POINTS OF SIGNIFICANCE

Split plot design

When some factors are harder to vary than others, a split plot design can be efficient.

We have already seen that varying two factors simultaneously provides an effective experimental design for exploring the main (average) effects and interactions of the factors[1]. However, in practice, some factors may be more difficult to vary than others at the level of experimental units. For example, drugs given orally are difficult to administer to individual tissues, but observations on different tissues may be done by biopsy or autopsy. When the factors can be nested, it is more efficient to apply a difficult-to-change factor to the units at the top of the hierarchy and then apply the easier-to-change factor to a nested unit. This is called a split plot design.

The term "split plot" derives from agriculture, where fields may be split into plots and subplots. It is instructive to review completely randomized design (CRD) and randomized complete block design (RCBD)[2] and show how these relate to split plot design. Suppose we are studying the effect of irrigation amount and fertilizer type on crop yield. We have access to eight fields, which can be treated independently and without proximity effects (**Fig. 1a**). If applying irrigation and fertilizer is equally easy, we can use a complete 2 × 2 factorial design and assign levels of both factors randomly to fields in a balanced way (each combination of factor levels is equally represented).

If our land is divided into two large fields that may differ in some way, we can use the field as a blocking factor (**Fig. 1b**). Within each block, we again perform a complete 2 × 2 factorial design: irrigation and fertilizer are assigned to each of the four smaller fields within the large field, leading to an RCBD with field as the block. Each combination of irrigation and fertilizer is balanced within the large field.

So far, we have not considered whether managing levels of irrigation and fertilizer require the same effort. If varying irrigation on a small scale is difficult, it makes more sense to irrigate larger areas of land than in **Figure 1a** and then vary the fertilizer accordingly to maintain a balanced design. If our land is divided into four fields (whole plots), each of which can be split into two subplots (**Fig. 1c**), we would assign irrigation to whole plots using CRD. Within a whole plot, fertilizer would be distributed across subplots using RCBD,

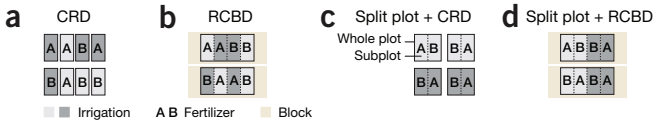

Figure 1 | Split plot design examples from agriculture. (**a**) In CRD, levels of irrigation and fertilizer are assigned to plots of land (experimental units) in a random and balanced fashion. (**b**) In RCBD, similar experimental units are grouped (for example, by field) into blocks and treatments are distributed in a CRD fashion within the block. (**c**) If irrigation is more difficult to vary on a small scale and fields are large enough to be split, a split plot design becomes appropriate. Irrigation levels are assigned to whole plots by CRD and fertilizer is assigned to subplots using RCBD (irrigation is the block). (**d**) If the fields are large enough, they can be used as blocks for two levels of irrigation. Each field is composed of two whole plots, each composed of two subplots. Irrigation is assigned to whole plots using RCBD (blocked by field) and fertilizer assigned to subplots using RCBD (blocked by irrigation).

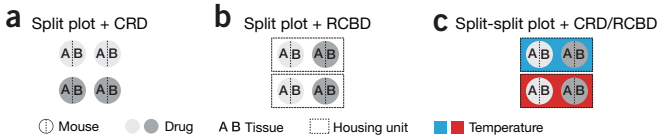

Figure 2 | In biological experiments using split plot designs, whole plot experimental units can be individual animals or groups. (**a**) A two-factor, split plot animal experiment design. The whole plot is represented by a mouse assigned to drug, and tissues represent subplots. (**b**) Biological variability coming from nuisance factors, such as weight, can be addressed by blocking the whole plot factor, whose levels are now sampled using RCBD. (**c**) With three factors, the design is split-split plot. The housing unit is the whole plot experimental unit, each subject to a different temperature. Temperature is assigned to housing using CRD. Within each whole plot, the design shown in **b** is performed. Drug and tissue are subplot and sub-subplot units. Replication is done by increasing the number of housing units.

randomly and balanced within whole plots with a given irrigation level. Irrigation is the whole plot factor and fertilizer is the subplot factor. It is important to note that all split plot experiments include at least one RCBD subexperiment, with the whole plot factor acting as a block.

Assigning levels of irrigation to fields at random neglects any heterogeneity among the fields. For example, if the land is divided into two large fields (**Fig. 1b**), it is best to consider each as a block. Within each block, we consider half of the field as a whole plot and irrigate using RCBD (**Fig. 1d**). As before, the fertilizer is assigned to subplots using RCBD. The designs in **Figure 1c** and **Figure 1d** vary only in how the whole plot factor levels are assigned: by CRD or RCBD.

Because split plot designs are based on RCBD, the two can be easily confused. For example, why is **Figure 1b** not considered a split plot design with field index being the whole plot factor? The answer involves whether we are interested in specific levels of the factor or are using it for blocking purposes. In **Figure 1b**, the field is a blocking factor because it is used to control the variability of the plots, not as a systematic effect. We use these two fields to generalize to all fields. In **Figure 1c**, irrigation is a whole plot factor and not a blocking factor because we are studying the specific levels of irrigation.

The terms "whole plot" and "subplot" translate naturally from agricultural to biological context, where split plot designs are common. Many factors, such as diet or housing conditions, are more easily applied to large groups of experimental subjects, making them suitable at the whole plot level. In other experiments, factors that are sampled hierarchically or from the same individual (tissue, cell or time points) can act as subplot factors. **Figure 2** illustrates split plot designs in a biological context.

Suppose that we wish to determine the *in vivo* effect of a drug on gene expression in two tissues. We assign mice to one of two drug treatments using CRD. The mouse is the whole plot experimental unit and the drug is the whole plot factor. Both tissues are sampled from each mouse. The tissue is the subplot factor and each mouse acts as a block for the tissue subplot factor; this is the RCBD component (**Fig. 2a**). The mouse itself can be considered a random factor used to sample biological variability and increase the external validity of the experiment. If we suspect environmental variability, we can group the mice by their housing unit (**Fig. 2b**), just as we did whole plots by field (**Fig. 1d**). The housing unit is now a blocking factor for the drug, which is applied to mice using RCBD. Other ways to group mice might be by weight, familial relationship or genotype.

Sensitivity in detecting effects of the subplot factor as well as interactions is generally greater than for a corresponding completely

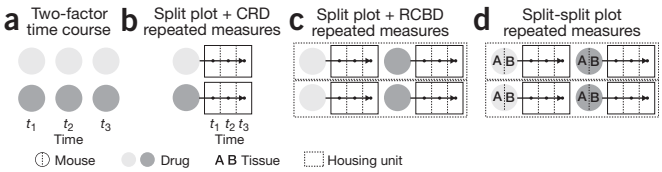

Figure 3 | The split plot design with CRD is commonly applied to a repeated measures time course design. (**a**) Basic time course design, in which time is one of the factors. Each measurement uses a different mouse. (**b**) In a repeated measures design, mice are followed longitudinally. Drug is assigned to mice using CRD. Time is the subplot factor. (**c**) Drug is blocked by housing. (**d**) A three-factor, repeated measures split-split plot design, now including tissue. Tissue is subplot and time is sub-subplot.

Table 1 | Split plot ANOVA table for two-factor split plot designs

	CRD			RCBD		
	d.f.	MS	*F*-ratio	d.f.	MS	*F*-ratio
Block, *bl*				n'	MS_{bl}	MS_{bl}/MS_{wp}
A	a'	MS_A	MS_A/MS_{wp}	a'	MS_A	MS_A/MS_{wp}
Error wp	an'	MS_{wp}		$n'a'$	MS_{wp}	
B	b'	MS_B	MS_B/MS_{sp}	b'	MS_B	MS_B/MS_{sp}
$A \times B$	$a'b'$	$MS_{A\times B}$	MS_{AB}/MS_{sp}	$a'b'$	$MS_{A\times B}$	$MS_{A\times B}/MS_{sp}$
Error sp	$ab'n'$	MS_{sp}		$ab'n'$	MS_{sp}	
Total	$abn - 1$			$abn - 1$		

Split plot ANOVA table for two factor split plot designs using CRD (**Fig. 1c**) and RCBD (**Fig. 1d**) with a levels of whole plot factor A and b levels of subplot factor B. For CRD n is measurements per subplot and for RCBD n is number of blocks. Whole plot and subplot errors are indicated by wp and sp subscripts, respectively. For RCBD, interaction between blocking factor *bl* and B is usually included in the subplot error term. $a' = a - 1$, $b' = b - 1$, $n' = n - 1$. d.f., degrees of freedom; *F*-ratio, test statistic for *F* test.

randomized factorial design in which only one tissue is measured in each mouse. This is because tissue comparisons are within mouse. However, because comparing the whole plot factor (drug) is done between subjects, the sensitivity for the whole plot factor is similar to that of a completely randomized design. Applying blocking at the whole plot level, such as housing (**Fig. 2b**), can improve sensitivity for the whole plot factor similarly to using a RCBD. Compared to a split plot design, the completely randomized design is both more expensive (twice as many mice are required) and less efficient (mouse variability will not cancel, and thus the tissue and interaction effects will include mouse-to-mouse variability).

The experimental unit at the whole plot level does not have to correspond to an individual. It can be one level above the individual in the hierarchy, such as a group or enclosure. For example, suppose we are interested in adding temperature as one of the factors to the study in **Figure 2b**. Since it is more practical to control the temperature of the housing unit than of individual mice, we use cage as the whole plot (**Fig. 2c**). Temperature is the whole plot factor and cage is the experimental unit at the whole plot level. As in **Figure 2a**, we use CRD to assign the whole plot factor (temperature) levels to whole plots (cages). Mice are now experimental units at the subplot level and the drug is now a subplot factor. Because we have three layers in the hierarchy of factors, tissue is at the sub-subplot level and the design is split-split plot. In **Figure 2b**, the cage is a block used to control variability because the effects of housing are not of specific interest to us. By contrast, in **Figure 2c**, specific levels of the temperature factor are of interest so it is part of the plot factor hierarchy.

Care must be taken to not mistake a split plot design for CRD. For example, an inadvertent split plot[3] can result if some factor levels are not changed between experiments. If the analysis treats all experiments as independent, then we can expect mistakes in conclusions about the significance of effects.

With two factors, more complicated designs are also possible. For example, we might expose the whole mouse to a drug (factor *A*) *in vivo* and then expose two liver samples to different *in vitro* treatments (factor *B*). In this case, the two liver samples from the same mouse form a block, which is nested in mouse[4].

The split plot CRD design (**Fig. 2a**) is commonly used as the basis for a repeated measures design, which is a type of time course design. The most basic time course includes time as one of the factors in a two-factor design. In a completely randomized time course experiment, different mice are used at each of the measurement times t_1, t_2 and t_3 after initial treatment (**Fig. 3a**). If the same mouse is used at each time and the mice are assigned at random to the levels of a (time-invariant) factor, the design becomes a repeated measures design (**Fig. 3b**) because the measurements are nested within mouse. The time of measurement is the subplot factor. The corresponding repeated measures of the design that uses housing as a block in **Figure 2b** is shown in **Figure 3c**. As before, housing is the block and drug is the whole plot factor, but now time is the subplot factor. If we include tissue type, the design becomes a split-split plot, with tissue being subplot and time sub-subplot (**Fig. 3d**).

Split plot designs are analyzed using ANOVA. Because comparisons at the whole plot level have different variability than those at the subplot level, the ANOVA table contains two sources of error, MS_{wp} and MS_{sp}, the mean square associated with whole plots and subplots, respectively (**Table 1**). This difference occurs because the subplot factor is always compared within a block, while the whole plot factor is compared between the whole plots. For example, in **Figure 2a**, variation between mice cancels out when comparing tissues but not when comparing drugs. Analogously to a two-factor ANOVA[1], we calculate the sums of squares and mean squares in a split plot ANOVA. For example, in a split plot with RCBD, given n blocks of blocking factor *bl* (**Table 1**) at the whole plot level and a and b levels of whole plot factor A and subplot factor B, $MS_{bl} = SS_{bl}/(n-1)$, where SS_{bl} is the sum of squared deviations of the average across each block relative to the grand mean times the number of measurements contributing to each average ($a \times b$). Similarly, SS_A uses the average across levels of A and the multiple is $n \times b$. The analysis at the whole plot level is essentially the same as in a one-way ANOVA with blocking: the subplot values are considered subsamples. The associated MS_{sp} is usually lower than in a factorial design, which improves the sensitivity in detecting $A \times B$ interactions.

Split plot designs are helpful when it is difficult to vary all factors simultaneously, and, if factors that require more time or resources can be identified, split plot designs can offer cost savings. This type of design is also useful for cases when the investigator wishes to expand the scope of the experiment: a factor can be added at the whole plot level without sacrificing sensitivity in the subplot factor.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Naomi Altman & Martin Krzywinski

1. Krzywinski, M., Altman, N. & Blainey, P. *Nat. Methods* **11**, 1187–1188 (2014).
2. Krzywinski, M. & Altman, N. *Nat. Methods* **11**, 699–700 (2014).
3. Ganju, J. & Lucas, J.M. *J. Stat. Plan. Infer.* **81**, 129–140 (1999).
4. Krzywinski, M., Altman, N. & Blainey, P. *Nat. Methods* **11**, 977–978 (2014).

Naomi Altman is a Professor of Statistics at The Pennsylvania State University. Martin Krzywinski is a staff scientist at Canada's Michael Smith Genome Sciences Centre.