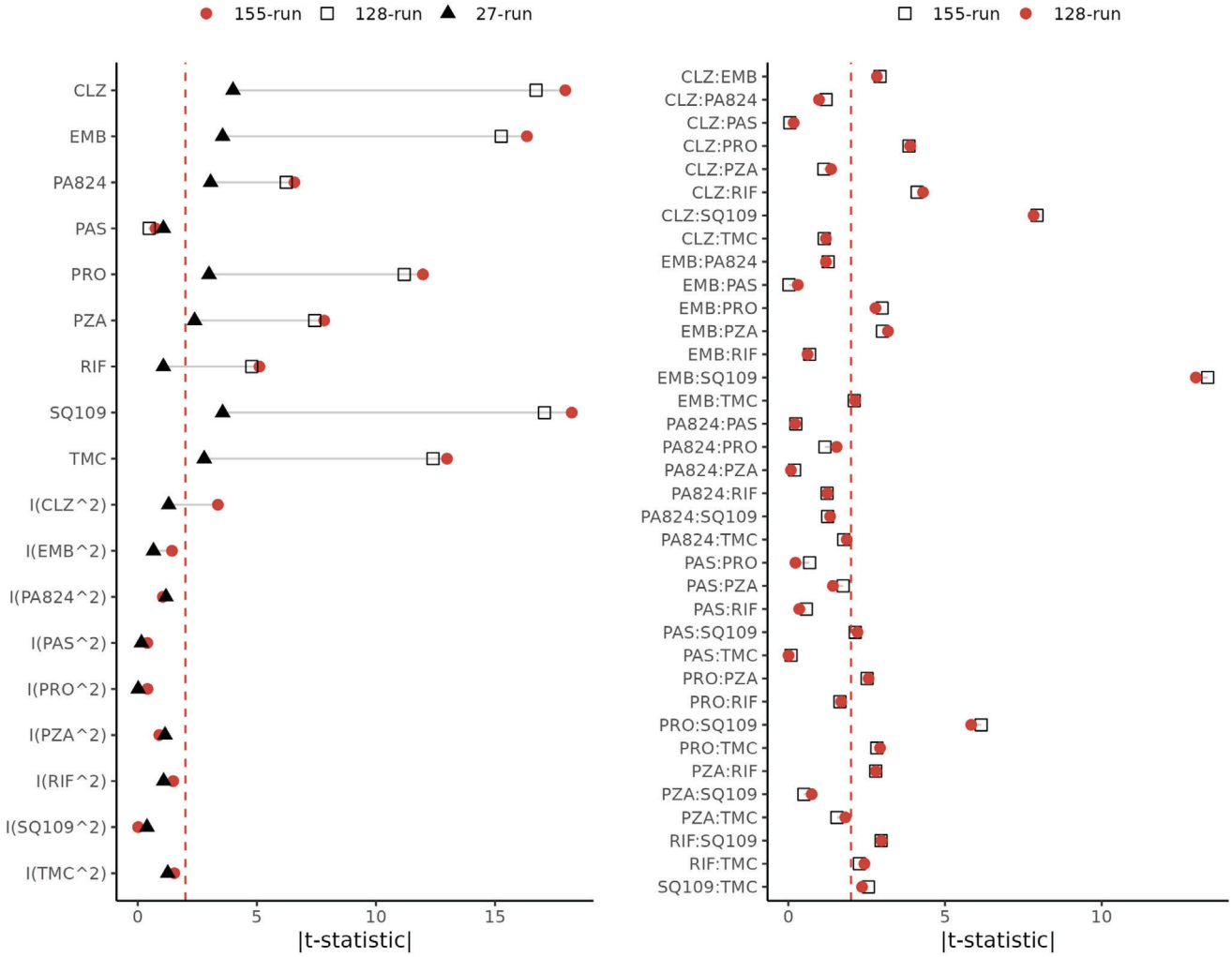

FIGURE 6 Nine-drug experiment $|t|$-statistics

Table 3 presents the parameter estimates and significance levels for these three models. In all three models, the linear effect of each drug is consistent in terms of both sign and magnitude of the parameter estimate. Eight out of the nine drugs are statistically significant at the 5% level in increasing the percent of inhibition in both the 155-run and 128-run models. The drug, PAS, is not statistically significant. The remaining eight drugs all have positive coefficients implying they contribute to higher percentages of inhibition. Of the nine quadratic effects estimated in the 155-run model and 27-run model, only CLZ is statistically significant at the 5% level. While, the sign and magnitude for this effect are consistent in both the 155-run and 27-run model, it is only significant for the full 155-run model. Among the 36 bilinear effects, 14 are identified as statistically significant at the 5% level and consistent in both the 155-run and 128-run models.

Figure 6 presents a comparison of the three models in terms of the $|t|$-statistics, which can be used to further check consistency and significance among the parameter estimates. The vertical axis indicates the corresponding linear, quadratic, and bilinear effects. The horizontal axis represents the $|t|$-statistic. The vertical dashed line represents a $|t|$-statistic of 2.5. In all three models the individual effect of each drug is consistent for the 155-run and 128-run models, with more variation for the 27-run model. The variation in the 27-run model can be explained by the fact that this model only includes linear and quadratic effects, and as a result the estimates have larger standard errors and smaller $|t|$-statistics. Of the nine linear effects, eight of them have a $|t|$-statistic greater than 2.5. The drug, PAS, does not appear to have an effect on the inhibition, which is consistent with the parameter estimates and significance levels provided in Table 3. Of the nine quadratic effects estimated, only CLZ has a $|t|$-statistic above 2.5 in the 155-run model. Among the 36 bilinear effects, 16 are identified with $|t|$-statistics above 2.5 in the 155-run and 128-run models.

TABLE 4 Parameter estimates for the six-drug experiment

	50-Run	32-Run	18-Run
CLZ	0.0556***	0.0566***	0.0875**
EMB	0.0545***	0.0491**	0.0717**
PRO	0.0538***	0.0509**	0.0675*
PZA	0.0782***	0.0716***	0.1183***
RIF	0.0485***	0.0528**	0.0358
TMC	0.0556***	0.0584***	0.0675*
CLZ^2	−0.0147		−0.0542
EMB^2	0.0272		−0.0117
PRO^2	0.0283		0.0158
PZA^2	−0.0184		−0.0217
RIF^2	0.0398		0.0158
TMC^2	0.0219		0.0058
CLZ:EMB	−0.0009	−0.0022	
CLZ:PRO	−0.0032	0.0009	
CLZ:PZA	−0.0117	−0.0122	
CLZ:RIF	−0.0275**	−0.0259	
CLZ:TMC	−0.0007	−0.0016	
EMB:PRO	−0.0037	0.0009	
EMB:PZA	−0.0149	−0.0109	
EMB:RIF	0.0125	0.0116	
EMB:TMC	−0.0079	−0.0053	
PRO:PZA	−0.0229*	−0.0203	
PRO:RIF	0.0029	0.0034	
PRO:TMC	−0.0076	−0.0097	
PZA:RIF	−0.0131	−0.0134	
PZA:TMC	−0.0163	−0.0166	
RIF:TMC	−0.0355***	−0.0341*	
Residual error $\hat{\sigma}$	0.05519	0.06914	0.05915
R^2	0.9471	0.9384	0.9636
Significance levels:	0'***'	0.001'**'	0.01'*' 0.05' '

6.2 | Six-drug 50-run OACD

To further explore the TB drugs that contribute to the highest percentage of inhibition and to identify efficient three- and four-drug combinations, Silva et al[4] perform a follow up experiment in Iteration II such that the dosage for each drug is updated. Since PAS was considered insignificant in the 155-run OACD it is removed from the follow-up experiment. Furthermore, while both SQ109 and PA824 are statistically significant in the initial experiment, they are not considered in the follow-up experiment as they are considered experimental drugs that are still in clinical trials. Thus, a follow-up experiment is performed with the following six drugs: CLZ, EMB, PRO, PZA, RIF, and TMC. Similar to Iteration I, we call these drugs: $x_1, x_2, x_3, x_4, x_5, x_6$. Each drug is tested at three levels: zero dose, half dose, and full dose, and the dosage for each drug is updated such that the concentration of the drug that achieved 20% inhibition in the screening test is used,

TABLE 5 Projected top three- and four-drug combinations using the 50-run OACD

CLZ	EMB	PRO	PZA	RIF	TMC	Predicted inhibition	Standard error
Top three-drug combinations							
1	−1	−1	1	−1	1	0.725	0.0260
−1	1	−1	1	1	−1	0.702	0.0279
1	−1	1	−1	−1	1	0.671	0.0264
−1	1	−1	1	−1	1	0.665	0.0289
Top four-drug combinations							
1	1	−1	1	−1	1	0.834	0.0256
1	−1	1	1	−1	1	0.787	0.0264
1	1	1	−1	−1	1	0.780	0.0260
	0	−1	1	−1	1	0.779	0.0243

see Silva et al[4] for additional details. For model fitting purposes each drug is coded as −1 for the zero dose, 0 for the half dose, and +1 for the full dose.

The design for the 50-run OACD with $k = 6$ is presented in Table 1. This OACD is a combination of a 32-run two-level fractional factorial design, with levels coded as −1 and +1, and an 18-run OA, with levels coded as $-1, 0$, and +1. The 32-run fractional factorial design is a half-fraction of a full 2^6 factorial design with design generator, $x_6 = x_1x_2x_3x_4x_5$. This fractional factorial design is a resolution VI design, which enables the estimation of all six linear effects plus 15 bilinear effects. The 18-run OA is a subset of the commonly used OA$(18, 3^7)$, which is an 18×6 matrix such that all 3^2 level combinations appear equally often in all 18×2 submatrices. This 18-run OA enables the estimate all of six linear plus six quadratic effects. Table S2 in the supplementary material includes the 50-run OACD and data.

Like in the 155-run OACD, we fit three models using the 50-run OACD. Table 4 presents the parameter estimates and significance levels for these three models. Comparing the three models, with the exception of one drug in one model, all six linear effects are statistically significant at the 5% level and consistent across the three models in terms of their magnitude and positive sign. While RIF is statistically significant at the 5% level for the 50-run and 32-run models, it is only significant at the 10% level in the 18-run model. However, the coefficient of RIF is consistent in terms of magnitude and positive sign. None of the quadratic effects appear to have an effect on the inhibition. As for the 15 bilinear effects, three are statistically significant at the 5% level for the 50-run model. Stepwise regression is then performed as a means for variable selection. The final model at the 5% significance level is as follows:

$$\hat{y} = 0.600 + 0.057x_1 + 0.054x_2 + 0.053x_3 + 0.080x_4 + 0.050x_5 + 0.059x_6 - 0.027x_1x_5 - 0.022x_3x_4 - 0.037x_5x_6. \quad (5)$$

All six drugs $x_1, x_2, x_3, x_4, x_5, x_6$ (CLZ, EMB, PRO, PZA, RIF, and TMC) have a positive significant effect on the inhibition. Three bilinear effects between two drugs are statistically significant (CLZ:RIF, PRO:PZA, RIF:TMC), all of which have a negative coefficient. This model results in an R^2 of 0.90.

We use the model in Equation (5) to predict top three- and four-drug combinations, which are of practical interest and importance. Table 5 presents the top three- and four-drug combinations based on predictions from a full 3^6 factorial design. Recall that −1, 0, and +1 represent zero, half, and full dose, respectively. The top three-drug combinations have projected inhibitions between 66.5% and 72.5% whereas the top four-drug combinations all have projected inhibitions greater than 77% and include both CLZ and TMC. None of the top four-drug combinations include RIF.

We compare the predicted results with the experimental inhibitions based on the standard regimens for TB treatment from the 1970s and 1980s reported in Table 5 of Silva et al.[4] The 1970s regimen consisted of the following four drugs: INH (isoniazid), RIF, EMB, and STR (streptomycin), and resulted in 71% of inhibition. The 1980s regimen consisted of the following four drugs: INH, RIF, EMB, and PZA, which resulted in 62% of inhibition, but at the same time shortened the length of treatment by almost half. Thus, all of the top four-drug combinations in Table 5 have projected inhibitions higher than the standard regimens from the 1970s and 1980s. Silva et al[4] performed follow-up experiments to validate the effectiveness of select top three-drug and four-drug combinations.

TABLE 6 Robustness of parameter estimates for the six-drug experiment

	Missing		
	None	Factorial	Axial
Terms	Coef (SE)	Coef (SE)	Coef (SE)
CLZ	0.0556 (0.0086)	0.0608 (0.0078)	0.0556 (0.0088)
EMB	0.0545 (0.0085)	0.0486 (0.0078)	0.0546 (0.0088)
PRO	0.0538 (0.0086)	0.0587 (0.0078)	0.0538 (0.0089)
PZA	0.0782 (0.0086)	0.0840 (0.0079)	0.0783 (0.0089)
RIF	0.0485 (0.0086)	0.0418 (0.0080)	0.0485 (0.0089)
TMC	0.0556 (0.0086)	0.0625 (0.0080)	0.0557 (0.0089)
CLZ^2	−0.0147 (0.0279)	−0.0220 (0.0248)	−0.0148 (0.0286)
EMB^2	0.0272 (0.0279)	0.0176 (0.0249)	0.0272 (0.0285)
PRO^2	0.0283 (0.0275)	0.0253 (0.0243)	0.0283 (0.0282)
PZA^2	−0.0184 (0.0273)	−0.0289 (0.0244)	−0.0184 (0.0280)
RIF^2	0.0398 (0.0277)	0.0500 (0.0247)	0.0398 (0.0283)
TMC^2	0.0219 (0.0268)	0.0230 (0.0237)	0.0219 (0.0275)
CLZ:EMB	−0.0009 (0.0093)	0.0072 (0.0087)	−0.0009 (0.0096)
CLZ:PRO	−0.0032 (0.0093)	−0.0091 (0.0085)	−0.0033 (0.0096)
CLZ:PZA	−0.0117 (0.0092)	−0.0174 (0.0084)	−0.0117 (0.0095)
CLZ:RIF	−0.0275 (0.0093)	−0.0209 (0.0086)	−0.0276 (0.0097)
CLZ:TMC	−0.0007 (0.0090)	−0.0071 (0.0083)	−0.0008 (0.0093)
EMB:PRO	−0.0037 (0.0090)	0.0016 (0.0082)	−0.0037 (0.0093)
EMB:PZA	−0.0150 (0.0090)	−0.0083 (0.0083)	−0.0149 (0.0093)
EMB:RIF	0.0125 (0.0090)	0.0075 (0.0082)	0.0126 (0.0093)
EMB:TMC	−0.0079 (0.0090)	−0.0019 (0.0083)	−0.0078 (0.0093)
PRO:PZA	−0.0229 (0.0091)	−0.0285 (0.0083)	−0.0228 (0.0094)
PRO:RIF	0.0030 (0.0090)	0.0078 (0.0082)	0.0031 (0.0093)
PRO:TMC	−0.0076 (0.0090)	−0.0139 (0.0083)	−0.0075 (0.0093)
PZA:RIF	−0.0131 (0.0090)	−0.0074 (0.0082)	−0.0130 (0.0093)
PZA:TMC	−0.0164 (0.0090)	−0.0233 (0.0084)	−0.0163 (0.0093)
RIF:TMC	−0.0355 (0.0091)	−0.0276 (0.0085)	−0.0354 (0.0094)

For this six-drug 50-run OACD, Table 6 illustrates the robustness of an OACD with one missing observation. The first column represents the factorial effects included in the full second-order model, which contains 6 linear effects, 6 quadratic effects, and 15 bilinear effects. The next six columns provide the parameter estimates and standard errors of the coefficients from the full second-order model based on the OACD with no missing observations, one missing factorial point with the largest Cook's distance (run 10), and one missing axial point with the largest percent of inhibition (run 35). Comparing the parameter estimates from the OACD with no missing observations to those with one missing observation, we can see the robustness of the OACD as missing one observation does not substantially affect the consistency of the parameter estimates. Overall, for the majority of the parameter estimates there is clear consistency among all three columns, with the exception of two bilinear effects, CLZ:PRO and PRO:RIF, which are not identified as statistically significant in Table 4.

7 | CONCLUDING REMARKS

The OACD introduced in the present study served as a good example of combining a two-level factorial design and three-level OA through an application for drug combination determination for TB. We effectively examined single drug effects as well as multi-drug interaction effects, simultaneously. Furthermore, we demonstrated many appealing properties of OACDs such as robustness to missing observations and allowing multiple analyses for cross-validation. The ability to perform cross-validation in biological applications is a valuable property since they often possess internal variances, and thus cross-validation between different models would assure more accurate and reliable conclusions.

We found that generally OACDs are more effective in estimating the parameters than the commonly used CCDs, have more precise model predictions than CCDs, and the overall performance of an OACD is more robust when there are missing observations than the CCD. While OACDs have various advantages, there are other second-order designs in the literature that may provide a smaller number of runs and fit a full second-order model. Examples include, but are not limited to, small composite designs (SCD),[38] augmented pairs designs (APD),[39] definitive screening composite designs,[7] and more. If the run size of a design is of concern, it is important to note that smaller OACDs than those provided in this article can be constructed. Xu et al[6] compare three sizes of OACDs to CCDs, APDs, and SCDs, and explore their properties. Additionally, Xu et al[6] illustrate how an OACD permits the use of resolution III or IV designs for the two-level portion, resulting in an OACD with a smaller number of runs and a higher efficiency for estimating the parameters in a second-order model compared to a CCD. Furthermore, Zhou and Xu[7] compare various sizes of OACDs to many existing and popular designs. Overall, OACDs can be used as an alternative to the popular CCDs for response surface modeling.

ACKNOWLEDGEMENTS

We are grateful to an associate editor and three reviewers for their insightful comments and constructive suggestions. Jaynes was supported by NSF HDR:DSC grant 2123380.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Jessica Jaynes https://orcid.org/0000-0002-6247-2033
Hongquan Xu https://orcid.org/0000-0002-3502-0664
Weng Kee Wong https://orcid.org/0000-0001-5568-3054

REFERENCES

1. World Health Organization. Global tuberculosis report; 2019. http://apps.who.int/iris/bitstream/handle/10665/329368/9789241565714-eng.pdf.
2. Lee BY, Clemens DL, Silva A, et al. Drug regimens identified and optimized by output-driven platform markedly reduce tuberculosis treatment time. *Nature Commun*. 2017;8(1):1-11.
3. Horwitz MA, Clemens DL, Lee BY. AI-enabled parabolic response surface approach identifies ultra short-course near-universal TB drug regimens. *Adv Ther*. 2020;3(4):1900086.
4. Silva A, Lee BY, Clemens DL, et al. Output-driven feedback system control platform optimizes combinatorial therapy of tuberculosis using a macrophage cell culture model. *Proc Nat Acad Sci*. 2016;113(15):E2172-E2179.
5. Vazquez AR, Schoen E, Goos P. Two-level orthogonal screening designs with 80, 96, and 112 runs, and up to 29 factors. *J Qual Technol*. 2021.
6. Xu H, Jaynes J, Ding X. Combining two-level and three-level orthogonal arrays for factor screening and response surface exploration. *Stat Sin*. 2014;24(1):269-289.
7. Zhou YD, Xu H. Composite designs based on orthogonal arrays and definitive screening designs. *J Am Stat Assoc*. 2017;112(520):1675-1683.
8. Zhang XR, Liu MQ, Zhou YD. Orthogonal uniform composite designs. *J Stat Plan Infer*. 2020;206:100-110.
9. Zhang XR, Qi ZF, Zhou YD, Yang F. Orthogonal-array composite design for the third-order models. *Commun Stat Theory Methods*. 2018;47(14):3488-3507.
10. Chen XP, Guo B, Liu MQ, Wang XL. Robustness of orthogonal-array based composite designs to missing data. *J Stat Plan Infer*. 2018;194:15-24.
11. Bailey RA, Cameron PJ. Multi-part balanced incomplete-block designs. *Stat Pap*. 2019;60(2):405-426.
12. Jaynes J, Ding X, Xu H, Wong WK, Ho CM. Application of fractional factorial designs to study drug combinations. *Stat Med*. 2013;32(2):307-318.

13. D'Almeida M, Sire O, Lardjane S, Duval H. Development of a new approach using mathematical modeling to predict cocktail effects of micropollutants of diverse origins. *Environ Res*. 2020;188:109897.
14. Blasiak A, Lim JJ, Seah SGK, et al. IDentif.AI: rapidly optimizing combination therapy design against severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) with digital drug development. *Bioeng Transl Med*. 2021;6(1):e10196.
15. de Mel S, Rashid MBM, Zhang XY, et al. Application of an ex-vivo drug sensitivity platform towards achieving complete remission in a refractory T-cell lymphoma. *Blood Cancer J*. 2020;10(1):1-5.
16. Zhang T, Sharma A, Li Y, Zhou Y, Ding X. Orthogonal Array composite design to study and optimize antioxidant combinations in the prevention of UVB-induced HSF damage. *J Photochem Photobiol B Biol*. 2018;178:568-576.
17. Jia X, Li Y, Sharma A, et al. Application of sequential factorial design and orthogonal array composite design (OACD) to study combination of 5 prostate cancer drugs. *Comput Biol Chem*. 2017;67:234-243.
18. Jaynes J, Zhao Y, Xu H, Ho CM. Use of orthogonal array composite designs to study lipid accumulation in a cell-free system. *Qual Reliab Eng Int*. 2016;32(5):1965-1974.
19. Lim JJ, Goh J, Rashid MBMA, Chow EKH. Maximizing efficiency of artificial intelligence-driven drug combination optimization through minimal resolution experimental design. *Adv Ther*. 2020;3(4):1900122.
20. Zoetemelk M, Ramzy GM, Rausch M, Nowak-Sliwinska P. Drug-drug interactions of irinotecan, 5-fluorouracil, folinic acid and oxaliplatin and its activity in colorectal carcinoma treatment. *Molecules*. 2020;25(11):2614.
21. Zoetemelk M, Ramzy GM, Rausch M, et al. Optimized low-dose combinatorial drug treatment boosts selectivity and efficacy of colorectal carcinoma treatment. *Mol Oncol*. 2020;14(11):2894-2919.
22. Weiss A, Roux-Bourdieu L, Zoetemelk M, et al. Identification of a synergistic multi-drug combination active in cancer cells via the prevention of spindle pole clustering. *Cancers*. 2019;11(10):1612.
23. Liu W, Li YL, Feng MT, et al. Application of feedback system control optimization technique in combined use of dual antiplatelet therapy and herbal medicines. *Front Physiol*. 2018;9:491.
24. Lin J, Jiang H, Ding X. Synergistic combinations of five single drugs from Centella asiatica for neuronal differentiation. *Neuroreport*. 2017;28(1):23-27.
25. Clemens DL, Lee BY, Silva A, et al. Artificial intelligence enabled parabolic response surface platform identifies ultra-rapid near-universal TB drug treatment regimens comprising approved drugs. *PLoS One*. 2019;14(5):e0215607.
26. Wu CFJ, Hamada MS. *Experiments: Planning, Analysis, and Optimization*. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2009.
27. Tang B, Deng LY. Minimum G_2-aberration for nonregular fractional factorial designs. *Ann Stat*. 1999;27(6):1914-1926.
28. Xu H, Wu CFJ. Generalized minimum aberration for asymmetrical fractional factorial designs. *Ann Stat*. 2001;29(4):1066-1077.
29. Tang Y, Xu H. Permuting regular fractional factorial designs for screening quantitative factors. *Biometrika*. 2014;101(2):333-350.
30. Dueck G, Scheuer T. Threshold accepting: a general purpose optimization algorithm appearing superior to simulated annealing. *J Comput Phys*. 1990;90(1):161-175.
31. Kiefer J. Optimum designs in regression problems, II. *Ann Math Stat*. 1961;32:298-325.
32. Kiefer J. General equivalence theory for optimum designs (approximate theory). *Ann Stat*. 1974;2(5):849-879.
33. Farrell R, Kiefer J, Walbran A. Optimum multivariate designs. Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability; Vol. 1, 1967:113-138.
34. Hardin R, Sloane N. A new approach to the construction of optimal designs. *J Stat Plan Infer*. 1993;37(3):339-369.
35. Smucker BJ, Jensen W, Wu Z, Wang B. Robustness of classical and optimal designs to missing observations. *Comput Stat Data Anal*. 2017;113:251-260.
36. Akhtar M, Prescott P. Response surface designs robust to missing observations. *Commun Stat Simul Comput*. 1986;15(2):345-363.
37. Ahmad T, Gilmour SG. Robustness of subset response surface designs to missing observations. *J Stat Plan Infer*. 2010;140(1):92-103.
38. Draper NR, Lin DKJ. Small response-surface designs. *Technometrics*. 1990;32(2):187-194.
39. Morris MD. A class of three-level experimental designs for response surface modeling. *Technometrics*. 2000;42(2):111-121.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Luna J, Jaynes J, Xu H, Wong WK. Orthogonal array composite designs for drug combination experiments with applications for tuberculosis. *Statistics in Medicine*. 2022;1-18. doi: 10.1002/sim.9423

APPENDIX

TABLE A1 Three-level OAs with 9, 18, and 27 runs

OA(9)												
−1		−1		−1		−1						
−1		0		0		1						
−1		1		1		0						
0		−1		0		0						
0		0		1		−1						
0		1		−1		1						
1		−1		1		1						
1		0		−1		0						
1		1		0		−1						
OA(18)												
−1	−1	−1	−1	−1	−1	−1						
−1	0	0	0	0	0	0						
−1	1	1	1	1	1	1						
0	−1	−1	0	0	1	1						
0	0	0	1	1	−1	−1						
0	1	1	−1	−1	0	0						
1	−1	0	−1	1	0	1						
1	0	1	0	−1	1	−1						
1	1	−1	1	0	−1	0						
−1	−1	1	1	0	0	−1						
−1	0	−1	−1	1	1	0						
−1	1	0	0	−1	−1	1						
0	−1	0	1	−1	1	0						
0	0	1	−1	0	−1	1						
0	1	−1	0	1	0	−1						
1	−1	1	0	1	−1	0						
1	0	−1	1	−1	0	1						
1	1	0	−1	0	1	−1						
OA(27)												
−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
−1	−1	0	0	−1	1	0	1	1	−1	0	0	1
−1	−1	1	1	−1	0	1	0	0	−1	1	1	0
−1	0	−1	0	1	0	−1	0	1	0	0	1	−1
−1	0	0	1	1	−1	0	−1	0	0	1	−1	1
−1	0	1	−1	1	1	1	1	−1	0	−1	0	0
−1	1	−1	1	0	1	−1	1	0	1	1	0	−1
−1	1	0	−1	0	0	0	0	−1	1	−1	1	1
−1	1	1	0	0	−1	1	−1	1	1	0	−1	0
0	−1	−1	0	0	0	0	−1	0	0	−1	0	0
0	−1	0	1	0	−1	1	1	−1	0	0	1	−1
0	−1	1	−1	0	1	−1	0	1	0	1	−1	1
0	0	−1	1	−1	1	0	0	−1	1	0	−1	0
0	0	0	−1	−1	0	1	−1	1	1	1	0	−1
0	0	1	0	−1	−1	−1	1	0	1	−1	1	1
0	1	−1	−1	1	−1	0	1	1	−1	1	1	0
0	1	0	0	1	1	1	0	0	−1	−1	−1	−1
0	1	1	1	1	0	−1	−1	−1	−1	0	0	1
1	−1	−1	1	1	1	1	−1	1	1	−1	1	1
1	−1	0	−1	1	0	−1	1	0	1	0	−1	0
1	−1	1	0	1	−1	0	0	−1	1	1	0	−1
1	0	−1	−1	0	−1	1	0	0	−1	0	0	1
1	0	0	0	0	1	−1	−1	−1	−1	1	1	0
1	0	1	1	0	0	0	1	1	−1	−1	−1	−1
1	1	−1	0	−1	0	1	1	−1	0	1	−1	1
1	1	0	1	−1	−1	−1	0	1	0	−1	0	0
1	1	1	−1	−1	1	0	−1	0	0	0	1	−1