





This item is the archived peer-reviewed author-version of:

Design and analysis of a microplate assay in the presence of multiple restrictions on the randomization

Reference:

Bohyn Alexandre, Schoen Eric D., Ng Chee Ping, Bishard Kristina, Haarmans Manon, Trietsch Sebastian J., Goos Peter.- Design and analysis of a microplate assay in the presence of multiple restrictions on the randomization
Journal of agricultural, biological, and environmental statistics / American Statistical Association; International Biometric Society - ISSN 1537-2693 - New york, Springer, 29(2024), p. 745-765
Full text (Publisher's DOI): <https://doi.org/10.1007/S13253-023-00570-1>
To cite this reference: <https://hdl.handle.net/10067/2020210151162165141>

Design and analysis of a microplate assay in the presence of multiple restrictions on the randomization

Alexandre Bohyn ¹, Eric D. Schoen ¹, Chee Ping Ng ², Kristina Bishard², Manon Haarmans², Sebastian J. Trietsch², and Peter Goos ^{1,3}

¹Department of Biosystems, Faculty of Bioscience Engineering, KU Leuven, Leuven, Belgium

²Mimetas BV, Leiden, The Netherlands

³Department of Engineering Management, Faculty of Business and Economics, University of Antwerp, Antwerp, Belgium

October 18, 2023

Abstract

Experiments using multi-step protocols often involve several restrictions on the randomization. For a specific application to *in vitro* testing on microplates, a design was required with both a split-plot and a strip-plot structure. On top of two-level treatment factors and the factors that define the randomization restrictions, a multi-level fixed blocking factor not involving further restrictions on the randomization had to be added. We develop a step-by-step approach to construct a design for the microplate experiment and analyse a response. To consolidate the approach, we study various alternative scenarios for the experiment.

Keywords: Split-plot experiment; Strip-plot experiment; Blocking; Mixed model; Regular two-level designs

1 Introduction

Empirical research and development projects frequently involve two-level screening designs to make out which factors in a design are active and which are not (see the textbooks of, e.g., Box et al., 1978; Mee, 2009; Wu and Hamada, 2009; Montgomery, 2019). One of the strong points of these designs is their run size economy. Indeed, it is feasible to include as many as $N - 1$ factors in N runs, where N is a multiple of four. A further strong point of two-level screening designs is their potential to include restrictions in the randomization and still allow for a simple analysis of the data. There is a substantial body of literature on restrictions involving nested random factors. These restrictions result in blocked designs (see the aforementioned textbooks, as well as Sartono et al., 2015a), row-column designs (Vo-Thanh et al., 2020), split-plot designs (Huang et al., 1998; Bingham and Sitter, 1999, 2001; Bingham et al., 2004; Sartono et al., 2015b), and split-split-plot designs (Schoen, 1999; Jones and Goos, 2009). Other papers address strip-plot designs that involve crossed random factors (see Miller, 1997; Mee and Bates, 1998; Vivacqua and Bisgaard, 2004).

The construction of designs in the presence of a combination of block restrictions, split-plot restrictions and strip-plot restrictions was addressed by Bingham et al. (2008) and Cheng and Tsai (2011). However, it is not clear how to use their approaches when the design requires a multilevel fixed blocking factor, on top of the two-level treatment factors and the factors that define the randomization restrictions. One such a case is the microplate experiment at the core of the present paper. The experiment involved cells grown on plates with eight columns of wells. There were eight two-level treatment factors. Two weeks were reserved for the work, and one of the treatment factors was varied over the weeks. In each week two plates could be processed; the second treatment factor was varied over the plates within a week. Then, eight tubes with extra-cellular matter had to be prepared in each week, while each tube was used on each of the two plates for that week. There were four factors that governed the contents of the tubes. In addition, there were two treatment factors applied to the columns of a plate. Finally, we needed to ensure that all the treatment factors were orthogonal to an eight-level factor indicating the column position on a plate.

The purpose of this paper is to provide a step-by-step approach to design a fractional factorial experiment with multiple randomization restrictions and an extra multilevel factor. In Section 2, we introduce the microplate experiment. In Section 3, we describe the protocol

for the experimental conduct. We then discuss the specific restrictions on the randomization of the experiment and outline a step by step approach to construct a suitable design for the experiment. In Section 4, we analyse the results of the experiment first by identifying active effects, and, second, by building a model that captures the full complexity of the error structure. In Section 5, we revisit the step-by-step approach to consider four alternative scenarios of the experiment. Finally, Section 6, offers some concluding remarks.

2 The microplate experiment

Organ-on-a-chip systems contain miniaturized tissues grown inside microfluidic chips. They have rapidly gained popularity over the last decade because they mirror the morphology of *in vitro* cells, which makes them more physiologically relevant than two-dimensional cell cultures (Sung et al., 2009; van Duinen et al., 2015; Leung et al., 2022). However, until recently, the scalability of these systems was limited, because most setups use a single chip. Furthermore, most systems use a permeable membrane for culturing cells, which can be an additional source of variation on top of the biological variation. Recently, Mimetas BV, a company based in Leiden, The Netherlands, introduced the OrganoPlate[®], a microfluidic multi-chip microplate where the cells are grown directly against an extra-cellular matrix (ECM), without the use of an artificial membrane.

The OrganoPlate[®] consists of a microplate with 64 microfluidic chips, each consisting of six interconnected wells; see Figure 1. The 64 chips are arranged in eight rows and eight columns. The inlay shows that a single chip comprises three channels, the central ECM channel and two perfusion channels, joining in the center of the chip. To prepare a plate for an experiment, the ECM is first prepared independently as a gel and then loaded on the chip through the gel inlet, A2. Once the gel is set, experimenters may choose to study cells on both sides of the ECM or only on one side. To this end, a medium containing cells is prepared and introduced in one or both of the perfusion channels of the chip through the perfusion inlets, A1 and A3 respectively. Once the cells have fully grown, the plate is ready for testing. All the measurements are made in the observation window of a chip, B2. Each plate thus allows for 64 observational units.

The ECM is a complex network of macro-molecules created using a mix of different polymers. The main polymer used in the ECM is type-I collagen because it allows the creation of unique geometries. These geometries mimic three-dimensional (3D) tissues *in vivo* and

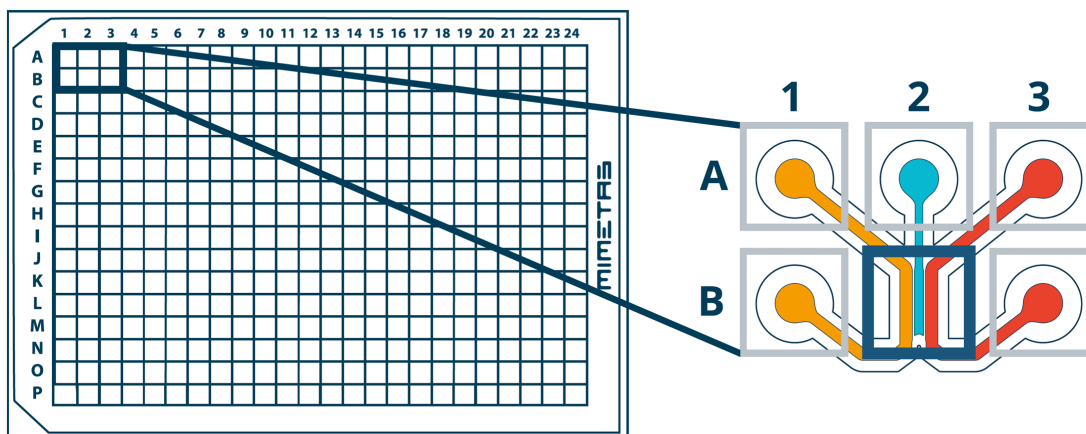


Figure 1: Schematic representation of an OrganoPlate[®] containing 64 chips. The inlay represents a single chip made up of six wells: (A1) left perfusion inlet, (B1) left perfusion outlet, (A2) gel inlet, (B2) observation window, (A3) right perfusion inlet and (B3) right perfusion outlet. Each chip contains three channels: a central extra-cellular matrix channel (blue), and two perfusion channels (red and orange).

provide support to the surrounding cells (Theocharis et al., 2016). However, the physical properties of collagen are strongly influenced by its polymerization conditions, such as pH and temperature (Roeder et al., 2002). In order to make this 3D system a reliable culture platform, it is essential that results are uniform within a plate, as well as reproducible between plates and between laboratories. For this reason, Mimetas wants to detect the factors that affect the quality of their extra-cellular matrix. In particular, the goal of the experimental study is to detect which factors minimize variation in the quality of the ECM. A secondary goal specifically for the OrganoPlate[®] was to establish factor settings that maximize that quality.

3 Design construction

In this section, we describe the protocol for preparing and loading the ECM gel and list the factors considered in this protocol. Then, we detail the randomization restrictions in the experiment and explain how we built a suitable experimental design compatible with all these restrictions.

3.1 Treatment factors

We identified eight factors that could affect the quality of the ECM gel; they are shown in Table 1. The ECM gel is prepared by mixing several ingredients in an Eppendorf tube. Before the actual mixing, the tube is cooled in an ice bath where it can stay for 1 or 30 minutes (factor **h**). The collagen solution that is added to the tube is taken from a large batch, which is aliquoted, i.e., divided into smaller batches. The aliquot used can either come from the start or from the end of the original batch. The origin of the aliquot is denoted by a two-level categorical factor with levels “Early” and “Late” (factor **a**). From each aliquot, either 100 μL or 300 μL of collagen can be added to the mix (factor **b**). Next, the mix is homogenized with a pipette, and this can be repeated either 20 or 50 times (factor **c**). The pH is known to influence the structure of collagen fibers and it can be adjusted by changing the NaHCO_3 concentration in the basic solution. The final pH of the gel is either 7.1 or 8.3 (factor **d**). When the preparation is finished, the mix needs to rest for 10 or 60 minutes (factor **g**) before being pipetted into the plate. Finally, once the gel has been loaded into the plate and has had time to polymerize, 50 μL of Hanks’ Balanced Salt Solution (HBSS), with either low or high levels of Ca^{++} and Mg^{++} (factor **e**), is added in the gel inlet to prevent the gel from drying. Before seeding the cells, the HBSS can either be removed or not. This is studied using another two-level factor with levels “No” and “Yes” (factor **f**).

Table 1: Experimental factors and their settings.

Stage	Label	Factor	Settings		Applied to
Mixing	a	Aliquot	Early	Late	Tube
	b	Collagen gel volume (μL)	100	300	
	c	Amount of mixing	20	50	
	d	pH of solution	7.1	8.3	
	g	Starting time (min)	10	60	Plate
	h	Time on ice (min)	1	30	Week
Post-Loading	e	HBSS Ca^{++} and Mg^{++} content	Low	High	Column
	f	HBSS removal	No	Yes	

Week	Plate	Tube															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	•	•	•	•	•	•	•	•								
	2	•	•	•	•	•	•	•	•								
2	3									•	•	•	•	•	•	•	•
	4									•	•	•	•	•	•	•	•

Figure 2: Schematic representation of the distribution of the tubes over the weeks and the plates. A black bullet in row i and column j means that the j th tube was used for the i th plate. Plates 1 and 2 are processed in week 1, plates 3 and 4 are processed in week 2.

3.2 Restrictions on the randomization

To perform the experiment, four plates were available. There are 64 chips per plate, resulting in a total of 256 individual chips. For logistic reasons, the factors \mathbf{g} and \mathbf{h} could only be varied between the plates. Since the ECM gel is prepared in an Eppendorf tube, factors \mathbf{a} , \mathbf{b} , \mathbf{c} and \mathbf{d} were all applied to an individual tube. However, the ECM gel is loaded on the plate using a micro-pipette that fills an entire column of a plate at a time, so that these factors can only be varied over an entire column. The same is true for factors \mathbf{e} and \mathbf{f} , but since these factors are applied in the post-loading stage, they are not applied to individual tubes but to the columns of the plates. Therefore, only eight different combinations of the factors \mathbf{a} to \mathbf{f} can be applied to a single plate, so that the experiment involves 32 different treatments combinations or experimental runs.

The manual processing of a single plate with many treatment combinations is time consuming. Therefore, only two plates can be processed in a single week so that the experiment involves two weeks and the plates are nested within the weeks. Since the ECM gel that is loaded into all chips within a column of a plate is prepared in an Eppendorf tube, a natural setup would have been to use 32 different tubes, one for each combination of column and plate. However, only eight tubes could be prepared per week, so that each tube had to be used twice in a given week. Our solution was to cross the tubes with the plates and nest them within the weeks. We discuss the option of using the tubes twice on the same plate in Section 5.2. The diagram in Figure 2 summarizes the experimental layout.

It was logistically convenient to vary factors \mathbf{g} and \mathbf{h} between entire plates, with one factor varied between the weeks and another varied between the two plates within a week.

Therefore, the four combinations of the levels of **g** and **h** define the four plates, and the main effect of **g**, the main effect of **h**, and the interaction effect **gh** are confounded with the random variation between plates or weeks. This is unavoidable given the randomization restrictions. Of the six remaining factors, four (**a** to **d**) are varied between tubes, while the other two, **e** and **f**, are varied between the columns of a plate.

In the design of experiments jargon, the fact that plates are nested within weeks and that one factor is applied to weeks and another is applied to individual plates creates a split-plot design. Here, the weeks are the whole plots, and within each whole plot, both the tubes and the plates are subplots. On top of this, having the plates crossed with the tubes results in a strip-plot design in each of the two weeks. Therefore, the design is a split-strip-plot design.

A common source of unwanted variability in microplate assays is the “edge effect”, due to which the measurements obtained from chips on the edge of a plate may differ from those obtained from chips in the center of the plate. An edge effect can be caused by heterogeneous experimental conditions over the plate, such as a thermal gradient or a difference in the evaporation rate. To minimize the impact of edge effects on the estimates of the factors’ effects, we want to ensure that, over the four plates, the levels of the factors **a**, **b**, **c**, **d**, **e** and **f** are balanced across the eight column positions. The experiment should therefore be blocked in eight blocks of size four.

3.3 Treatment design

Our challenge is to create a 32-run design involving the eight two-level factors in Table 1, and one eight-level blocking factor corresponding to the column positions within the plates, obeying all logistical restrictions. To create the design, we proceeded in four steps: (1) creating the design for the six factors **a** to **f** varied over the columns of a plate, (2) arranging the design in eight blocks corresponding to the eight column positions of a plate, (3) adding the two remaining factors **g** and **h** varied over the plates and the weeks, respectively, and (4) ensuring that in each week, eight different tubes are used.

In step 1, we select the resolution-VI 2^{6-1} design for the six factors **a** to **f** with **f** = **abcde** as the generator to define the treatments for the factors varied between the columns of a plate.

In step 2, we arrange the 32 treatments of the 2^{6-1} design into eight blocks, such that

the number of two-factor interactions that are confounded with the blocks is minimized. The eight-level blocking factor can be defined by taking three independent factorial effects of the 2^{6-1} design. This confounds the three effects, the products of each pair of these effects, and the product of all three effects with the blocks, making up a total of seven confounded effects. We call these seven effects “pseudo-factors”. Choosing **ab**, **ce** and **acf** (which is aliased with **bde**) as the independent effects minimizes the number of two-factor interactions confounded with the blocks (Xu and Lau, 2006). Table 2 shows the aliasing of each of the seven resulting pseudo-factors with two- and three-factor interactions among the factors **a** – **f**. The pseudo-factors **p₁**, **p₂** and **p₃** correspond to the three independent factorial effects, **p₄** = **p₁p₂**, **p₅** = **p₁p₃**, **p₆** = **p₂p₃** and **p₇** = **p₁p₂p₃**. The column positions, determined by the grouping scheme presented in Table 3, are now orthogonally blocked for the main effects of factors **a** – **f**.

Table 2: Aliasing of the pseudo-factors **p_i** ($i = 1, \dots, 7$), defining the column positions, with interactions among the six two-level factors in the 2^{6-1} design.

Pseudo-factor	Aliasing
p₁	ab
p₂	ce
p₃	acf + bde
p₄	df
p₅	ade + bcf
p₆	aef + bcd
p₇	acd + bef

In step 3, we use the OA package of Eendebak and Vazquez (2019) to generate all possible ways of adding an orthogonal four-level factor to the design with one eight-level blocking factor and six two-level factors, with the four levels of the new factor corresponding to the four plates and thus to the four level combinations of factors **g** and **h**. We obtained three different designs where the effects of the six two-level factors are either orthogonal to or completely confounded with the new four-level factor. Table 4 shows the assignments that minimize the aliasing of the main effects of factors **g** and **h** with the two- or three-factor interactions among the other factors. Option 1 is the best of the three because none of the

Table 3: Grouping scheme used to assign treatment combinations to the eight column positions of a plate.

$\mathbf{p}_1 = \mathbf{ab}$	$\mathbf{p}_2 = \mathbf{ce}$	$\mathbf{p}_3 = \mathbf{acf}$	Column position
–	+	–	1
–	–	+	2
–	–	–	3
–	+	+	4
+	–	–	5
+	–	+	6
+	+	–	7
+	+	+	8

main effects of \mathbf{g} or \mathbf{h} are aliased with a two-factor interaction.

In the final step of the design creation, we need to assign each of the 16 possible level combinations of factors $\mathbf{a} - \mathbf{d}$ to the 16 tubes available and ensure that eight of the 16 tubes are allocated to one week and the rest to the other week. The design in Step 1 includes all 16 level-combinations of the factors $\mathbf{a} - \mathbf{d}$, so that the experiment requires 16 tubes. To define the weeks we can either use the factor $\mathbf{g} = \mathbf{ace} + \mathbf{bdf}$ or the factor $\mathbf{h} = \mathbf{abc} + \mathbf{def}$. If we use factor \mathbf{g} , we do not invoke further restrictions on the tube factors $\mathbf{a} - \mathbf{d}$, so that each of the 16 tubes is used once each week. However, if we use factor \mathbf{h} , the eight tubes for which $\mathbf{abc} = -1$ are used twice within the first week, and the eight remaining tubes are used twice within the second week. Thus, using the factor \mathbf{h} to identify the weeks and \mathbf{g} to identify the plates within a week results in the required eight tubes per week. This is shown in Table 5. Finally, the full experimental design is presented in Table 6.

Based on the 2^{6-1} design for factors \mathbf{a} to \mathbf{f} , and the aliasing of the pseudo-factors \mathbf{p}_i ($i = 1, \dots, 7$) and the factors \mathbf{g} and \mathbf{h} with the factors of the 2^{6-1} design, we can compute the aliasing between all the main effects, the two-factor interactions and the pseudo-factors corresponding to the column positions. The full aliasing pattern is presented in Table 7. Furthermore, since we know the allocation of the treatments to the tubes, the plates, and the weeks, we can group the factorial effects according to their treatment unit and thus

Table 4: Aliasing of factors **g** and **h** and their interaction **gh** for each of the three options for adding a four-level factor to the 2^{6-1} design.

Option	Factor	Aliasing
1	g	ace + bdf
	h	abc + def
	gh	be
2	g	cd
	h	ad
	gh	ac
3	g	ef
	h	ad
	gh	ce

according to the random error they are subjected to. The resulting groups are called “strata” and there are four of them in this experiment: *Week*, *Plate*, *Tube*, and *Unit*. Both *Plate* and *Tube* are nested within *Week*, while *Plate* is crossed with *Tube*. Finally, *Unit* corresponds to the smallest treatment unit, the column within a plate.

Table 5: Definition of the tube numbers using the coded levels of the factors **a**, **b**, **c** and **d**. The vertical dashed line shows the two subdesigns used in the two weeks.

Week	1								2							
Tube	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
a	-	-	-	-	+	+	+	+	-	-	-	-	+	+	+	+
b	-	-	+	+	-	-	+	+	-	-	+	+	-	-	+	+
c	-	-	+	+	+	+	-	-	+	+	-	-	-	-	+	+
d	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
h	-								+							

Table 6: Final 32-run experimental design, along with the corresponding week, plate, tube, and column for each experimental run.

Week	Plate	Column	Tube	a	b	c	d	e	f	g	h
1	1	1	3	-	+	+	-	+	+	-	-
1	1	2	6	+	-	+	+	-	+	-	-
1	1	3	5	+	-	+	-	-	-	-	-
1	1	4	4	-	+	+	+	+	-	-	-
1	1	5	7	+	+	-	-	+	+	-	-
1	1	6	8	+	+	-	+	+	-	-	-
1	1	7	1	-	-	-	-	-	-	-	-
1	1	8	2	-	-	-	+	-	+	-	-
1	2	1	6	+	-	+	+	+	-	+	-
1	2	2	3	-	+	+	-	-	-	+	-
1	2	3	4	-	+	+	+	-	+	+	-
1	2	4	5	+	-	+	-	+	+	+	-
1	2	5	2	-	-	-	+	+	-	+	-
1	2	6	1	-	-	-	-	+	+	+	-
1	2	7	8	+	+	-	+	-	+	+	-
1	2	8	7	+	+	-	-	-	-	+	-
2	3	1	12	-	+	-	+	-	-	-	+
2	3	2	13	+	-	-	-	+	-	-	+
2	3	3	14	+	-	-	+	+	+	-	+
2	3	4	11	-	+	-	-	-	+	-	+
2	3	5	16	+	+	+	+	-	-	-	+
2	3	6	15	+	+	+	-	-	+	-	+
2	3	7	10	-	-	+	+	+	+	-	+
2	3	8	9	-	-	+	-	+	-	-	+
2	4	1	13	+	-	-	-	-	+	+	+
2	4	2	12	-	+	-	+	+	+	+	+
2	4	3	11	-	+	-	-	+	-	+	+
2	4	4	14	+	-	-	+	-	-	+	+
2	4	5	9	-	-	+	-	-	+	+	+
2	4	6	10	-	-	+	+	-	-	+	+
2	4	7	15	+	+	+	-	+	-	+	+
2	4	8	16	+	+	+	+	+	+	+	+

Table 7: Aliasing of main effects, two-factor interactions, and pseudo-factors of the column positions.

<i>Week</i>	<i>Plate</i>	<i>Tube</i>	<i>Unit</i>
h	g	a	e
	gh + be	b	f
		c	p₂ + ag + ce
		d	p₃
		p₁ + ch + ab	p₄ + bg + df + eh
		p₆	p₅
		p₇	ae + cg
		ac + bh + eg	af
		ad	bf + dg
		ah + bc	cf
		bd + fg	de + fh
		cd	
		dh + ef	

4 Analysis and Interpretation

In this section, we describe the response variable used in the analysis. Next, we present our analysis strategy. Since we have eight factors, a model with an intercept, all main effects and all interactions would contain 37 terms. However, only 32 runs are available. For this reason, we have to run our analysis in two steps. We first identify the active effects in each error stratum. Next, we use the active effects to build the final model and interpret the results.

4.1 Response variable

An important characteristic of the ECM in the central channel of a chip is the fibrosity of the collagen network in the matrix. To assess the fibrosity of the ECM in this experiment, a small sample of the ECM was collected in the central channel of each chip, and its fibrosity was measured using texture analysis (Mostaço-Guidolin et al., 2013), that results in a dimensionless response. A high measured value implies a low fibrosity so that the response is inversely related to the fibrosity of the chip. In the context of the OrganoPlate[®], low fibrous or even non-fibrous ECM is desirable, which translates to a high response value. The main goal of this experiment is to understand which factors are the most influential on the fibrosity, in order to limit the variation in the measured response.

The fibrosity was measured in 256 chips. However, in 17 of them, the ECM gel did not fill the central channel correctly. Therefore, these chips were excluded from the analysis. The final data set thus contains 239 observations, and the response ranges from 244 to 378, with a mean value of 329. The complete data set containing the fibrosity measurements for all 256 chips, along with the experimental design, is available in the supplementary material of this paper. The code used for the analysis and to generate the figures is available in the Github repository https://github.com/ABohynDOE/microplate_assays.

4.2 Phase I: Identification of the active effects

Among the eight rows present on all plates, there were two rows that did not contain excluded chips. We averaged the fibrosity response over these two rows for all 32 columns of the experiment, and use these 32 averages to identify the active factor effects in our analysis. This approach was chosen to ensure that the 32 averages were generated from

the same number of observations.

Table 8 shows the effect sizes for fibrosity per stratum, in decreasing order of their absolute value. For the *Tube* and *Plate* strata, the table shows an estimate of the PSE(50) (Schoen and Kaul, 2000) robust standard error of the effects, along with a threshold to declare an effect active ($\alpha = 10\%$). Since the *Week* and *Plate* strata include only one and two degrees of freedom, respectively, it is not feasible to obtain a robust standard error estimate for these strata and a threshold to identify active effects. We also see that the robust standard error for the *Tube* stratum is lower than for the *Unit* stratum. This suggests that the random variation between the columns of a plate dominates the variation between the different tubes used within a week.

In view of their size, we tentatively assume that the factors \mathbf{g} and \mathbf{h} and their interaction \mathbf{gh} are active. The estimates for \mathbf{cd} , \mathbf{a} , \mathbf{d} , \mathbf{ah} and \mathbf{c} are statistically significant in the *Tube* stratum, and the estimates for \mathbf{p}_5 and \mathbf{p}_3 are statistically significant in the *Plate* stratum, implying that the column position is also significant. Table 7 shows that the significant two-factor interaction \mathbf{ah} is aliased with the two-factor interaction \mathbf{bc} . However, given the effect sizes of the corresponding main effects, \mathbf{ah} is more likely to be active than \mathbf{bc} , due to the heredity principle (Wu and Hamada, 2009, Chap. 8)

4.3 Phase II: Model building and data analysis

Now that we have identified the active effects, we can build a model for the data at the individual chip level. This has the additional benefit that we can assess the effect of the row positions on the fibrosity.

The experiment was run over two weeks. We use the index i , taking the values 1 or 2, to indicate the weeks. Within each week, there are two plates and eight tubes. We use the index j , taking the values 1 or 2, to indicate the plates within a week, and the index k , taking the values 1 to 8, to indicate the tubes within a week. However, each plate and tube combination within a week also corresponds to a specific column of the plate (see the full design in Table 6 as well as Figure 2). Therefore, the eight columns of a plate are defined by the indices i , j , and k . Each plate also contains eight rows, for which we use the index l , with $l = 1, \dots, 8$, to define the eight rows within a plate. As a result, the 256 chips used in the experiment can be uniquely identified by the indices i , j , k , and l , and we can use Y_{ijkl} to denote the fibrosity measured in week i , on plate j , in column ijk and row l , with

Table 8: Effect sizes for fibrosity per stratum. The effects are arranged in decreasing order of absolute size. The figures in parentheses are the robust standard error and the critical value corresponding to a 10% significance level, for the *Tube* and *Unit* strata.

<i>Week</i>		<i>Tube</i> (0.9, 1.55)		<i>Unit</i> (1.32, 2.26)	
Effect	Estimate	Effect	Estimate	Effect	Estimate
h	-13.83	cd	-5.45	p₅	5.2
<i>Plate</i>		a	-4.89	p₃	-2.77
Effect	Estimate	d	-4.55	cgh	-1.95
gh	16.27	ah	-3.02	cg	-1.33
g	5.39	c	2.27	cf	-1.27
		fg	-1.45	agh	1.02
		p₇	1.42	adg	0.92
		b	0.77	fh	0.89
		eg	-0.61	f	-0.8
		ad	0.58	p₂	0.7
		p₁	-0.55	p₄	-0.58
		dh	0.39	af	-0.42
		p₆	-0.3	dg	-0.39
		afg	-0.27	e	-0.14

an ECM gel coming from tube ik . The appropriate model for the fibrosity has the form

$$Y_{ijkl} = \beta_0 + \mathbf{x}'_{ijkl}\boldsymbol{\beta} + \delta_i + \gamma_{ij} + \lambda_{ik} + \phi_{ijk} + \rho_{ijl} + \epsilon_{ijkl}, \quad (1)$$

where β_0 is a fixed intercept, \mathbf{x}_{ijkl} represents the model expansion of the levels of the treatment factors for chip $ijkl$, $\boldsymbol{\beta}$ is a vector of fixed parameters, $\delta_i \sim N(0, \sigma_\delta^2)$ is the random effect of the weeks, $\gamma_{ij} \sim N(0, \sigma_\gamma^2)$ is the random effect for the plates, $\lambda_{ik} \sim N(0, \sigma_\lambda^2)$ is the random effect for the tubes, $\phi_{ijk} \sim N(0, \sigma_\phi^2)$ is the random effect for the columns of the plates, $\rho_{ijl} \sim N(0, \sigma_\rho^2)$ is a random effect for the rows of the plates and $\epsilon_{ijkl} \sim N(0, \sigma_\epsilon^2)$ is the random error of the individual measurement in each chip. To estimate the variance components of the linear mixed model in Equation (1), we use restricted maximum likelihood (see Goos et al. (2006) for more details).

The vector \mathbf{x}_{ijkl} , representing the model expansion of the levels of the treatment factors for chip $ijkl$, contains the levels of the factors **a**, **c**, **d**, **g** and **h** and the cross-products corresponding to the interactions **ah**, **cd** and **gh**. To account for the potential edge effects, we also add the main effects of the eight column positions, represented by **p**₁ to **p**₇, and the main effects of the eight row positions using **q**₁ to **q**₇.

There are five variance components, two of which (σ_δ^2 and σ_γ^2) are not estimable because the corresponding strata do not contain enough degrees of freedom. Estimates of the three estimable variance components, the tube variance σ_λ^2 , the column variance σ_ϕ^2 , and the error variance σ_ϵ^2 , are presented in Table 9. The total variance is dominated by the error component, σ_ϵ^2 . This suggests that the major source of random variation is related to measurements at the individual chip level or to the variation among these chips.

Using F-tests, we compute p-values for the fixed effects; see Table 10. Since effects **g**, **h**, and **gh** are not testable, they are not shown in the table. We see that, except for the factor **row**, all effects are significant at the 10 % level.

The small p-value for the column factor and the large p-value for the row factor suggests that an edge effect only seems to occur over the columns of a plate. To visualize the magnitude of the effects of the row and column positions, we compute the fitted means for each row and column position, along with the least significant difference (LSD) at the 10% level; see Figure 3. None of the pairs of fitted means of the row positions are significantly different. It confirms that there does not seem to be edge effects over the rows of a plate. However, there are significant differences between the means of the column positions, although there

Table 9: Point estimates and standard errors of the five variance components in the model in Equation (1).

Stratum	Variance component	Estimate	Standard error
Week	σ_{δ}^2	–	–
Plate	σ_{γ}^2	–	–
Tube	σ_{λ}^2	2.5	12.1
Column	σ_{ϕ}^2	6.2	13.6
Row	σ_{ρ}^2	1.9	8.3
Error	σ_{ϵ}^2	179.8	18.7

Table 10: Significance test for the fixed effects.

Effect	DF num.	DF den.	F statistic	p-value
a	1	5.8	8.59	0.027
c	1	5.9	3.88	0.097
d	1	5.9	10.97	0.017
cd	1	5.9	36.13	0.001
ah	1	5.8	10.71	0.018
row	7	20.1	0.46	0.850
column	7	9.9	3.17	0.049

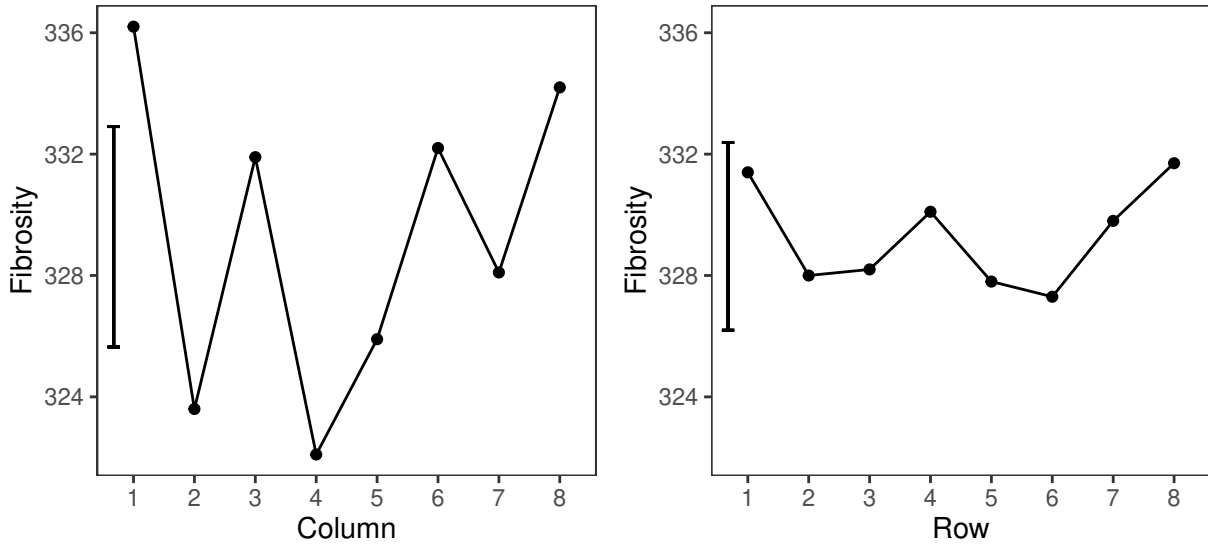


Figure 3: Fitted means of the column and row positions, with an error bar representing the least significant difference ($\alpha = 10\%$).

is no clear division between the edge columns, 2 and 23, and the rest.

Interaction plots for the interactions **cd**, **ah** and **gh** are shown in Figure 4, along with the LSD at the 10% level for **cd** and **ah**. The first interaction, shown in Figure 4a, is **gh**, i.e., the interaction between starting time and time on ice. This interaction is clearly the strongest of all three interactions. However, since it is estimated in the *Plate* stratum, we cannot determine an LSD for that interaction. We see that a starting time of ten minutes combined with a time on ice of one minute maximizes the fibrosity response. However, we see that a starting time of one minute makes the response highly sensitive to the level of time on ice. So, a starting time of 60 minutes minimizes the effect of time on ice on the fibrosity response. The second interaction, shown in Figure 4b, is **ah**, i.e., the interaction between aliquot and time on ice. We see that the level of the factor aliquot only matters if the time on ice is 30 minutes, and that a time on ice of one minute maximizes the fibrosity response, as in Figure 4a. This is in line with the results presented in Table 8, where the estimated effect size of time on ice is strongly negative.

Finally, for the third interaction **cd** shown in Figure 4c, i.e., the interaction between amount of mixing and pH, we see that a pH of 7.1 in combination with an amount of mixing of 50 stirrings maximizes the fibrosity response but a pH of 8.3 limits the variation in that response.

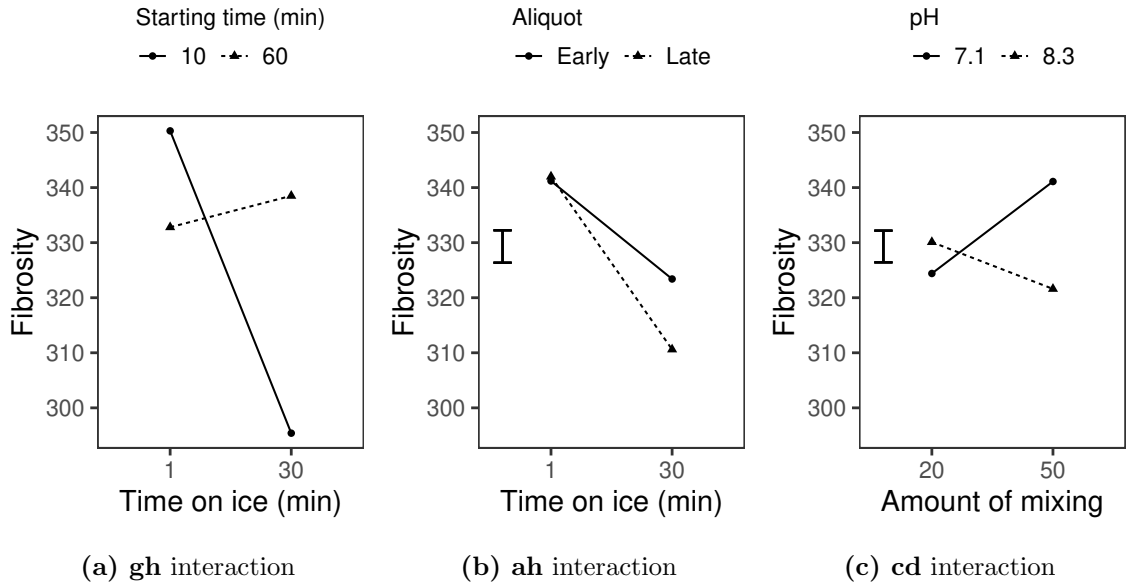


Figure 4: Interaction plots with their least significant difference ($\alpha = 10\%$) represented by error bars on the left. No LSD can be determined for the first interaction since the variance components σ_{δ}^2 and σ_{γ}^2 cannot be estimated.

4.4 Conclusions on the fibrosity response

The goal of this study was to detect which factors minimize variation in the fibrosity of the ECM. A secondary goal specifically for the OrganoPlate[®] was to establish factor settings with high values of the response (this minimizes fibrosity; see Section 4.1). Based on our analysis, we recommend two treatment combinations, depending on which of the two goals is more important. To minimize the variation in the fibrosity, use a time on ice of one minute, starting time of 60 minutes, an early aliquot, a pH of 8.3, and mix the preparation 20 times. To maximize the fibrosity response use a time on ice of one minute, a starting time of ten minutes, an early aliquot, pH of 7.1, and mix the preparation 50 times. We observed a strong effect of the column positions on the results, which should be taken into account when performing assays. For this reason, it is important to replicate treatment combinations over different column positions.

5 Alternative scenarios

In this section, we consider four alternative scenarios for the experiment. For each scenario, we justify why it could make sense, describe how the randomization restrictions of the experiment would change, and explain which experimental design we would consider. We provide the final design for each alternative scenario in the supplementary materials.

The four scenarios considered are:

1. Each of the 16 available tube is used twice on a single plate instead of once on each plate within a week. Using a tube twice on a single plate creates a split-split-split-plot structure instead of the current split-strip-plot structure.
2. Thirty-two instead of 16 tubes are available so that each individual column would get its own tube instead of sharing a tube with a column on another plate.
3. All factors can be varied within the plates. Currently, factors \mathbf{g} and \mathbf{h} are varied between the plates and the weeks, respectively, because of logistic restrictions. Without these restrictions, all factors could be varied from column to column.
4. The blocking factor corresponding to the column positions has four levels. It is conceivable that the columns at the same distance of the left- and right-most edges of the plate are alike. In that case, a potential edge effect would be symmetrical and a four-level blocking factor would be sufficient to capture the effect of the column positions.

Before considering the four alternative scenarios, we summarize the process used to generate the experimental design for the actual experiment.

5.1 Initial design

The initial design generation process can be summarized in four steps (see Section 3.3):

1. Find a 2^{6-1} design involving 32 runs for the factors \mathbf{a} to \mathbf{f} . For this purpose, we chose the minimum aberration design with $\mathbf{f} = \mathbf{abcde}$.
2. Arrange the design in eight orthogonal blocks corresponding to the eight column positions on a plate using the blocking scheme of Xu and Lau (2006). In our case, the three defining pseudo-factors were $\mathbf{p}_1 = \mathbf{ab}$, $\mathbf{p}_2 = \mathbf{ce}$ and $\mathbf{p}_3 = \mathbf{acf}$.

3. Add a four-level factor orthogonal to factors $\mathbf{a} - \mathbf{f}$ and to the eight-level blocking factor, whose two independent pseudo-factors define the two factors \mathbf{g} and \mathbf{h} . We considered all options with regular aliasing and selected the one that minimized the aliasing between the interactions among the six two-level factors \mathbf{a} to \mathbf{f} , and between the main effects and two-factor interactions of the factors \mathbf{g} and \mathbf{h} . In our case, the two additional factors were defined by the relations $\mathbf{g} = \mathbf{ace}$ and $\mathbf{h} = \mathbf{abc}$.
4. Ensure that eight tubes per week are used. In our case, this was done by selecting the factor \mathbf{h} to define the weeks.

We now discuss how each of the four steps change when we consider an alternative scenario, along with the allocation of the factorial effects to the different error strata.

5.2 Scenario 1: tubes are used twice on each plate

If each tube were used twice on a single plate, the logistic restrictions concerning the blocks for the eight column positions and the factors \mathbf{g} and \mathbf{h} would still be in place. This would imply that steps 1 to 3 of the design generation process would be the same as for the initial design, and only the assignment of the tubes to the plates and weeks in Step 4 would change, because we now need to divide the 16 tubes into groups corresponding to the four plates. We would still use $\mathbf{h} = \mathbf{abc}$ to define the weeks, but, instead of using the combinations of factors \mathbf{a} , \mathbf{b} , \mathbf{c} and \mathbf{d} as the four tube factors (see Table 5), we would use the factors \mathbf{a} , \mathbf{b} , \mathbf{c} and \mathbf{e} . Using these four factors to define the tubes would mean that the tubes are confounded with the weeks (defined by $\mathbf{h} = \mathbf{abc}$) and with the plates (defined by $\mathbf{g} = \mathbf{ace}$), so that the 16 tubes would be divided in four groups of four tubes corresponding to the four plates.

Under this scenario, the *Week* and *Plate* strata would still contain the factors \mathbf{g} , \mathbf{h} and their interaction \mathbf{gh} . However, the *Tube* stratum would contain 12 degrees of freedom instead of 14, and the *Unit* stratum would contain 16 degrees of freedom instead of 14. This would make the effects in the *Tube* stratum harder to detect. The effects contained in the *Tube* and the *Unit* strata are presented in the first panel of Table 13, along with the aliasing between the main effects and two-factor interactions of the eight factors \mathbf{a} to \mathbf{h} and the pseudo-factors for the columns positions.

5.3 Scenario 2: 32 tubes available

If 32 tubes were available, the logistic restrictions concerning factors \mathbf{g} and \mathbf{h} would still be in place, and steps 1 to 3 of the design generation process would be the same as for the current design. Now that 32 tubes are available, we need to ensure that all 16 level combinations of the tube factors $\mathbf{a} - \mathbf{d}$ are used within each week. In the original design, $\mathbf{h} = \mathbf{abc}$ was used to define the weeks. This imposed a restriction on the level combinations of the tube factors \mathbf{a} , \mathbf{b} and \mathbf{c} , so that half of them were associated with one week and the other half with the other week. For a design with 32 tubes, we use $\mathbf{g} = \mathbf{ace}$ to define the weeks, so that no restriction on level combinations of tube factors is invoked. The factor \mathbf{h} now defines the plates within a week.

The new arrangement does not change the aliasing of the treatment factors and their interactions, but it changes the randomization structure of the experiment. As there are now three nested error strata, this is a split-split-plot design. Indeed, the *Tube* stratum is merged into the *Unit* stratum because no factors are varied only at the tube level; they are all varied between weeks, plates, or column positions, respectively. The table for the aliasing of the main effects, two-factor interactions, and pseudo-factors, and their allocation to the error strata, is similar to Table 7, but with the columns *Tube* and *Unit* merged. In principle, the 28 effects in the *Unit* stratum would make for a more efficient robust standard error estimate compared to the case of 14 effects. However, the effects in that stratum now have a variance that is larger than before, but smaller than the error variance of the effects in the *Tube* stratum of the original design. This is an advantage for the detection of effects in the former *Tube* stratum and a disadvantage for the effects in the former *Unit* stratum.

5.4 Scenario 3: no logistic restrictions on the treatment factors

In this scenario, all factors can be varied from column to column on each plate. In Step 1, we would consider a 2^{8-3} design for factors \mathbf{a} to \mathbf{h} . More specifically, we would use the minimum aberration design with $\mathbf{f} = \mathbf{abcd}$, $\mathbf{g} = \mathbf{abe}$, and $\mathbf{h} = \mathbf{ace}$. To block the design in Step 2, we would use the blocking scheme of Xu and Lau (2006) with $\mathbf{p}_1 = \mathbf{abc}$, $\mathbf{p}_2 = \mathbf{ad}$ and $\mathbf{p}_3 = \mathbf{ae}$. In Step 3, we need to add a four-level factor defining the four plates and the two weeks. The four levels of the factor can be decomposed in two independent pseudo-factors \mathbf{b}_1 and \mathbf{b}_2 , and we define their interaction as $\mathbf{b}_3 = \mathbf{b}_1\mathbf{b}_2$. Clearly, we again need to consider all possible ways of adding a four-level factor. The three options with regular

aliasing (effects are either fully aliased or orthogonal) are shown in Table 11. Option 1 is the best because only three to-factor interactions are aliased with pseudo-factors for the column positions, as opposed to four and six for Options 2 and 3. Now, the allocation of the 32 treatment combinations to the columns are defined through the eight levels of the eight-level blocking factor, but we still need to assign the experimental runs to the four plates, the two weeks, and the 16 tubes. To restrict the number of tubes to eight per week, we still need a half-fraction of the 2^4 design for factors **a**, **b**, **c** and **d** inside each week. Defining the weeks with the second pseudo-factor $\mathbf{b}_2 = \mathbf{abd}$ ensures that half of the treatment combinations of **a**, **b** and **d** occur in one week and the other half in the remaining week. The eight level combinations of factors **a**, **b**, and **d** thus define the eight tubes within a week. Finally, we define the plates using the first pseudo-factor $\mathbf{b}_1 = \mathbf{ade}$. This scenario does not change the randomization structure of the experiment. However, the allocation of the effects to the error strata is different. The first two strata, namely *Week* and *Plate*, would contain the interactions **ade**, **cf** and **ag** rather than factors **g**, **h** and their interaction **gh**. The effects contained in the *Tube* and *Unit* strata are presented in the second panel of Table 13, along with the aliasing between the main effects and two-factor interactions of the eight factors **a** to **h** and the pseudo-factors for the columns positions. The main advantage of this allocation is that all the main effects of the eight factors are now testable with a robust standard error.

5.5 Scenario 4: four distinct column positions

If there were four distinct column positions instead of eight due to the symmetry of the edge effect, the second step of the design process would change because we would need to arrange the design in four blocks instead of eight. Using the blocking scheme from Xu and Lau (2006), the two independent pseudo-factors that create the four-level factor would be $\mathbf{p}_1 = \mathbf{ab}$ and $\mathbf{p}_2 = \mathbf{acd}$. In Step 3, there would again be three regular options to add a four-level factor. All of these are shown in Table 12. In the first option, interactions **abd** and **abc** used to define factors **g** and **h**, partition the 16 tubes into four sets of four tubes so that each tube is used twice on a plate. This option is feasible, but, as Scenario 1 shows, the *Tube* and *Unit* strata then have 12 and 16 degrees of freedom, respectively. We prefer options with equal number of degrees of freedom in each of these strata.

For the second option in Table 12, the main effect of factor **h** is confounded with the two-

Table 11: Aliasing of the two pseudo-factors (\mathbf{b}_i with $i = 1, \dots, 3$) defining the weeks and the plates with interactions among the eight factors of the 2^{8-3} design for each of the three options considered.

Option	Factor	Aliasing
1	\mathbf{b}_1	$\mathbf{ade, bdg, cdh}$
	\mathbf{b}_2	$\mathbf{cf, abd, deg}$
	\mathbf{b}_3	$\mathbf{ag, be, dfh}$
2	\mathbf{b}_1	$\mathbf{bc, gh, adf}$
	\mathbf{b}_2	$\mathbf{cf, abd, deg}$
	\mathbf{b}_3	$\mathbf{bf, acd, deh}$
3	\mathbf{b}_1	$\mathbf{bc, gh, adf}$
	\mathbf{b}_2	$\mathbf{ac, eh, bdf}$
	\mathbf{b}_3	$\mathbf{ab, eg, cdf}$

factor interaction \mathbf{ac} . This results in a design of resolution III, which is undesirable. By relabelling the factors, we could use $\mathbf{h} = \mathbf{abe}$, $\mathbf{gh} = \mathbf{ac}$, and $\mathbf{g} = \mathbf{adf}$, which would create a resolution IV design. However, none of the factors \mathbf{g} or \mathbf{h} could be used to define the weeks where 8 of the 16 tubes are involved.

Finally, the third option in Table 12 is suitable for our purpose. It retains a resolution of IV, and, because $\mathbf{h} = \mathbf{abc}$, it restricts the number of tubes to 8 per week. In this scenario, both the randomization structure and the allocation of the factor effects would be similar to the design actually used. The main difference would be that the column position effects are now modelled with three instead of seven pseudo-factors, so that they are estimated more precisely.

Under this scenario, the *Week* and *Plate* strata would still contain the factors \mathbf{g} , \mathbf{h} and their interaction \mathbf{gh} . The effects contained in the *Tube* and the *Unit* strata are presented in the third panel of Table 13, along with the aliasing between the main effects and two-factor interactions of the eight factors \mathbf{a} to \mathbf{h} and the pseudo-factors for the four columns positions.

Table 12: Aliasing of the two two-level factors **g** and **h**, and their interaction **gh**, for each of the three options for adding a four-level factor to the 2^{6-1} design in case a four-level factor suffices to model column positions.

Option	Factor	Aliasing
1	g	abd, cef
	h	abc, def
	gh	cd
2	g	adf, bce
	h	ac
	gh	abe, cdf
3	g	abe, cdf
	h	abc, def
	gh	ce

Table 13: Aliasing of main effects, two-factor interactions, some three-factor interactions and pseudo-factors of the column positions, in the *Tube* and *Unit* strata, for Scenarios 1, 3, and 4.

Scenario 1: each tube is used twice on a single plate.

Scenario 3: all factors can be varied within plates.

Scenario 4: the blocking factor corresponding to the column positions has four levels.

26

<i>Tube</i>	<i>Unit</i>
a	d
b	f
c	p₃
e	p₅
p₁ + ab + ch	p₆
p₂ + ag + ce	p₇
p₄ + bg + df + eh	ad
ac + bh + eg	af
ae + cg	bd + fg
ah + bc	bf + dg
abe	cd
abg	cf
	de + fh
	dh + ef
	abd
	abf
12 df	16 df

<i>Tube</i>	<i>Unit</i>
a	e
b	g
c	h
d	p₃ + ae + bg + ch
f	p₅
p₁ + df	p₆ + de
p₂ + ad	p₇ + ef
p₄ + af	ah + ce
ab + eg	bh + cg
ac + eh	dg
bc + gh	dh
bd	fg
bf	fh
cd	aef
14 df	14 df

<i>Tube</i>	<i>Unit</i>
a	e
b	f
c	ae + bg
d	af
p₁ + ab + ch + eg	ag + be
p₂	bf
p₃	cf + dg
ac + bh	cg + df + eh
ad	de + fh
ah + bc	abf
bd	ace
cd	acf
dh + ef	acg
abd	ade
14 df	14 df

5.6 Conclusions for the alternative scenarios

Each alternative scenario had its own specificity which made it different from the original one. In Scenario 1, some degrees of freedom were moved from the *Tube* stratum to the *Unit* stratum so that, compared with the original scenario, there are more effects estimated in the *Unit* stratum and less effects in the *Tube* stratum. This makes the effects in the *Tube* stratum harder to detect. In Scenario 2, the 14 degrees of freedom from both the *Tube* and the *Unit* strata were merged so that the tube factors' effects are estimated more precisely than before and the remaining effects in that stratum are estimated less precisely than before. However the main problem in both these scenarios is that factors **g** and **h** are fully aliased with the effects of the weeks and the plates. Scenario 3 would remedy this problem because the main effects of all eight factors are located in strata with 14 degrees of freedom. They can thus all be evaluated using robust standard errors. Finally, Scenario 4 provides more information on the factorial effects, because there are only four column positions to account for. Overall, all four scenarios considered demonstrate how the methodology applied to the initial design changes with the different restrictions.

6 Conclusions

This paper showed that a small eight-factor experiment can be challenging to design when there are multiple restrictions on the randomization. While split-plotting, strip-plotting, and blocking are fairly common on their own, the combination of the three in the same design is unusual. This complicates the generation of the design as there are no catalogs or templates of designs with similar restrictions on the randomization. We developed a step-by-step approach to create a good experimental design while taking these restrictions into account. This approach allowed us to tackle the problem of including an extra multi-level factor orthogonal to the main randomization structure.

The design allowed us to detect the active effects in the protocol for microplate assays, and to estimate their contribution to the final quality of the microplate. These insights will be used to improve the protocol for the OrganoPlate[®].

The alternative scenarios in Section 5 should be helpful for practitioners to apply the step by step approach for building in multiple restrictions in the randomization in their own experiments.

Supplementary material

data_table.csv: File containing the fibrosity measurement data for all 256 chips.

alternative_scenarios.xlsx: Spreadsheet containing the final design chosen for each of the four alternative scenarios described in Section 5. (Excel file)

Declarations

No conflicts of interests or competing interests are declared.

Acknowledgements

This research was financially supported by the Fonds Wetenschappelijk Onderzoek (FWO, Flanders, Belgium).

References

- Bingham, D., Sitter, R., Kelly, E., Moore, L., and Olivas, J. D. (2008). Factorial Designs with Multiple Levels of Randomization. *Statistica Sinica*, 18(2):493–513.
- Bingham, D. and Sitter, R. R. (1999). Minimum-Aberration Two-Level Fractional Factorial Split-Plot Designs. *Technometrics*, 41(1):62–70.
- Bingham, D. R., Schoen, E. D., and Sitter, R. R. (2004). Designing fractional factorial split-plot experiments with few whole-plot factors. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 53(2):325–339.
- Bingham, D. R. and Sitter, R. R. (2001). Design Issues in Fractional Factorial Split-Plot Experiments. *Journal of Quality Technology*, 33(1):2–15.
- Box, G. E., Hunter, W. H., and Hunter, S. (1978). *Statistics for Experimenters*, volume 664. John Wiley and sons New York.
- Cheng, C.-S. and Tsai, P.-W. (2011). Multistratum fractional factorial designs. *Statistica Sinica*, 21(3):1001–1021.

- Eendebak, P. and Vazquez, A. (2019). OAPackage: A Python package for generation and analysis of orthogonal arrays, optimal designs and conference designs. *Journal of Open Source Software*, 4(34):1097.
- Goos, P., Langhans, I., and Vandebroek, M. (2006). Practical Inference from Industrial Split-Plot Designs. *Journal of Quality Technology*, 38(2):162–179.
- Huang, P., Chen, D., and Voelkel, J. O. (1998). Minimum-Aberration Two-Level Split-Plot Designs. *Technometrics*, 40(4):314–326.
- Jones, B. and Goos, P. (2009). D-optimal design of split-split-plot experiments. *Biometrika*, 96(1):67–82.
- Leung, C. M., de Haan, P., Ronaldson-Bouchard, K., Kim, G.-A., Ko, J., Rho, H. S., Chen, Z., Habibovic, P., Jeon, N. L., Takayama, S., Shuler, M. L., Vunjak-Novakovic, G., Frey, O., Verpoorte, E., and Toh, Y.-C. (2022). A guide to the organ-on-a-chip. *Nature Reviews Methods Primers*, 2(1):1–29.
- Mee, R. W. (2009). *A Comprehensive Guide to Factorial Two-Level Experimentation*. Springer, Dordrecht.
- Mee, R. W. and Bates, R. L. (1998). Split-Lot Designs: Experiments for Multistage Batch Processes. *Technometrics*, 40(2):127–140.
- Miller, A. (1997). Strip-Plot Configurations of Fractional Factorials. *Technometrics*, 39(2):153–161.
- Montgomery, D. C. (2019). *Design and Analysis of Experiments, 9th Edition*. John Wiley & Sons.
- Mostaço-Guidolin, L. B., Ko, A. C.-T., Wang, F., Xiang, B., Hewko, M., Tian, G., Major, A., Shiomi, M., and Sowa, M. G. (2013). Collagen morphology and texture analysis: From statistics to classification. *Scientific Reports*, 3:2190.
- Roeder, B. A., Kokini, K., Sturgis, J. E., Robinson, J. P., and Voytik-Harbin, S. L. (2002). Tensile Mechanical Properties of Three-Dimensional Type I Collagen Extracellular Matrices With Varied Microstructure. *Journal of Biomechanical Engineering*, 124(2):214–222.

- Sartono, B., Goos, P., and Schoen, E. (2015a). Blocking Orthogonal Designs With Mixed Integer Linear Programming. *Technometrics*, 57(3):428–439.
- Sartono, B., Goos, P., and Schoen, E. (2015b). Constructing General Orthogonal Fractional Factorial Split-Plot Designs. *Technometrics*, 57(4):488–502.
- Schoen, E. D. (1999). Designing fractional two-level experiments with nested error structures. *Journal of Applied Statistics*, 26(4):495–508.
- Schoen, E. D. and Kaul, E. A. A. (2000). Three Robust Scale Estimators to Judge Unreplicated Experiments. *Journal of Quality Technology*, 32(3):276–283.
- Sung, K. E., Su, G., Pehlke, C., Trier, S. M., Eliceiri, K. W., Keely, P. J., Friedl, A., and Beebe, D. J. (2009). Control of 3-dimensional collagen matrix polymerization for reproducible human mammary fibroblast cell culture in microfluidic devices. *Biomaterials*, 30(27):4833–4841.
- Theocharis, A. D., Skandalis, S. S., Gialeli, C., and Karamanos, N. K. (2016). Extracellular matrix structure. *Advanced Drug Delivery Reviews*, 97:4–27.
- van Duinen, V., Trietsch, S. J., Joore, J., Vulto, P., and Hankemeier, T. (2015). Microfluidic 3D cell culture: From tools to tissue models. *Current Opinion in Biotechnology*, 35:118–126.
- Vivacqua, C. A. and Bisgaard, S. (2004). Strip-Block Experiments for Process Improvement and Robustness. *Quality Engineering*, 16(3):495–500.
- Vo-Thanh, N., Schoen, E. D., and Goos, P. (2020). Row-column arrangements of regular and nonregular fractional factorial two-level designs. *Journal of Quality Technology*, 52(3):304–322.
- Wu, C. F. J. and Hamada, M. (2009). *Experiments: Planning, Analysis, and Optimization*. Wiley Series in Probability and Statistics. Wiley, Hoboken, N.J, 2nd ed edition.
- Xu, H. and Lau, S. (2006). Minimum aberration blocking schemes for two- and three-level fractional factorial designs. *Journal of Statistical Planning and Inference*, 136(11):4088–4118.